



Quasi-Stationary Distributions for the Voter Model on Complete Bipartite Graphs

Iddo Ben-Ari, Hugo Panzo, Philip Speegle and R. Oliver VandenBerg

Department of Mathematics
University of Connecticut
Storrs, CT 06269, USA
E-mail address: iddo.ben-ari@uconn.edu
URL: <http://iddo.ben-ari.uconn.edu/wordpress/>

Faculties of Electrical Engineering and Mathematics
Technion – Israel Institute of Technology
Haifa 32000, Israel
E-mail address: panzo@campus.technion.ac.il
URL: <https://sites.google.com/view/hugopanzo/>

University of Alabama
Tuscaloosa, AL 35487, USA
E-mail address: pspeegle@crimson.ua.edu

University of Minnesota Minneapolis, MN 55455, USA
E-mail address: vand1716@umn.edu

Abstract. We consider the discrete-time voter model on complete bipartite graphs and study the quasi-stationary distribution (QSD) for the model as the size of one of the partitions tends to infinity while the other partition remains fixed. We show that the QSDs converge weakly to a nontrivial limit which features a consensus with the exception of a random number of dissenting vertices in the “large” partition. Moreover, we explicitly calculate the law of the number of dissenters and show that it follows the heavy-tailed Sibuya distribution with parameter depending on the size of the “small” partition. Our results rely on a discrete-time analogue of the well-known duality between the continuous-time voter model and coalescing random walks which we develop in the paper.

Received by the editors April 21th, 2020; accepted December 30th, 2020.

2010 Mathematics Subject Classification. 60F99, 60J10, 60K35, 82C22.

Key words and phrases. Quasi-Stationary Distributions, Voter Model, Coalescing Random Walks, Complete Bipartite Graphs.

Research performed during [Markov Chains REU](#), partially supported by NSA grant H98230-19-1-0022 to Iddo Ben-Ari. Hugo Panzo was supported at the Technion by a Zuckerman Fellowship.

1. Introduction and Main Results

The *voter model* is an interacting particle system which models the evolution of opinions in a population of voters. In the classical version, the voters are vertices of \mathbb{Z}^d which can hold the opinion “0” or “1” and the model evolves in continuous-time by having each vertex change its opinion at a rate proportional to the number of dissenting neighbors, see [Liggett \(1985\)](#). With these dynamics, it is clear that *consensus* is an absorbing state and early investigations of the model were interested in the time required to reach consensus and characterization of nontrivial invariant measures. Besides regular lattices, the voter model has also been studied in discrete-time on heterogeneous graphs where it was shown in [Sood and Redner \(2005\)](#) to display markedly different behavior, see also [Cooper and Rivera \(2016\)](#) where it is referred to as the *asynchronous pull model*. Additionally, there are further extensions of the model which more closely approximate reality, see [Redner \(2019\)](#) for a recent survey.

In this paper, we study the discrete-time voter model on the *complete bipartite graphs* $K_{n,m}$. These are heterogeneous graphs whose vertex set can be partitioned into two groups, a “large” group L of size n and a “small” group S of size m , where each vertex of L is connected to all of the vertices of S and vice versa, and there are no connections between vertices in the same group. While the time required to reach consensus in the voter model on $K_{n,m}$ and its behavior on the way to consensus have already been studied in [Sood and Redner \(2005\)](#), we investigate what happens when consensus is conditioned to never occur. More specifically, we study the *quasi-stationary distribution* (QSD) for the voter model on $K_{n,m}$.

Loosely speaking, a QSD for a Markov chain with absorbing states is a stationary distribution for the chain conditioned on nonabsorption, see [Section 2](#) for a precise definition. In the case of the voter model on $K_{n,m}$, one might ask what the distribution of opinions typically looks like if consensus hasn’t been reached after a long time. Is the lack of consensus due to a small minority of dissenters or are the opinions relatively balanced? If $n \gg m$, does the distribution of opinions on L differ qualitatively from that of S ? In order to give a concise answer to these questions, we fix m and find the limit in distribution of the sequence of QSDs as $n \rightarrow \infty$. We point to the case of the voter model on the complete graphs K_n in [Example 4.3](#) as evidence that even the existence of a limiting distribution is not obvious.

Before stating our main results, we recall from [Devroye \(1993\)](#); [Johnson et al. \(2005\)](#) that the *Sibuya distribution* with parameter $\gamma \in (0, 1]$ is a probability distribution on $\mathbb{Z}_+ = \{1, 2, \dots\}$ with probability mass function f_γ and probability generating function ϕ_γ given by

$$f_\gamma(k) = \frac{\gamma}{k!} \prod_{j=1}^{k-1} (j - \gamma), \quad k \in \mathbb{Z}_+, \quad (1.1)$$

$$\phi_\gamma(z) = 1 - (1 - z)^\gamma, \quad |z| \leq 1.$$

When $\gamma \in (0, 1)$, the Sibuya distribution is heavy tailed and we note from [Christoph and Schreiber \(2000\)](#) that in this case f_γ decays according to a power law with

$$f_\gamma(k) \sim \frac{1}{\pi} \sin(\gamma\pi) \Gamma(1 + \gamma) \frac{1}{k^{\gamma+1}} \quad \text{as } k \rightarrow \infty.$$

We denote this probability distribution by $\text{Sib}(\gamma)$. See [Kozubowski and Podgórski \(2018\)](#); [Letac \(2019\)](#); [Huillet \(2020\)](#) for some recent applications of the Sibuya distribution. We also use $\text{Bern}(p)$ to denote a Bernoulli random variable which takes the values 1 and 0 with probability p and $1 - p$, respectively.

Theorem 1.1. *Let $C \sim \text{Bern}(1/2)$ and $D \sim \text{Sib}(\gamma_m)$ be independent, with*

$$\gamma_m = 2 \left(1 - \sqrt{1 - \frac{1}{2m}} \right).$$

Then the distribution of opinions under the QSD for the voter model on $K_{n,m}$ as $n \rightarrow \infty$ converges weakly to the following:

- (1) *All vertices of S have opinion C .*
- (2) *All but D vertices in L have opinion C .*

Corollary 1.2. *The distribution of the number of disagreements along edges in $K_{n,m}$ under the QSD tends to mD , where $D \sim \text{Sib}(\gamma_m)$.*

The proof of [Theorem 1.1](#) will be given at the end of [Section 5](#). The rest of the paper is organized as follows. In [Section 2](#) we recall some important facts from the theory of quasi-stationary distributions for finite state Markov chains. We describe the voter model on general finite graphs in [Section 3](#). In [Section 4](#), we develop for general finite graphs a discrete-time analogue of the well-known duality between the continuous-time voter model and coalescing random walks. We use this duality in [Section 5](#) to compute explicitly the geometric tail of the time to consensus in the voter model on $K_{n,m}$ and subsequently prove our main results.

2. Quasi-Stationary Distributions

Here we give a quick summary of the theory of quasi-stationary distributions for finite state Markov chains, see [Collet et al. \(2013\)](#) for more details. To this end, suppose that $\mathbf{Y} = (Y_t : t \in \mathbb{Z}_+)$ is a Markov chain on a finite state space Ω with transition function \bar{S} . Recall that a state i is absorbing if

$$P_i(Y_1 = i) = 1.$$

We will assume

- (1) Δ , the set of all absorbing states, is nontrivial: it is not empty and its complement is not empty; and
- (2) Δ is accessible from every state. That is, for every state i , there exists $t \in \mathbb{Z}_+$ such that

$$P_i(Y_t \in \Delta) > 0.$$

Letting

$$\tau = \inf\{t \in \mathbb{Z}_+ : Y_t \in \Delta\},$$

it follows from our assumptions that $\tau < \infty$ almost surely under any initial distribution. Let S denote the substochastic transition function obtained by restricting \bar{S} to the complement of Δ in $\bar{\Omega}$. We denote this complement by Ω . In other words, S is the principal submatrix obtained from \bar{S} by removing all rows and columns corresponding to states in Δ . Let μ be an initial distribution for \mathbf{Y} whose support is contained in Ω . As there is no risk of ambiguity, we will abuse notation and also denote its restriction to Ω by μ . We have

$$P_\mu(Y_t = j, \tau > t) = \mu S^t(j).$$

In particular,

$$P_\mu(\tau > t) = \mu S^t \mathbf{1} = \sum_{j \in \Omega} \mu S^t(j).$$

Furthermore,

$$\lim_{t \rightarrow \infty} P_\mu(\tau > t)^{1/t}$$

exists. Moreover, if one considers the restriction of S to the states accessible from the support of μ , then the limit above coincides with spectral radius of the resulting principal submatrix.

A probability distribution ν on $\bar{\Omega}$ is called a quasi-stationary distribution (QSD) if

$$P_\nu(Y_t = j | \tau > t) = \nu(j), \quad j \in \Omega. \quad (2.1)$$

for all $t \in \mathbb{Z}_+$. Clearly, ν is then supported on Ω , and therefore will be viewed as a probability measure on Ω . Furthermore, (2.1) can be rewritten as

$$\frac{\nu S^t e_j}{\nu S^t \mathbf{1}} = \frac{\nu S^t(j)}{C_\nu(t)} = \nu(j),$$

where $e_j(k) = 1$ if $k = j$ and 0 otherwise, and $C_\nu(t) = \sum_{k \in \Omega} \nu S^t(k)$. Equivalently,

$$\nu S^t = C_\nu(t) \nu.$$

Plugging $t = 1$ into the above equation leads to the following well-known result.

Proposition 2.1.

- (1) *A probability vector ν on Ω is a QSD if and only if ν is a left eigenvector for S with a strictly positive eigenvalue λ . That is, if*

$$\nu S = \lambda \nu,$$

with λ being the spectral radius of S restricted to the linear space spanned by the indicators of the support of ν .

- (2) *If ν is a QSD, then the distribution of τ under P_ν is geometric with parameter $1 - \lambda$.*

As for existence and uniqueness of the QSD, as well as for convergence of the conditioned chain to the QSD, the Perron-Frobenius theorem (Horn and Johnson, 2013, Theorem 8.4.4) and the limit theorem for primitive matrices (Horn and Johnson, 2013, Theorems 8.5.1 and 8.5.3) yield the following result.

Theorem 2.2.

- (1) *If S is irreducible then there exists a unique QSD.*
 (2) *If S is irreducible and aperiodic (also known as primitive), then for any initial distribution μ ,*

$$\lim_{t \rightarrow \infty} P_\mu(Y_t \in \cdot | \tau > t) = \nu,$$

where ν is the unique QSD.

3. Voter Model on a Finite Graph

Let $G = (V, E)$ be a finite, connected graph with vertex set V and edge set E . A *coloring* of G is a function f from the V to the nonnegative integers. The number assigned to vertex $v \in V$ by f , namely $f(v)$, is the color of v or the opinion of v . We will use the terms “color” and “opinion” interchangeably. The model we will study in this paper is a discrete-time version of the voter model on G . This is a Markov chain on the colorings of G which evolves during each unit of time as follows. A vertex is picked uniformly, this vertex samples a neighbor uniformly, then the former vertex adopts the opinion of the latter. We say that two neighboring vertices are in agreement if their colors are identical. Otherwise, they are in disagreement. States with no disagreements are *consensus states*. Note that the set of absorbing states is the set of consensus states.

Since G is connected, the model reaches a consensus with probability 1. This is because by construction there is positive probability to reach a consensus from any state within a certain number of steps. Since the graph is finite, this implies that the time to reach a consensus is dominated by a geometric random variable.

We write $\eta = (\eta_t : t \in \mathbb{Z}_+)$ for the discrete-time voter model on G , with $\eta_t(v)$ representing the color (or opinion) of vertex v at time t . A state of the system is therefore a coloring $\eta : V \rightarrow \mathbb{Z}_+$ of G . The probability of a transition from η to η' is positive if and only if there exists $(v, v') \in V \times V$ such that

- (1) $\{v, v'\} \in E$;
- (2) $\eta'(v) = \eta(v')$; and
- (3) $\eta'(u) = \eta(u)$ for all $u \neq v$.

Now if the pair η and η' satisfy the above conditions, then a transition is obtained by first uniformly sampling the vertex v among all those for which a matching v' exists, and then adopting the opinion of v' . This leads to the following transition function:

$$p(\eta, \eta') = \frac{1}{|V|} \sum_{v \in V} \sum_{v' \sim v} \frac{\mathbf{1}_{\eta'(v)}(\eta(v'))}{\deg(v)} \prod_{u \neq v} \mathbf{1}_{\eta'(u)}(\eta(u)). \tag{3.1}$$

All other transitions are not allowed.

Bringing the discussion from Section 2 into the context of the voter model on G , the absorption time τ is the time of first consensus, that is

$$\tau = \inf\{t \in \mathbb{Z}_+ : \eta_t(v) = \eta_t(v') \text{ for all } v, v' \in V\}. \tag{3.2}$$

Using the subscript V to designate voter model, we write $\lambda_V(G, \mu)$ for the spectral radius associated with the initial distribution μ ,

$$\lambda_V(G, \mu) = \lim_{t \rightarrow \infty} P_\mu(\tau > t)^{1/t}. \tag{3.3}$$

4. Coalescing Random Walks and Time to Absorption

The key to our analysis is based on the duality between the continuous-time voter model and coalescing random walks. This duality is well known and the reader is directed to [Durrett \(1988\)](#); [Aldous and Fill \(2002\)](#); [Oliveira \(2012\)](#) and references therein for an exposition. In this section, we develop a discrete-time analogue of this duality. For the purposes of presentation and in order to make it useful for future work, we will consider the voter model on a general finite connected graph

$G = (V, E)$. As far as the authors know, such a general treatment of this duality in discrete-time hasn't appeared in the literature before.

The first step towards finding a QSD is identifying $\lambda_V(G, \mu)$. This is a non-trivial problem in general. The connection with coalescing random walks that will be described in this section simplifies the analysis of the time until consensus by identifying the distribution of τ with the distribution of the time until two random walks on the graph first meet. The idea is to describe the “flow” or propagation of opinions back in time, tracing whose opinion each vertex inherited from previous steps, going all the way back to time zero. Following the origin of an opinion of a given vertex backward in time is a random walk on the graph, and the family of resulting random walks, indexed by the vertices of the graph, is a process known as coalescing random walks.

In passing from time $t - 1$ to t in the voter model, we first uniformly select a vertex v , then uniformly select a neighbor u and assign $\eta_t(v) = \eta_{t-1}(u)$. For each $t \in \{1, 2, \dots\}$, the sampling of vertex v and its neighbor is independent of and identically distributed as the respective sampling for other times. Furthermore, this sampling is also independent of the actual opinions up to time $t - 1$. Fix some time $T \in \mathbb{Z}_+$. We will construct a random directed graph \mathcal{G}_T on $V \times \{0, \dots, T\}$ which would represent the same process, but with time reversed. Why reverse time? Because we eventually want to trace whose original opinion (opinion at time 0) each of the vertices holds at time T . Below we denote a directed arrow from $(u, n) \in V \times \{0, \dots, T\}$ to $(u', n') \in V \times \{0, \dots, T\}$ by $(u, n) \rightarrow (u', n')$.

We now describe the construction. This is done in three steps.

1. *Adopting others' opinions.* If at time $t = T - n$ the vertex v is selected to adopt the opinion of vertex u at time $t - 1 = T - (n + 1)$, we will draw a directed arrow from (v, n) to $(u, n + 1)$. We begin from $t = T$, and end at $t = 1$. This describes which vertex got whose opinion and when. See Figure 4.1 for an illustration of this procedure on a star graph with $T = 6$. Note that the t -time of the voter model runs from left to right while the n -time of the random directed graph runs from right to left.

2. *Keeping one's opinion.* Since all vertices but one keep their opinions from time $t - 1$ to time t , we add arrows to represent this as well. To do that let $h(n)$ denote the unique vertex (v, n) with an arrow to some $(u, n + 1)$, as obtained in Step 1. For all $v \in V - \{h(n)\}$, we draw an arrow from (v, n) to $(v, n + 1)$. See Figure 4.2 for an illustration of this stage.

3. *Removing useless arrows.* At the end of Step 2, for every $v \in V$ there exists a unique path from $(v, 0)$ to (\cdot, T) . That is, given $v \in V$, the unique path is a sequence $(v_0 = v, 0) \rightarrow (v_1, 1) \rightarrow \dots \rightarrow (v_T, T)$ where v_{n+1} is the unique vertex $v \in V$ satisfying $(v_n, n) \rightarrow (v, n + 1)$. Since the path is determined by the choice of v and T , we denote it by $\mathbf{X}^T(v) = (X_n^T(v) : n = 0, \dots, T)$, where $X_n^T(v) = v_n$. An arrow is useless if no path $(v, 0) \rightarrow \dots \rightarrow (v_T, T)$ uses it. We will remove those from our graph \mathcal{G}_T . With this our construction is complete. This stage is illustrated in Figure 4.3.

Now that we have completed the construction of the random graphs, we do some analysis. Each of the random graphs is determined by Step 1 while Steps 2 and 3 are deterministic functions of it. From our construction, for each $v \in V$, $\mathbf{X}^T(v)$

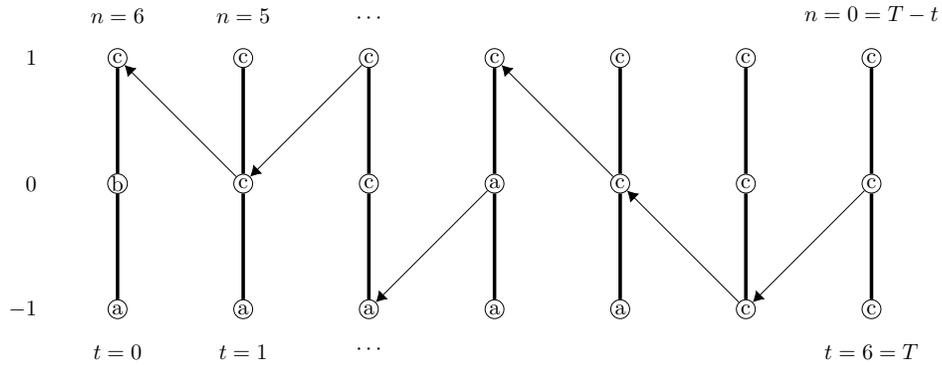


FIGURE 4.1. Realization of the voter model on a simple star graph as a function of time. Vertices are the circles, labeled $0, \pm 1$ and edges are the vertical line segments. Opinion at each vertex is the letter inscribed in the circle. The time for the voter model appears at the bottom, while reversed time, the second component in the vertices of the random graph \mathcal{G}_T , appears at the top. An arrow from (v, n) to vertex $(u, n+1)$ represents vertex v adopting at time $t = T - n$ the opinion of vertex u at time $t - 1 = T - (n + 1)$.

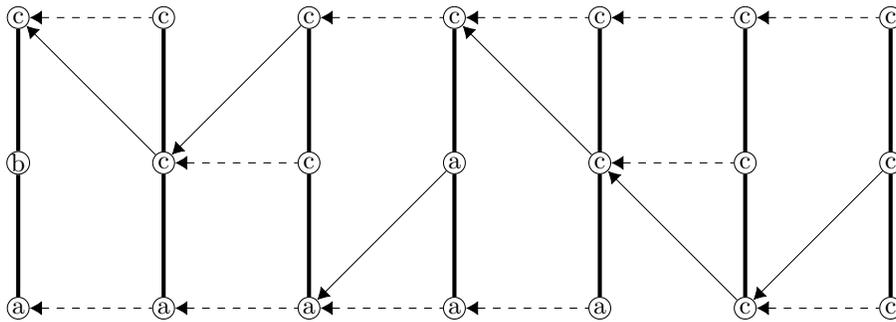


FIGURE 4.2. Same realization as Figure 4.1, with dashed horizontal arrows representing vertices keeping their opinion.

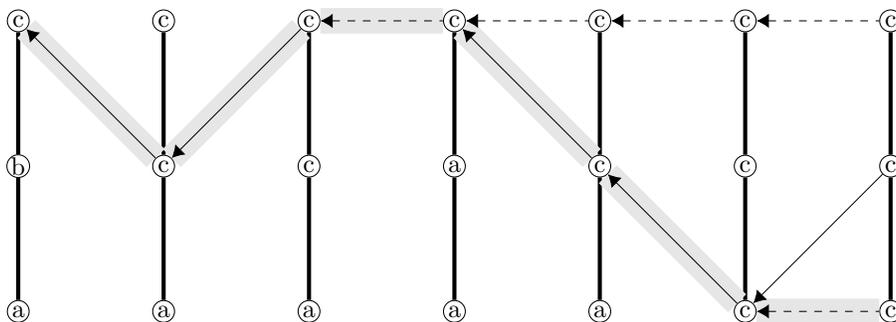


FIGURE 4.3. The random graph \mathcal{G}_6 , obtained after removing useless arrows from Figure 4.2. The path of the random walk $\mathbf{X}^6(-1)$ has a shadow.

is a Markov chain on G whose initial distribution is the point-mass at v and with transition function

$$p(u, u) = \frac{|V| - 1}{|V|}, \quad p(u, w) = \frac{1}{|V|} \frac{1}{\deg(u)}, \quad \{u, w\} \in E.$$

Equivalently, $\mathbf{X}^T(v)$ is a lazy random walk on G which starts at v and has probability $1 - \frac{1}{|V|}$ of staying put at each step.

Next we define a system of coupled random walks $\mathbf{Y} = (Y(v) : v \in V)$ whose distribution up to time T coincides with that of \mathbf{X}^T . We begin by setting $Y_0(v) = v$ for each $v \in V$. Assuming that $Y_s(\cdot)$ is defined for $s \leq t$, we uniformly and independently sample $\mathbf{v} \in V$ and a neighbor \mathbf{u} of \mathbf{v} . If $Y_t(\cdot) \neq \mathbf{v}$, then we set $Y_{t+1}(\cdot) = Y_t(\cdot)$. Otherwise we set $Y_{t+1}(\cdot) = \mathbf{u}$. We will refer to \mathbf{Y} as the *coalescing random walks*. Note that the joint distribution of the walks $\mathbf{X}^T(v), v \in V$ coincides with the restriction of the coalescing random walks to the time interval $\{0, \dots, T\}$.

For any two distinct vertices $v, v' \in V$, let

$$\sigma_{v,v'}^T = \inf\{n \in \{0, \dots, T\} : X_n^T(v) = X_n^T(v')\}, \quad \text{with } \inf \emptyset = \infty,$$

and let

$$\sigma_{v,v'} = \inf\{n \in \mathbb{Z}_+ : Y_n(v) = Y_n(v')\}.$$

Then

$$P(\sigma_{v,v'}^T = t) = P(\sigma_{v,v'} = t), \quad t \leq T. \tag{4.1}$$

Also, let

$$\sigma = \max_{v,v'} \sigma_{v,v'}.$$

Continuing, assume that initially all opinions are distinct. Without loss of generality this can be expressed as $\eta_0(v) = v, v \in V$. We denote this initial state of the system by \mathbf{i} . Observe the following:

- (1) Under $P_{\mathbf{i}}$, the distribution of τ , the consensus time, and of σ coincide.
- (2) Let μ be any initial distribution for the voter model on G . Then the distribution of τ under P_{μ} is stochastically dominated by its distribution under $P_{\mathbf{i}}$.

Before we continue, we would like to recall a fundamental but useful fact. Suppose that η is a \mathbb{Z}_+ -valued random variable with the property that for some $\lambda \in [0, 1]$,

$$\lambda = \lim_{n \rightarrow \infty} P(\eta > n)^{1/n}.$$

Letting $\rho \geq 1$, we can write

$$E[\rho^\eta] = \sum_{k=1}^{\infty} \rho^k (P(\eta > k - 1) - P(\eta > k)) = \rho + (\rho - 1) \sum_{k=1}^{\infty} \rho^k P(\eta > k), \tag{4.2}$$

and therefore it follows from the Cauchy-Hadamard theorem that the radius of convergence of the power series on the right-hand side of (4.2) is $1/\lambda$. In particular,

$$\frac{1}{\lambda} = \sup\{\rho : E[\rho^\eta] < \infty\}. \tag{4.3}$$

Now define

$$\lambda_{\text{CRW}}(G) = \lim_{n \rightarrow \infty} P(\sigma > n)^{1/n}$$

and recall that

$$\lambda_{\text{V}}(G, \mu) = \lim_{n \rightarrow \infty} P_{\mu}(\tau > n)^{1/n}.$$

Both limits exists as σ and τ are hitting times of finite-state Markov chains and decay geometrically (possibly with a polynomial correction). Therefore it follows from (4.3) that

$$\frac{1}{\lambda_{\text{CRW}}(G)} = \sup\{\rho : E[\rho^\sigma] < \infty\}, \tag{4.4}$$

as well as

$$\frac{1}{\lambda_V(G, \mu)} = \sup\{\rho : E_\mu[\rho^\tau] < \infty\}. \tag{4.5}$$

The observations above also imply

$$\lambda_V(G, \mu) \leq \lambda_{\text{CRW}}(G), \tag{4.6}$$

and $\mu = \mathbf{i}$ is a sufficient condition for equality.

We will now examine other sufficient conditions for an equality. For $\rho \geq 1$,

$$\max_{v,v'} E[\rho^{\sigma_{v,v'}}] \leq E[\rho^\sigma] \leq \sum_{v,v'} E[\rho^{\sigma_{v,v'}}]. \tag{4.7}$$

It therefore follows that $E[\rho^\sigma] < \infty$ if and only if $\max_{v,v'} E[\rho^{\sigma_{v,v'}}] < \infty$. Now we can write

$$\begin{aligned} \lambda_{\text{CRW}}(G) &= 1/\sup\{\rho : \max_{v,v'} E[\rho^{\sigma_{v,v'}}] < \infty\} \\ &= \max_{v,v'} 1/\sup\{\rho : E[\rho^{\sigma_{v,v'}}] < \infty\} \\ &= \max_{v,v'} \lim_{n \rightarrow \infty} P(\sigma_{v,v'} > n)^{1/n} \\ &= \lim_{n \rightarrow \infty} \left(\max_{v,v'} P(\sigma_{v,v'} > n) \right)^{1/n} \end{aligned} \tag{4.8}$$

where the first equality follows from (4.4) and (4.7) and the third equality uses (4.3).

Fix $t \geq 0$ and let $v, v' \in V$ be distinct. If $X_t^t(v) = u \neq u' = X_t^t(v')$ and $\eta_0(u) \neq \eta_0(u')$, then necessarily $\eta_t(v) = \eta_0(u)$ and $\eta_t(v') = \eta_0(u')$, hence $\tau > t$. Summing, we have

$$P_\mu(\tau > t) \geq \sum_{u,u'} P_\mu(X_t^t(v) = u, X_t^t(v') = u', \eta_0(u) \neq \eta_0(u')),$$

and we can sum over all u, u' because $\eta_0(u) \neq \eta_0(u')$ implies $u \neq u'$. Since η_0 is independent of the random walk, we can decouple the condition on the random walk from the condition on the initial opinions, and limit the summation only to pairs u, u' where $u \neq u'$. Using (4.1), this gives

$$\begin{aligned} P_\mu(\tau > t) &\geq \sum_{u \neq u'} P(X_t^t(v) = u, X_t^t(v') = u') P_\mu(\eta_0(u) \neq \eta_0(u')) \\ &\geq \sum_{u \neq u'} P(X_t^t(v) = u, X_t^t(v') = u') c \\ &= P(\sigma_{v,v'} > t) c \end{aligned}$$

where

$$\begin{aligned} c &= \min_{u \neq u'} P_\mu(\eta_0(u) \neq \eta_0(u')) \\ &\geq P_\mu(\text{all initial opinions are distinct}). \end{aligned} \tag{4.9}$$

As long as $c > 0$, this implies that the geometric decay of τ starting from μ is at least as slow as that of $\sigma_{v,v'}$. Since v, v' were arbitrary, it follows from (4.8) that

$$\lambda_{\text{CRW}}(G) \leq \lambda_V(G, \mu). \quad (4.10)$$

In view of (4.6) and (4.9), we have established an equality in the case where with positive probability, all initial opinions are distinct.

Next, we relax the condition for equality a little further. Suppose that μ_0 is an initial distribution on any number of opinions such that for some t_0 , $P_{\mu_0}(\eta_{t_0}(u) \neq \eta_{t_0}(u')) > 0$ for all $u \neq u'$. Denote the distribution of η_{t_0} by μ and note that $\lambda_V(G, \mu) = \lambda_{\text{CRW}}(G)$ follows from (4.6), (4.9), and (4.10). For any $\rho \geq 1$, we can use the Markov property to write

$$\begin{aligned} E_{\mu_0}[\rho^\tau] &\geq E_{\mu_0}[\rho^\tau, \tau > t_0] \\ &= E_{\mu_0}[(1 - \mathbf{1}_{\{\tau \leq t_0\}}) \rho^{t_0} E_{\eta_{t_0}}[\rho^\tau]] \\ &= \rho^{t_0} (E_{\mu_0}[E_{\eta_{t_0}}[\rho^\tau]] - E_{\mu_0}[\mathbf{1}_{\{\tau \leq t_0\}} E_{\eta_{t_0}}[\rho^\tau]]) \\ &\geq \rho^{t_0} (E_\mu[\rho^\tau] - 1). \end{aligned}$$

Hence if $E_\mu[\rho^\tau]$ is infinite, then so is $E_{\mu_0}[\rho^\tau]$. Therefore, it follows from (4.5) that $\lambda_V(G, \mu_0) \geq \lambda_V(G, \mu)$. Thus $\lambda_V(G, \mu_0) = \lambda_{\text{CRW}}(G)$ and we have proved the next proposition.

Proposition 4.1. *Let $G = (V, E)$ be a finite connected graph.*

(1) *For any initial opinion distribution μ*

$$\lambda_V(G, \mu) \leq \lambda_{\text{CRW}}(G).$$

(2) *An equality holds in each of the following cases:*

- (a) *With positive probability, all initial opinions are distinct.*
- (b) *There exists $t \geq 0$, such that for every distinct $u, u' \in V$, $P_\mu(\eta_t(u) \neq \eta_t(u')) > 0$.*

An important observation is that we can always attain an equality with only two opinions.

Corollary 4.2. *Let $G = (V, E)$ be a finite connected graph and suppose that μ is the initial distribution on the opinions $\{0, 1\}$ where all vertices have opinion 0 except for one uniformly chosen vertex which has opinion 1. Then*

$$\lambda_V(G, \mu) = \lambda_{\text{CRW}}(G).$$

Example 4.3. This case was treated in Dickman and Vidigal (2002, Section 5) for the continuous-time analogue. Let K_n be the complete graph with n vertices and consider the voter model on K_n for $n \geq 3$ with two opinions: “yes” and “no”. This Markov chain, when restricted to the nonabsorbing states, is irreducible and aperiodic. Hence it follows from Theorem 2.2 that the chain conditioned on nonabsorption converges to the unique QSD. Writing down the eigenvalue equation from Proposition 2.1, we have

$$\begin{aligned} \lambda \nu_n(k) &= \frac{k(k-1) + (n-k)(n-k-1)}{n(n-1)} \nu_n(k) \\ &\quad + \frac{(k+1)(n-k-1)}{n(n-1)} \nu_n(k+1) + \frac{(n-k+1)(k-1)}{n(n-1)} \nu_n(k-1), \end{aligned}$$

where k represents the number of “yes” opinions, and $\nu_n(0) = \nu_n(n) = 0$. Letting $A = k$ and $B = n - k$, the numerators on the right-hand side are equal to $A^2 - A + B^2 - B$, $AB + B - A - 1$, $AB + A - B - 1$, and adding them up gives us $A^2 + B^2 + 2AB - (A + B) - 2 = n(n - 1) - 2$. Therefore, the equation is solved by choosing $\nu_n(k) = \frac{1}{n-1}$ and $\lambda = \frac{n(n-1)-2}{n(n-1)} = 1 - \frac{2}{n(n-1)}$. Proposition 4.1-2b gives $\lambda_V(K_n, \mu) = \lambda_{\text{CRW}}(K_n)$, whenever μ is not supported on consensus states, due to the above claimed irreducibility. Given two walkers at different vertices, the probability that they meet in the next step is $\frac{2}{n(n-1)}$, so it follows from (4.8) that $\lambda_{\text{CRW}}(K_n) = 1 - \frac{2}{n(n-1)}$, as established above through a direct calculation of the QSD.

Summarizing, the QSD for the voter model on the complete graph K_n is uniform on the nonabsorbing states. In particular, the QSDs do not converge to a probability distribution as $n \rightarrow \infty$. As we will see in Section 5, the situation is more interesting for the complete bipartite graphs $K_{n,m}$.

5. Voter model on Complete Bipartite Graphs

Let $K_{n,m} = (V, E)$ be the bipartite graph whose vertex V set is the disjoint union of L and S , where $|L| = n$, $|S| = m$, $m \leq n$, and its edge set is $E = \{\{l, s\} : l \in L, s \in S\}$. Though these graphs are quite simple, an interesting structure appears when considering the behavior of the voter model conditioned on not reaching consensus for a long time.

We will study QSDs for the voter model on $K_{n,m}$ with two opinions, “0” and “1”, also referred to as “no” and “yes”, respectively. As noted before, we assume $m \leq n$. We will also impose the following additional constraints which we need in order to guarantee irreducibility:

$$\begin{aligned} m &\geq 2 \text{ or} \\ m = 1 \text{ and } n &\geq 3. \end{aligned} \tag{5.1}$$

The set Δ of absorbing states for the voter model on $K_{n,m}$ is given by

$$\Delta = \{\eta : \eta \equiv 0 \text{ or } \eta \equiv 1\}.$$

In addition, the states in

$$BP := \{\eta : \eta(l) = 1 - \eta(s), l \in L, s \in S\}$$

are not accessible from any state not in BP . We will therefore eliminate the subsets Δ and BP from our state space for the model. A routine argument shows that under (5.1), the 2-opinion voter model is now irreducible and aperiodic. Hence it follows from Theorem 2.2 and Proposition 4.1 that starting from any initial distribution μ supported on $(\Delta \cup BP)^c$ and conditioning on not reaching consensus, the model converges to the unique QSD which is also supported on $(\Delta \cup BP)^c$. We denote this QSD by $\pi_{n,m}$ and note that it is a left eigenvector corresponding to the eigenvalue $\lambda_{n,m} = \lambda_V(K_{n,m}, \mu) = \lambda_{\text{CRW}}(K_{n,m})$ for the restriction of the transition function of the voter model to $(\Delta \cup BP)^c$.

To continue our analysis, we will exploit the symmetry among vertices within each group. Instead of following the opinion on each vertex, we will follow the number of “yes” opinions in each of the groups S and L . This leads to a Markov chain on the state space $\{0, \dots, n\} \times \{0, \dots, m\}$. Each state is an ordered pair

(k, h) , with k representing the number of “yes” in group L and h representing the number of “yes” in S . Observe that the only allowed transitions are the following:

- (1) $(k, h) \rightarrow (k + 1, h)$. This happens if a “no” vertex in L is sampled and adopts a “yes” from S . The probability of such a transition is therefore $\frac{n-k}{n+m} \frac{h}{m}$.
- (2) $(k, h) \rightarrow (k - 1, h)$. This happens if a “yes” vertex in L is sampled and adopts a “no” from S . The probability is therefore $\frac{k}{n+m} \frac{m-h}{m}$.
- (3) $(k, h) \rightarrow (k, h + 1)$. This is item 1. above with the roles of L and S interchanged and hence occurs with probability $\frac{m-h}{n+m} \frac{k}{n}$.
- (4) $(k, h) \rightarrow (k, h - 1)$. This is item 2. above with the roles of L and S interchanged. Similarly, this happens with probability $\frac{h}{n+m} \frac{n-k}{n}$.
- (5) $(k, h) \rightarrow (k, h)$. This happens with probability $\frac{k}{n+m} \frac{h}{m} + \frac{h}{n+m} \frac{k}{n} + \frac{n-k}{n+m} \frac{m-h}{m} + \frac{m-h}{n+m} \frac{n-k}{n} = \frac{kh+(n-k)(m-h)}{nm}$.

Of course, $(0, 0), (n, m)$ are the unique absorbing states, and the set BP collapses into two states, $(0, m)$ and $(n, 0)$, not accessible from any other state. Thus eliminating these four states, the chain is irreducible. As a result, it possesses a unique QSD which we denote by $\mu_{n,m}$. Recall that $\lambda_{n,m} = \lambda_{\text{CRW}}(K_{n,m})$ coincides with the geometric rate of the time to absorption in the voter model. The absorption time for the new chain from any initial state coincides with the time to absorption for the voter model starting from any state with matching numbers of opinions in both S and L , therefore, it follows from Proposition 2.1 that the eigenvalue corresponding to $\mu_{n,m}$ is equal to $\lambda_{n,m}$. Now fix any state (k, h) for our new chain. By looking at all possible transitions we obtain the following equation for $\mu_{n,m}$:

$$\begin{aligned} \lambda_{n,m} \mu_{n,m}(k, h) &= \mu_{n,m}(k, h) \frac{kh + (n - k)(m - h)}{nm} \\ &+ \mu_{n,m}(k - 1, h) \frac{(n - k + 1)h}{m(n + m)} + \mu_{n,m}(k + 1, h) \frac{(k + 1)(m - h)}{m(n + m)} \\ &+ \mu_{n,m}(k, h - 1) \frac{(m - h + 1)k}{n(n + m)} + \mu_{n,m}(k, h + 1) \frac{(h + 1)(n - k)}{n(n + m)}. \end{aligned} \tag{5.2}$$

In order to extract more information, we first compute $\lambda_{n,m}$.

Proposition 5.1.

$$\begin{aligned} \lambda_{n,m} = \lambda_{\text{CRW}}(K_{n,m}) &= 1 - \frac{2}{n + m} \left(1 - \sqrt{1 - \frac{1}{2n} - \frac{1}{2m}} \right) \\ &= 1 - \frac{\gamma_{n,m}}{n + m}, \end{aligned}$$

where

$$\gamma_{n,m} = 2 \left(1 - \sqrt{1 - \frac{1}{2n} - \frac{1}{2m}} \right).$$

Proof: We assume first $m > 1$. From (4.8) it is enough to consider only two coalescing random walks on $K_{n,m}$. The two CRW paths can be in either one of the following states: both walks are in different vertices of L , both walks are in different vertices of S , one walk is in S and another in L , or they are both at the same vertex. Label these four states of the system as 1, 2, 3, 4, respectively. Of course, 4 is the absorbing state for the CRW, so we will omit it from our calculations. From each

of the states 1, 2, 3, the system will stay put with probability $1 - \frac{2}{n+m}$. From state 1 the system can transition to 3 with probability $\frac{2}{n+m}$, and similarly, from state 2 the system can transition to state 3 with probability $\frac{2}{n+m}$. Finally, from state 3, the system can transition to 1 or to 2 with respective probabilities $\frac{1}{n+m} \frac{n-1}{n}$ and $\frac{1}{n+m} \frac{m-1}{m}$. As a result the substochastic transition function on states 1, 2, 3 is

$$\begin{pmatrix} \frac{n+m-2}{n+m} & 0 & \frac{2}{n+m} \\ 0 & \frac{n+m-2}{n+m} & \frac{2}{n+m} \\ \frac{1}{n+m} \frac{n-1}{n} & \frac{1}{n+m} \frac{m-1}{m} & \frac{n+m-2}{n+m} \end{pmatrix}. \tag{5.3}$$

Since from both states 1 and 2 the transitions are either to themselves with the same probability, or to 3 with the complementary probability, we can consolidate these states into one, leading to the matrix

$$\begin{pmatrix} \frac{n+m-2}{n+m} & \frac{2}{n+m} \\ \frac{1}{n+m} \left(\frac{n-1}{n} + \frac{m-1}{m} \right) & \frac{n+m-2}{n+m} \end{pmatrix}. \tag{5.4}$$

The characteristic equation is

$$\left(\lambda - \frac{n+m-2}{n+m} \right)^2 - \frac{2}{(n+m)^2} \frac{2mn - m - n}{nm} = 0.$$

Therefore the two eigenvalues, λ_+ and λ_- , are given by

$$\begin{aligned} \lambda_{\pm} &= 1 - \frac{2}{n+m} \pm \frac{1}{n+m} \sqrt{4 - \frac{2}{n} - \frac{2}{m}} \\ &= 1 - \frac{2}{n+m} \left(1 \pm \sqrt{1 - \frac{1}{2n} - \frac{1}{2m}} \right), \end{aligned}$$

and the largest eigenvalue is obtained by choosing λ_- (using the “-” sign), giving the expression in the statement.

It remains to consider the case $m = 1$. In this case, state 2 is not possible. We therefore eliminate the second row and the second column from (5.3), ending up with the matrix (5.4) and then continue as before. \square

The next result gives a more direct connection between $\lambda_{n,m}$ and $\mu_{n,m}$.

Proposition 5.2.

$$\lambda_{n,m} = 1 - \frac{2}{n+m} (\mu_{n,m}(1, 0) + \mu_{n,m}(0, 1))$$

Proof: Let (K, H) be a random vector representing the number of “yes” in L and S , respectively, whose distribution is $\mu_{n,m}$. Recalling that $\mu_{n,m}(0, 0) = \mu_{n,m}(n, m) = 0$, we can sum both sides of (5.2) over $-1 \leq k \leq n + 1$ and $-1 \leq h \leq m + 1$ while

eliminating from the sum the pairs $(k, h) = (0, 0)$ and $(k, h) = (n, m)$ to obtain

$$\begin{aligned} \lambda_{n,m} &= \frac{1}{nm} (E[KH] + E[(n-K)(m-H)]) \\ &\quad + \frac{1}{m(n+m)} (E[(n-K)H] - \mu_{n,m}(n-1, m)m) \\ &\quad + \frac{1}{m(n+m)} (E[K(m-H)] - \mu_{n,m}(1, 0)m) \\ &\quad + \frac{1}{n(n+m)} (E[(m-H)K] - \mu_{n,m}(n, m-1)n) \\ &\quad + \frac{1}{n(n+m)} (E[H(n-K)] - \mu_{n,m}(0, 1)n) \\ &= 1 - \frac{1}{n+m} (\mu_{n,m}(n-1, m) + \mu_{n,m}(1, 0) + \mu_{n,m}(n, m-1) + \mu_{n,m}(0, 1)) \\ &= 1 - \frac{2}{n+m} (\mu_{n,m}(1, 0) + \mu_{n,m}(0, 1)) \end{aligned}$$

where the last equality follows from invariance under relabeling of the two opinions. \square

We now show how to use this proposition to compute $\mu_{n,1}(j, 0)$. As the state $(0, 1)$ is in BP and not in the support of $\mu_{n,1}$, using Proposition 5.2 along with Proposition 5.1 gives

$$\begin{aligned} \mu_{n,1}(1, 0) &= \frac{n+1}{2} (1 - \lambda_{n,1}) = 1 - \sqrt{\frac{1}{2} - \frac{1}{2n}} \\ &= \frac{\gamma_{n,1}}{2}. \end{aligned} \tag{5.5}$$

Writing out (5.2) for this case leaves us with

$$\left(\lambda_{n,1} - \frac{n-k}{n} \right) \mu_{n,1}(k, 0) = \mu_{n,1}(k+1, 0) \frac{k+1}{n+1} + \mu_{n,1}(k, 1) \frac{n-k}{n(n+1)}.$$

Due to invariance under relabeling of the two opinions, we have $\mu_{n,1}(k, 1) = \mu_{n,1}(n-k, 0)$. In addition, from Proposition 5.1 we have

$$\lambda_{n,1} - \frac{n-k}{n} = \frac{k}{n} - \frac{\gamma_{n,1}}{n+1}.$$

We therefore obtain the equation

$$\left(\frac{k}{n} - \frac{\gamma_{n,1}}{n+1} \right) \mu_{n,1}(k, 0) = \mu_{n,1}(k+1, 0) \frac{k+1}{n+1} + \mu_{n,1}(n-k, 0) \frac{n-k}{n(n+1)}. \tag{5.6}$$

This nonlocal recurrence relation can be solved through iteration. Having calculated $\mu_{n,1}(1, 0)$ in (5.5) and recalling that $\mu_{n,1}(n, 0) = 0$, we can plug $k = n-1$ into (5.6) and obtain

$$\left(\frac{n-1}{n} - \frac{\gamma_{n,1}}{n+1} \right) \mu_{n,1}(n-1, 0) = \mu_{n,1}(1, 0) \frac{1}{n(n+1)},$$

or

$$\mu_{n,1}(n-1, 0) = \frac{\gamma_{n,1}}{2(n^2 - n\gamma_{n,1} - 1)}.$$

We can repeat this procedure inductively. Having calculated $\mu_{n,1}(j, 0)$ and $\mu_{n,1}(n - j, 0)$ for $j = 1, \dots, k < n$, we can use (5.6) to recover $\mu_{n,1}(k + 1, 0)$ and then $\mu_{n,1}(n - k - 1, 0)$.

Next we prove a technical lemma which gives the asymptotic behavior of $\mu_{n,m}$ as $n \rightarrow \infty$.

Lemma 5.3. *Suppose $k \geq 0$ and $h > 0$. Then*

$$\lim_{n \rightarrow \infty} \mu_{n,m}(k, h) = 0.$$

Proof: Assume $k, h > 0$. From (5.2) we have

$$\left(\lambda_{n,m} - \frac{kh + (n - k)(m - h)}{nm} \right) \mu_{n,m}(k, h) = (1 + o(1)) \frac{h}{m} \mu_{n,m}(k - 1, h) + O\left(\frac{1}{n}\right)$$

as $n \rightarrow \infty$. From Proposition 5.1, we know that $\lambda_{n,m} \rightarrow 1$ as $n \rightarrow \infty$. Hence

$$\mu_{n,m}(k, h) = (1 + o(1)) \mu_{n,m}(k - 1, h) + O\left(\frac{1}{n}\right) \text{ as } n \rightarrow \infty. \tag{5.7}$$

Now suppose for the sake of a contradiction that $\limsup_{n \rightarrow \infty} \mu_{n,m}(k', h') = \epsilon > 0$ for some $k' \geq 0$ and $h' > 0$. Then (5.7) implies that $\lim_{j \rightarrow \infty} \mu_{n_j,m}(k' + 1, h') = \epsilon$ along some subsequence n_1, n_2, \dots . Similarly, (5.7) can be used again to show that $\lim_{j \rightarrow \infty} \mu_{n_j,m}(k' + 2, h') = \epsilon$ along the same subsequence. Reasoning inductively both forwards and backwards in k , it follows that $\lim_{j \rightarrow \infty} \mu_{n_j,m}(k, h') = \epsilon$ for all $k \geq 0$. In particular, for j large enough we have $\mu_{n_j,m}(k, h') > \epsilon/2$ for $0 \leq k \leq \lceil 2/\epsilon \rceil$. Hence $\sum_k \mu_{n_j,m}(k, h') > 1$, a contradiction. \square

In the following proposition we calculate the pointwise limit of $\mu_{n,m}$ as the size of the large partition tends to infinity.

Proposition 5.4. *Let f_{γ_m} be as in (1.1) with γ_m defined in Theorem 1.1. Then for $k \geq 0$*

$$\mu_{\infty,m}(k, 0) := \lim_{n \rightarrow \infty} \mu_{n,m}(k, 0) = \frac{1}{2} f_{\gamma_m}(k).$$

Proof: From Lemma 5.3 we know that $\lim_{n \rightarrow \infty} \mu_{n,m}(0, 1) = 0$. Therefore it follows from Propositions 5.1 and 5.2 that

$$\mu_{\infty,m}(1, 0) = \frac{\gamma_m}{2}. \tag{5.8}$$

Returning to (5.2) with $k \geq 1$ and $h = 0$, we can write

$$\lambda_{n,m} \mu_{n,m}(k, 0) = \mu_{n,m}(k, 0) \frac{n - k}{n} + \mu_{n,m}(k + 1, 0) \frac{k + 1}{n + m} + \mu_{n,m}(k, 1) \frac{n - k}{n(n + m)}.$$

Rearranging this equation while using Proposition 5.1 again leads to

$$\mu_{n,m}(k + 1, 0) = \frac{n + m}{k + 1} \left(\mu_{n,m}(k, 0) \left(\frac{k}{n} - \frac{\gamma_{n,m}}{n + m} \right) - \mu_{n,m}(k, 1) \frac{n - k}{n(n + m)} \right). \tag{5.9}$$

Letting $n \rightarrow \infty$ in (5.9) while recalling that $\gamma_{n,m} \rightarrow \gamma_m$ and $\mu_{n,m}(k, 1) \rightarrow 0$ results in

$$\mu_{\infty,m}(k + 1, 0) = \mu_{\infty,m}(k, 0) \frac{k - \gamma_m}{k + 1}.$$

Now we can argue inductively starting from (5.8) to conclude that

$$\begin{aligned}\mu_{\infty,m}(k,0) &= \frac{\gamma_m}{2} \frac{1}{k!} \prod_{j=1}^{k-1} (j - \gamma_m) \\ &= \frac{1}{2} f_{\gamma_m}(k), \quad k \in \mathbb{Z}_+.\end{aligned}$$

□

Finally we can give the proof of our main result.

Proof of Theorem 1.1: By the invariance under relabeling of the two opinions, we know that $\mu_{n,m}(k,h) = \mu_{n,m}(n-k,m-h)$. Hence Proposition 5.4 implies that

$$\lim_{n \rightarrow \infty} \mu_{n,m}(n-k,m) = \lim_{n \rightarrow \infty} \mu_{n,m}(k,0) = \frac{1}{2} f_{\gamma_m}(k), \quad k \in \mathbb{Z}_+.$$

Since we know that f_{γ_m} is a probability mass function, it follows from the Portmanteau theorem that the QSDs for the voter model on $K_{n,m}$ converge weakly as $n \rightarrow \infty$ to the probability distribution where S attains a consensus of all “0” or all “1” each with probability 1/2, and conditioned on the opinion of S , the number of vertices in L which are of a different opinion has probability mass function f_{γ_m} . □

Acknowledgments

The authors would like to thank an anonymous referee for their useful comments.

References

- Aldous, D. and Fill, J. A. Reversible Markov Chains and Random Walks on Graphs (2002). Unfinished monograph, recompiled 2014, available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- Christoph, G. and Schreiber, K. Scaled Sibuya distribution and discrete self-decomposability. *Statist. Probab. Lett.*, **48** (2), 181–187 (2000). [MR1769823](#).
- Collet, P., Martínez, S., and San Martín, J. *Quasi-stationary distributions. Markov chains, diffusions and dynamical systems*. Probability and its Applications (New York). Springer, Heidelberg (2013). ISBN 978-3-642-33130-5; 978-3-642-33131-2. [MR2986807](#).
- Cooper, C. and Rivera, N. The linear voting model. In *43rd International Colloquium on Automata, Languages, and Programming*, volume 55 of *LIPICs. Leibniz Int. Proc. Inform.*, pp. Art. No. 144, 12. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern (2016). [MR3577205](#).
- Devroye, L. A triptych of discrete distributions related to the stable law. *Statist. Probab. Lett.*, **18** (5), 349–351 (1993). [MR1247445](#).
- Dickman, R. and Vidigal, R. Quasi-stationary distributions for stochastic processes with an absorbing state. *J. Phys. A*, **35** (5), 1147–1166 (2002). [MR1891634](#).
- Durrett, R. *Lecture notes on particle systems and percolation*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA (1988). ISBN 0-534-09462-7. [MR940469](#).
- Horn, R. A. and Johnson, C. R. *Matrix analysis*. Cambridge University Press, Cambridge, second edition (2013). ISBN 978-0-521-54823-6. [MR2978290](#).

- Huillet, T. E. On New Mechanisms Leading to Heavy-Tailed Distributions Related to the Ones Of Yule-Simon. *Indian J. Pure Appl. Math.*, **51** (1), 321–344 (2020). [MR4076215](#).
- Johnson, N. L., Kemp, A. W., and Kotz, S. *Univariate discrete distributions*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition (2005). ISBN 978-0-471-27246-5; 0-471-27246-9. [MR2163227](#).
- Kozubowski, T. J. and Podgórski, K. A generalized Sibuya distribution. *Ann. Inst. Statist. Math.*, **70** (4), 855–887 (2018). [MR3830290](#).
- Letac, G. Is the Sibuya distribution a progeny? *J. Appl. Probab.*, **56** (1), 52–56 (2019). [MR3981145](#).
- Liggett, T. M. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York (1985). ISBN 0-387-96069-4. [MR776231](#).
- Oliveira, R. I. On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.*, **364** (4), 2109–2128 (2012). [MR2869200](#).
- Redner, S. Reality-inspired voter models: A mini-review. *Comptes Rendus Physique*, **20** (4), 275 – 292 (2019). DOI: [10.1016/j.crhy.2019.05.004](#).
- Sood, V. and Redner, S. Voter Model on Heterogeneous Graphs. *Phys. Rev. Lett.*, **94**, 178701 (2005). DOI: [10.1103/PhysRevLett.94.178701](#).