Chapter 2

The Voter Model with Confidence Levels

2.1 Introduction

Groups of people often have to decide together on some issue or course of action. We wish to form a model from which we can make conclusions about the group as a whole. In the model, we will incorporate the group dynamics on a local scale, including how and when a person changes his or her opinion and how other people influence this decision. The *voter model* is one stochastic process that models such interactions. In the voter model, there is a group of voters, each of whom has an opinion 1 or 0 representing *yes* or *no* on some issue. Each voter has relationships to a subset of the other voters, whom we call neighbors, and as time passes, each voter's opinion is influenced by its neighbors. We wish to know how the voters' opinions evolve as a whole. Do they ever come to a consensus on the issue? If consensus occurs, which outcome (1 or 0) is more likely, given every voter's initial opinion? If the vertices of a graph G represent voters and the edges neighbor relationships between voters, then how does the structure of G affect the outcome?

Our work on the voter model was also motivated by an interest in the spread of infectious diseases. Here, the opinions correspond to disease states and a disease spreads from individuals to "neighbors," individuals with whom someone has social contact. We wish to determine how the distribution of infected individuals evolves over time. Do we ever end up in the state when no one is infected or the state where everyone is infected? How does the structure of the edge neighbor relationships affect the outcome? The voter model does not precisely capture disease dynamics since being infected and uninfected is not symmetric in the same way the 1 and 0 opinions are. However, insights into the voter model may lead to insights into more complicated epidemiological models.

The voter model was introduced independently by Clifford and Sudbury [3] and by Holley and Liggett [16]. Clifford and Sudbury were interested in modeling spatial conflicts and geographic dominance between two competing species, and Holley and Liggett viewed the voter model as describing the configuration of spins in a particle system. As the name suggests, the voter model also can be viewed as representing the opinions of a group of voters. Our terminology will come from this last interpretation.

In the extensive literature on the voter model, the limiting behavior of the process is completely described. Unfortunately, the dependence of the limiting behavior on the graph G is minimal and hence not very interesting. Because of the interaction mechanism usually used in such models for determining how a voter is influenced by its neighbors, the final outcome is based only a simple graph theoretic property: the sum of the degrees of the vertices where the voters' initial opinions are 1. In this chapter, we present a different interaction mechanism involving a voter's *confidence level*. In this modified voter model, the structure of G has a large impact on evolution of the voters' opinions. When G is finite and connected, consensus always occurs, and we can explicitly calculate how likely each outcome is. The remarkable feature of the voter model with confidence levels is that determining the likelihood of each outcome remains tractable *despite* this dependence on the structure of G.

The standard reference for the voter model and related interacting particle systems is Liggett [20] and its continuation [21]. Griffeath [14] and Durrett [9] are also widely used sources that provide slightly different perspectives. While the previous authors consider the voter model on infinite graphs, particularly the square lattices in various dimensions, Aldous and Fill [1] is a very readable account of the theory on finite graphs. Also on finite graphs, Donnelly and Welsh [7] consider the probability of each outcome and the time needed to reach consensus.

The area of social influence has an extensive literature on the models of opinion formulation. Poljak and Sûra [25] developed a model where each person has one of a finite set of opinions, each person has some measure of influence on the opinion of every other person, and these influences are symmetric. At each discrete time step, a person's opinion is updated to the opinion held by the most neighbors, weighted by their influences. DeGroot [5] and French [12] proposed a similar model with opinions taken from a real interval, and where a person's opinion is updated to an average of his or her neighbor's opinions, the average being weighted according to the neighbors' influence. Both of these models are related to the Delphi method ([4], [22]) developed at the RAND Corporation for consensus finding and problem solving in groups. Latané used neural networks to in [17], [18], and [19] to model social impact, and Merrill presented similar models in [24]. Falmagne and his colleagues in [6], [10], and [11] explored models for approval voting [2], where voters' opinions varied according to a stochastic stream of tokens or simple messages.

The idea of confidence levels was inspired by the work of Hoffman [15] and Roberts [27] who considered the use of confidence levels in deterministic models of opinion formulation on graphs. Confidence in models of opinion formulation is analogous to disease resistance in epidemiological models since both dynamically affect how quickly an individual can have his or her opinion changed, or be infected. Insights into the effect of confidence levels may lead to a greater understanding of how resistance affects epidemiological models.

In section 2.2 we provide a brief discussion of the original voter model to establish terminology and some of the tools that we will be using. In section 2.3 we present our version of the voter model with confidence levels and determine the limiting probability of entering the uniform 1 opinion, given the graph and the initial opinions. In section 2.4 we present some sample waiting functions and calculate the limiting probability for these functions on a sample graph. Section 2.5 mentions open questions and potential alternative ways to bring confidence levels into the voter model.

2.2 The Voter Model

The voter model is well studied, though often many authors focus on infinite graphs. We present here the voter model on arbitrary finite graphs, using our own terminology and notation. Let G be a finite, connected, undirected graph on n vertices labeled v_1 through v_n. We define a continuous-time Markov process $\{Z_t\}_{t\geq 0}$ on the finite state space $\{0,1\}^{V(G)}$ as follows. At every time $t \geq 0$, each vertex v_i has an *opinion* $Z_t(v_i)$ in $\{0,1\}$. Each vertex

v_i asynchronously updates its opinion at times $\{T_\ell(v_i)\}_{\ell=1}^{\infty}$, where the sequence $\{T_\ell(v_i)\}$ is a Poisson process with rate 1; that is, $T_{\ell+1}(v_i) - T_\ell(v_i)$ is independently exponentially distributed with mean 1 for each ℓ. The $\{T_\ell(v_i)\}$ are also independent for each v_i. We set $T_0(v_i) = 0$ for convenience.

Vertex v_i updates its opinion at the times $\{T_\ell(v_i)\}_{\ell=1}^{\infty}$ according to the following rule: for $T_\ell(v_i) < t \leq T_{\ell+1}(v_i)$,

$$Z_t(v_i) = \begin{cases} 0, & \text{with probability } \#\{x \in N_G(v_i) : Z_{T_\ell(v_i)}(x) = 0\}/\#N_G(v_i), \\ 1, & \text{with probability } \#\{x \in N_G(v_i) : Z_{T_\ell(v_i)}(x) = 1\}/\#N_G(v_i). \end{cases}$$

Note that $Z_t(v_i)$ is updated only at the times $\{T_\ell(v_i)\}$[1], and is constant on each interval $(T_\ell(v_i), T_{\ell+1}(v_i)]$. Also observe that the update mechanism is probabilistically the same as v_i adopting the opinion of a neighbor $N_\ell(v_i)$ chosen uniformly at random from $N_G(v_i)$: for $T_\ell(v_i) < t \leq T_{\ell+1}(v_i)$,

$$Z_t(v_i) = Z_{T_\ell(v_i)}(N_\ell(v_i)), \text{ where } N_\ell(v_i) \text{ is chosen uniformly from } N_G(v_i).$$

We will use this view of the update mechanism for analyzing the voter model.

Because G is finite and connected, the only absorbing states for the process $\{Z_t\}_{t\geq 0}$ are the states of uniform opinion where either $Z_t(v_i) = 0$ for all i, or $Z_t(v_i) = 1$ for all i. Since the process will eventually enter one of these absorbing states with probability 1, we wish to know the probability of entering the uniform 1 opinion given the initial configuration Z_0 of opinions. All vertices have the same opinion in the absorbing states, and so this probability is the same as the limiting probability of v_i's opinion being 1, for any vertex v_i. Thus,

$$\Pr\{Z_t \text{ enters the uniform 1 opinion state} \mid Z_0\} = \lim_{t\to\infty} \Pr\{Z_t(v_i) = 1 \mid Z_0\}.$$

This probability can be found by considering another continuous-time Markov process $\{X_s\}$ that is dual to $\{Z_t\}$. The process $\{X_s\}_{s\geq 0}$ is a system of random walks, with one random walk $X_s(v_i)$ starting at each vertex v_i. Formally, the process $\{X_s\}$ is defined on the finite state space $V(G)^{V(G)}$, where at every $s \geq 0$, each vertex has as its state $X_s(v_i)$

---

[1] Strictly speaking, this isn't true since v_i's opinion does not change until immediately *after* $T_\ell(v_i)$. However, we will stick to the terminology that "v_i updates at $T_\ell(v_i)$."

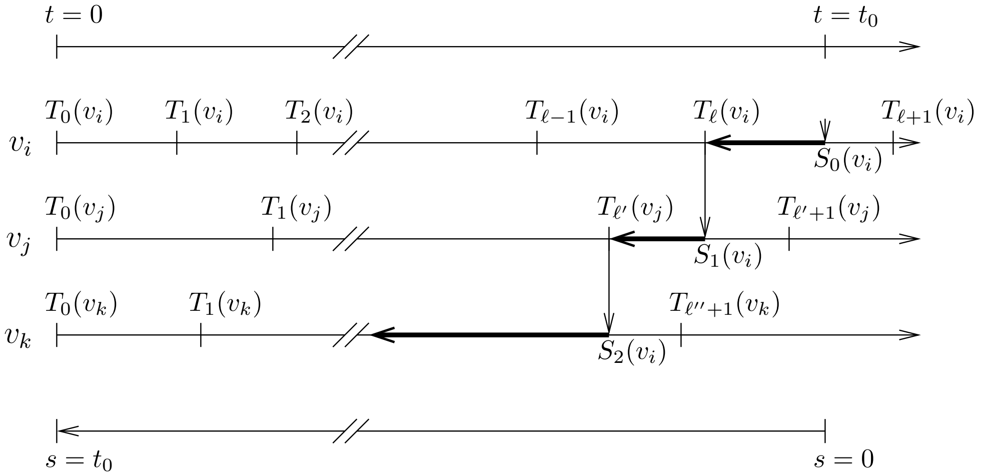

Figure 2.1: $\{X_s\}$ progresses backwards in time relative to $\{Z_t\}$.

a vertex in $V(G)$. The random walk $X_s(v_i)$ updates at the discrete times $\{S_m(v_i)\}_{m=0}^{\infty}$. We think of the dual process as progressing *backwards* in time relative to $\{Z_t\}$, as shown in Figure 2.1. Fix a time $t_0 > 0$. The time $t = t_0$ for $\{Z_t\}$ corresponds to the time $s = 0$ for $\{X_s\}$. For each vertex v_i, $X_s(v_i)$ traces where v_i's opinion at time t_0 came from. We start with $X_0(v_i) = v_i$ and $S_0(v_i) = 0$. At some time $T_\ell(v_i)$, where $T_\ell(v_i) < t_0 \leq T_{\ell+1}(v_i)$, v_i adopted the opinion of a neighbor $v_j := N_\ell(v_i)$. We set $S_1(v_i) := t_0 - T_\ell(v_i)$ and update $X_{S_1(v_i)}(v_i)$ to v_j. For those s such that $0 \leq s < S_1(v_i)$, $X_s(v_i)$ is constant. At some time $T_{\ell'}(v_j)$, where $T_{\ell'}(v_j) < T_\ell(v_i) \leq T_{\ell'+1}(v_j)$, v_j adopted the opinion of a neighbor $v_k := N_{\ell'}(v_j)$. We set $S_2(v_i) := t_0 - T_{\ell'}(v_j)$ and update $X_{S_2(v_i)}(v_i) = v_k$. For those s such that $S_1(v_i) \leq s < S_2(v_i)$, $X_s(v_i)$ is constant. We continue this process, following v_i's opinion backwards in time, for $s \leq t_0$. Since v_i's opinion at time t_0 is the opinion of $X_s(v_i)$ at time $t_0 - s$, we have that

$$Z_{t_0}(v_i) = Z_{t_0-s}(X_s(v_i)).$$

Since these random variables are equal, we also have

$$\Pr\{Z_{t_0}(v_i) = 1 \mid Z_0\} = \Pr\{Z_{t_0-s}(X_s(v_i)) = 1 \mid Z_0\}.$$

By setting $s = t_0$ and taking the limit as $t_0 \to \infty$, we conclude that

$$\lim_{t_0\to\infty} \Pr\{Z_{t_0}(v_i) = 1 \mid Z_0\} = \lim_{t_0\to\infty} \Pr\{Z_0(X_{t_0}(v_i)) = 1 \mid Z_0\}. \tag{2.1}$$

The left hand limit in equation (2.1) is the quantity we want, and thus we have reduced the problem to calculating the limiting distribution of $X_s(v_i)$. The process $X_s(v_i)$ is simply a random walk where $X_s(v_i)$ moves to a neighbor chosen uniformly at random at the discrete times $\{S_m(v_i)\}_{m=0}^{\infty}$. How is $\{S_m(v_i)\}$ distributed? Again, fix $t_0 > 0$, and let m be such that $S_{m+1}(v_i) < t_0$. Set $v_j := X_{S_m(v_i)}(v_i)$, and let ℓ be such that $T_\ell(v_j) < t_0 - S_m(v_i) \leq T_{\ell+1}(v_j)$. Note that $S_{m+1}(v_i) = t_0 - T_\ell(v_j)$. The length of the interval $T_{\ell+1}(v_j) - T_\ell(v_j)$ is exponentially distributed with mean 1, and indicates no information about the direction of time. Thus, the random variable $T_{\ell+1}(v_j) - T_\ell(v_j)$ conditioned on $T_\ell(v_j)$ and the random variable $T_{\ell+1}(v_j) - T_\ell(v_j)$ conditioned on $T_{\ell+1}(v_j)$ are both exponentially distributed with mean 1. Since exponential distributions are memoryless, the distribution of when all the random walks $X_s(v_i)$ move from a vertex v_j is independent of when the random walks arrived at v_j. Hence, $S_{m+1}(v_i) - S_m(v_i)$ is exponentially distributed with mean 1. Since the $\{T_{\ell+1}(v_j) - T_\ell(v_j)\}$ are independent for each v_j and ℓ, the $\{S_{m+1}(v_i) - S_m(v_i)\}$ are also independent for each v_i and m. Therefore, $\{X_s(v_i)\}_{s\geq 0}$ is a standard continuous-time random walk on G; that is, $\{X_s(v_i)\}_{s\geq 0}$ is a random walk that moves to a neighbor chosen uniformly at random according to a Poisson process with rate 1. From the theory of random walks (see, for example, [23]),

$$\lim_{s\to\infty} \Pr\{Z_0(X_s(v_i)) = 1\} = \frac{\displaystyle\sum_{x\in V(G):Z_0(x)=1} \deg_G(x)}{\displaystyle\sum_{x\in V(G)} \deg_G(x)}. \tag{2.2}$$

Note again that the graph structure does not enter into equation (2.2) and the degrees of the vertices do. The theory of random walks actually gives us more precise information, in that we can calculate $\lim_{s\to\infty} \Pr\{X_s(v_i) = x\}$. Summing these values over all x such that $Z_0(x) = 1$ gives us equation (2.2) above.

The random processes $\{X_s(v_i)\}$ for different vertices v_i are coupled together in the following way. Suppose that $X_s(v_i) = X_s(v_j) = v_k$ for $T_\ell(v_k) < t_0 - s \leq T_{\ell+1}(v_k)$. Then $X_{t_0 - T_\ell(v_k)}(v_i) = X_{t_0 - T_\ell(v_k)}(v_j) = N_\ell(v_k)$. In words, if $X_s(v_i)$ and $X_s(v_j)$ are both at v_k, then $X_s(v_i)$ and $X_s(v_j)$ move *together* to the vertex $N_\ell(v_k)$ that is chosen uniformly from v_k's neighborhood. From that point on, $X_s(v_i)$ and $X_s(v_j)$ remain coupled, in the sense that $X_s(v_i) = X_s(v_j)$ for $s \geq t_0 - T_\ell(v_j)$. The random walks $\{X_s\}$ are called *coalescing* *random walks*, and two random walks $X_s(v_i)$ and $X_s(v_j)$ are said to *coalesce* when the two random walks hit and move together thereafter. With probability 1, all of the random walks coalesce (see [20]); that is, there is some time S such that $X_s(v_i) = X_s(v_j)$ for $s > S$ and for all v_i and v_j. This provides another proof of the fact that the only absorbing states for the process $\{Z_t\}_{t \geq 0}$ are the states of uniform opinion.

2.3 The Voter Model with Confidence Levels

We now extend the voter model to include the *confidence* a voter has in its opinion. Specifically, we include the idea that a voter is less likely to consider the opinions of neighbors the higher its confidence level is. We present one specific set of assumptions for this extension, and in section 2.5 we point out other possibilities for the assumptions for future consideration.

To motivate this extension, we consider a group of voters in a two-party political system. All of the relationships between voters are known and are represented as edges in the graph G. We assume that the two parties have opposing views on some issue, and that the members of each party initially assume their party's position by default before discussion begins. As the discussion progresses, a voter v_i is equally swayed by all of its neighbors in G, whether they are in its party or not. However, members of the same party are trusted associates in the sense that if trusted associate v_j convinces v_i of its opinion, then v_j also convinces v_i of v_j's confidence level in that opinion. In fact, v_i's new confidence level is even one greater than v_j's confidence level $C_t(v_j)$. Not only does v_i accept v_j's confidence in the opinion, but v_j's own credibility is added to v_i's confidence level. When v_i adopts the opinion of a non-trusted associate, then its confidence level simply resets to a level of 0. The effect of a higher confidence level is that it takes longer for any of v_i's neighbors to convince v_i of a new opinion.

We now formally specify the voter model with confidence levels. As above, we have a finite connected undirected graph G and a continuous-time stochastic process $\{(Z_t, C_t)\}_{t \geq 0}$ defined on the state space $(\{0,1\} \times \mathbb{N})^{V(G)}$. At every time $t \geq 0$, each vertex v_i has an *opinion* $Z_t(v_i)$ in $\{0,1\}$ and a *confidence level* $C_t(v_i)$ in $\mathbb{N}$. We use the confidence level to count, in some sense, the number of times that v_i's current opinion has been affirmed by other voters. A vertex v_i asynchronously updates its opinion at times $\{T_\ell(v_i)\}_{\ell=1}^{\infty}$ by adopting the opinion of a neighbor $v_j := N_\ell(v_i)$ chosen uniformly at random; that is, for $T_\ell(v_i) < t \leq T_{\ell+1}(v_i)$,

$$Z_t(v_i) = Z_{T_\ell(v_i)}(v_j), \text{ where } v_j := N_\ell(v_i) \text{ is chosen uniformly from } N_G(v_i).$$

Vertex v_i's confidence level is updated at the same times $\{T_\ell(v_i)\}$. However, these times are not distributed as before, so we will specify their distribution below. To specify how the confidence level is updated, we define v_j to be a *trusted associate* of v_i if and only if v_j's initial opinion $Z_0(v_j)$ agrees with v_i's initial opinion $Z_0(v_i)$. The *trusted component* $K(v_i)$ of v_i is the component containing v_i of the subgraph of G induced by vertices whose initial opinions agree with v_i's. Note that while opinions and confidence levels change as the process progresses, the trusted relationship does not. For $T_\ell(v_i) < t \leq T_{\ell+1}(v_i)$, $C_t(v_i)$ is updated as follows:

$$C_t(v_i) = \begin{cases} C_{T_\ell(v_i)}(v_j) + 1 & \text{if } v_j = N_\ell(v_i) \text{ is a trusted associate of } v_i, \\ 0 & \text{if not.} \end{cases} \tag{2.3}$$

Notice that v_i increments v_j's confidence level if $v_j = N_\ell(v_i)$ is a trusted associate, and not v_i's own confidence level.

Rather than specifying the initial confidence level $C_0(v_i)$ as a single value, we instead define $C_0(v_i)$ to be a random variable with distribution

$$\Pr\{C_0(v_i) = \ell\} = \Pr\{X_s(v_i) \text{ leaves } K(v_i) \text{ in exactly } \ell \text{ steps}\}.$$

We will explain this assumption during our analysis of the voter model, particularly what the process $X_s(v_i)$ is. However, the reason for this choice is that $C_0(v_i)$ is a "typical" value representative of the values that occur as the process $\{(Z_t, C_t)\}$ is progressing. This assumption simplifies the analysis since the initial condition for $C_t(v_i)$ looks the same as when the process is running. When computing the limiting probability $\lim_{t\to\infty} \Pr\{Z_0(X_t(v_i)) = 1 \mid Z_0\}$, we condition on Z_0 (and intrinsically on G), but *not* on C_0.

The length of the time interval $(T_\ell(v_i), T_{\ell+1}(v_i)]$ is dependent purely on the confidence level $C_t(v_i)$ that v_i has during that time interval. In most cases, we want a higher confidence level to mean a longer time before opinions are updated. Let the *waiting function* $f\colon \mathbb{N} \to \mathbb{R}^+$ be a positive function whose differences $f(c) - f(c-1)$ are increasing for positive integers c. If $c = C_{T_\ell(v_i)}(v_i)$, then $T_{\ell+1}(v_i) - T_\ell(v_i)$ is exponentially distributed with mean $f(c) - f(c-1)$ if c is positive, and exponentially distributed with mean $f(0)$ if $c = 0$. We set $T_0(v_i) = 0$ for convenience. This choice of how confidence levels affects update times is again chosen to simplify the analysis of the voter model. However, it is flexible enough to consider several different waiting functions f, and we consider several examples in section 2.4.

To analyze $\{(Z_t, C_t)\}$, we form the dual process $\{(X_s, D_s)\}_{s\geq 0}$, where, as described above, the dual is progressing *backwards* in time relative to $\{(Z_t, C_t)\}$. Fix a time $t_0 > 0$. Again, $\{X_s(v_i)\}_{s\geq 0}^{s\leq t_0}$ traces where v_i's opinion at time t_0 came from, and so we have

$$\begin{aligned} Z_{t_0}(v_i) &= Z_{t_0-s}(X_s(v_i)), \text{ and} \\ Z_{t_0}(v_i) &= Z_0(X_{t_0}(v_i)), \text{ by setting } s = t_0. \end{aligned}$$

The process $\{X_s(v_i)\}_{s\geq 0}$ is a random walk on G that moves to a neighbor $N_m(v_i)$ chosen uniformly at random at the discrete times $\{S_m(v_i)\}_{m=0}^{\infty}$, where $S_0(v_i)$ is defined to be 0. In keeping with our view of the dual progressing backwards in time, the variable $D_s(v_i)$ is defined to be the confidence level $C_{t_0-s}(X_s(v_i))$ of the vertex $X_s(v_i)$. The distribution of $S_{m+1}(v_i) - S_m(v_i)$ is dependent on $c := D_s(v_i)$, for $S_m(v_i) \leq s < S_{m+1}(v_i)$. Note that c is constant for s in this interval. As before, since the exponential distribution is independent of the direction of time and is memoryless, $S_{m+1}(v_i) - S_m(v_i)$ is exponentially distributed with mean $f(c) - f(c-1)$ if c is positive, and exponentially distributed with mean $f(0)$ if $c = 0$.

Note that $D_s(v_i)$ updates only when $X_s(v_i)$ updates, and that when $X_s(v_i)$ does not leave the current trusted component $K(X_s(v_i))$, then $D_s(v_i)$ simply decrements by 1. In fact, $X_s(v_i)$ does not leave the trusted component until $D_s(v_i) = 0$. The critical observation is that $D_s(v_i)$ thus counts the number of steps that $X_s(v_i)$ *will take* in the current trusted component before leaving. It is slightly disconcerting that $D_s(v_i)$ is dependent on the future (in the sense of the s time variable). However, since the random variable $N_m(v_i)$ of a randomly chosen neighbor at time $S_m(v_i)$ is independent of $N_{m'}(v_j)$ for any m' and

$v_j \neq v_i$, of $N_{m'}(v_i)$ for any $m' \neq m$, and of $S_{m'}(v_j)$ for any m' and v_j, and since $D_s(v_i)$ is dependent only on the choices $\{N_m(v_i)\}$, we will be able to analyze the situation with only a little extra difficulty. This explains our choice of the distribution of $C_0(v_i)$:

$$\Pr\{C_0(v_i) = \ell\} = \Pr\{X_s(v_i) \text{ leaves } K(v_i) \text{ in exactly } \ell \text{ steps}\}.$$

The random walks $\{X_s\}$ are again coupled in that if $X_{s'}(v_i) = X_{s'}(v_j)$ for some s', then $X_s(v_i) = X_s(v_j)$ for all $s \geq s'$. The random variables D_s also are coupled since if $X_s(v_i) = X_s(v_j)$ for all $s \geq s'$, then so $D_s(v_i) = D_s(v_j)$ for all $s \geq s'$. The system of random walks $\{X_s\}$ do coalesce into a single random walk. However, this is not immediate, and will be proved in Theorem 2.3.

From the construction of the dual, we immediately have

Proposition 2.1 (Duality). *The process* $\{(X_s, D_s)\}_{s\geq 0}$ *is dual to* $\{(Z_t, C_t)\}_{t\geq 0}$ *in the sense that*

$$\Pr\{Z_{t_0}(v_i) = 1 \mid Z_0\} = \Pr\{Z_0(X_{t_0}(v_i)) = 1 \mid Z_0\}, \textit{ and so}$$

$$\lim_{t_0\to\infty} \Pr\{Z_{t_0}(v_i) = 1 \mid Z_0\} = \lim_{s\to\infty} \Pr\{Z_0(X_s(v_i)) = 1 \mid Z_0\},$$

if at least one of the limits exists.

Definition 2.2. Let $\{\widehat{X}_m(v_i)\}_{m=0}^{\infty}$ be the discrete-time Markov chain given by $\widehat{X}_m(v_i) = X_{S_m(v_i)}(v_i)$. Known as the *embedded chain* or *jump chain* for the process $\{(X_s(v_i), D_s(v_i)\}$, $\{\widehat{X}_m(v_i)\}$ is defined by the state transition probabilities without respect to the holding times at each state. In our case, $\{\widehat{X}_m(v_i)\}$ is a standard discrete-time random walk on G; that is, $\{\widehat{X}_m(v_i)\}_{m=0}^{\infty}$ is a random walk that moves to a neighbor chosen uniformly at random at each discrete time unit.

Theorem 2.3 (Coalescence). *If* Z_0 *is not a uniform opinion state, then the system* $\{(X_s, D_s)\}_{s\geq 0}$ *of random walks almost surely coalesces to a single random walk; that is, with probability* 1 *there exists a time* S_c *such that for all* $s \geq S_c$ *and any* v_i *and* v_j, $X_s(v_i) = X_s(v_j)$.

Proof. For the sake of clarity and simplicity of notation, we will refer to the random variables $X_s(v_i)$, $X_s(v_j)$, and $\widehat{X}_m(v_i)$ as X_s, X'_s, and $\widehat{X}_m$, respectively. For the rest of the proof,

v_i and v_j will simply be arbitrary vertices, and not necessarily the starting vertices of the random walks.

If G is connected and non-bipartite, then the state space of $\{\widehat{X}_m\}$ contains cycles of both even and odd length. By Corollary 1 of section 5.6 of [26], $\{\widehat{X}_m\}$ is a regular Markov chain. Thus, there exists an integer M such that

$$\Pr\{\widehat{X}_m = v_k \mid \widehat{X}_0 = v_i\} > \varepsilon \tag{2.4}$$

for any $m \geq M$, any v_i and v_k, and for some positive constant ε. Let $\mathcal{E}_{i,k,s'}$ be the event that the number of updates of X_s (or "steps" of $\widehat{X}_m$) that occur between $s = 0$ and $s = s'$ is at least M, given that $X_0 = v_i$ and that $X_{s'} = v_k$. Because Z_0 is not a uniform opinion state, $K(v_i)$ is not all of G. Thus, there exists a sequence of at most $2M$ steps of $\widehat{X}_m$ from v_i to v_k to some vertex $v_j \notin K(v_k)$ such that the sequence leaves $K(v_i)$. If v_i and v_k are in the same trusted component, then one possibility is that the sequence starts at v_i, visits v_k, and then exits $K(v_i)$ to v_j. If v_i and v_k are in different trusted components, then the sequence could start at v_i, leave $K(v_i)$ to visit v_k, and then exit $K(v_i)$ to v_j. Note that v_j is the first vertex visited when $\widehat{X}_m$ exits $K(v_k)$. Since the sequence leaves both $K(v_i)$ and $K(v_k)$, the confidence level at each step is bounded by $2M$, and so the time to take each of these steps has finite expectation bounded by $f(2M) - f(2M-1)$. Thus, there exists a time $S_{i,k}$ and a positive constant $\delta_{i,k}$, both dependent on v_i and v_k, such that the probability of $\mathcal{E}_{i,k,S_{i,k}}$ is greater than $\delta_{i,k}$. Let S be the maximum of $S_{i,k}$ over all i and k, and similarly let δ be the minimum of $\delta_{i,k}$ over all i and k. Then $\mathcal{E}_{i,k,S}$ has probability greater than δ, for any choice of v_i and v_k. Thus,

$$\begin{aligned}
&\Pr\{X_S = v_k \mid X_0 = v_i\} \\
&= \sum_{m=0}^{\infty} \Pr\{X_s \text{ takes } m \text{ steps during } s \in (0, S]\} \Pr\{\widehat{X}_m = v_k \mid \widehat{X}_0 = v_i\} \\
&\geq \sum_{m=M}^{\infty} \Pr\{X_s \text{ takes } m \text{ steps during } s \in (0, S]\} \Pr\{\widehat{X}_m = v_k \mid \widehat{X}_0 = v_i\} \\
&> \sum_{m=M}^{\infty} \Pr\{X_s \text{ takes } m \text{ steps during } s \in (0, S]\}\varepsilon \qquad \text{by (2.4),} \\
&> \delta\varepsilon,
\end{aligned}$$

and so

$$\Pr\{X_S = v_k \mid X_0 = v_i\} > \delta\varepsilon. \tag{2.5}$$

If G is bipartite, we need to consider the parity of the number of steps taken. By Theorem 5.13 of [26], $\{\widehat{X}_m\}$ is an ergodic Markov chain of period two. Thus, there exists an integer M such that

$$v_i, v_k \text{ in the same partite set} \Rightarrow \Pr\{\widehat{X}_m = v_k \mid \widehat{X}_0 = v_i\} > \varepsilon \quad \text{if } m \text{ is even,} \tag{2.6}$$

and

$$v_i, v_k \text{ in different partite sets} \Rightarrow \Pr\{\widehat{X}_m = v_k \mid \widehat{X}_0 = v_i\} > \varepsilon \quad \text{if } m \text{ is odd,} \tag{2.7}$$

where $m \geq M$ and ε is some positive constant. Let $\mathcal{E}^{\mathrm{e}}_{i,k,s'}$ be the event that the number of updates of X_s (or "steps" of $\widehat{X}_m$) that occur between $s = 0$ and $s = s'$ is at least M and of the appropriate parity, given that $X_0 = v_i$ and that $X_{s'} = v_k$. Because Z_0 is not a uniform opinion state, $K(v_i)$ is not all of G. Thus, there exists a sequence of at most $2M + 2$ steps of $\widehat{X}_m$ from v_i to v_k to some vertex $v_j \notin K(v_k)$ such that the sequence leaves $K(v_i)$. As before, if v_i and v_k are in the same trusted component, then one possibility is that the sequence starts at v_i, visits v_k, and then exits $K(v_i)$ to v_j. If v_i and v_k are in different trusted components, then the sequence could start at v_i, leave $K(v_i)$ to visit v_k, and then exit $K(v_i)$ to v_j. Note that v_j is the first vertex visited when $\widehat{X}_m$ exits $K(v_k)$. Since the sequence leaves both $K(v_i)$ and $K(v_k)$, the confidence level at each step is bounded by $2M + 2$, and so the time to take each of these steps has finite expectation bounded by $f(2M + 2) - f(2M + 1)$. Hence, there exists an $S_{i,k}$ and a positive constant $\delta_{i,k}$, both dependent on v_i and v_k, such that the probability of $\mathcal{E}^{\mathrm{e}}_{i,k,S_{i,k}}$ is greater than $\delta_{i,k}$. Let S be the maximum of $S_{i,k}$ over all i and k, and similarly let δ be the minimum of $\delta_{i,k}$ over all i

and k. Then $\mathcal{E}^{\mathrm{e}}_{i,k,S}$ has probability greater than δ, for any choice of v_i and v_k. Thus,

$$\begin{aligned}
&\Pr\{X_S = v_k \mid X_0 = v_i\} \\
&= \sum_{m=0,\ m \text{ even}}^{\infty} \Pr\{X_s \text{ takes } m \text{ steps during } s \in (0, S]\} \Pr\{\widehat{X}_m = v_k \mid \widehat{X}_0 = v_i\} \\
&\geq \sum_{m=M,\ m \text{ even}}^{\infty} \Pr\{X_s \text{ takes } m \text{ steps during } s \in (0, S]\} \Pr\{\widehat{X}_m = v_k \mid \widehat{X}_0 = v_i\} \\
&> \sum_{m=M,\ m \text{ even}}^{\infty} \Pr\{X_s \text{ takes } m \text{ steps during } s \in (0, S]\}\varepsilon \qquad \text{by (2.6)}, \\
&> \delta\varepsilon,
\end{aligned}$$

and so

$$\Pr\{X_S = v_k \mid X_0 = v_i\} > \delta\varepsilon. \tag{2.8}$$

Similarly, if v_i and v_k are in different partite sets, then the same proof works by considering only an odd number of steps.

Note that the bounds (2.5), (2.8), and the analogous bound when v_i and v_k are in different partite sets are time translation invariant, in the sense that for any $s > 0$,

$$\Pr\{X_{S+s} = v_k \mid X_s = v_i\} > \delta\varepsilon.$$

This is because X_s is independent of the past.

We now consider a second random walk $\{X'_s\}$. We wish to determine the probability of X_s and X'_s hitting; that is, of there existing a time S_c such that $X_{S_c} = X'_{S_c}$. Let S_h be the first time when $X_s = X'_s$ if X_s and X'_s hit, and let $S_h = \infty$ if X_s and X'_s never hit. Given that $X_0 = v_i$ and $X'_0 = v_j$, the probability that S_h is at most the parameter S from above is exactly equal to the probability that X_S and X'_S both equal a vertex v_k, for some vertex v_k. Thus, we have

$$\begin{aligned}
\Pr\{S_h \leq S \mid X_0 = v_i, X'_0 = v_j\} &= \sum_{k=1}^{n} \Pr\{X_S = v_k \mid X_0 = v_i\} \Pr\{X'_S = v_k \mid X'_0 = v_j\} \\
&> \sum_{k=1}^{n} (\delta\varepsilon) \Pr\{X'_S = v_k \mid X'_0 = v_j\} \qquad \text{by (2.5) and (2.8)}, \\
&= \delta\varepsilon.
\end{aligned}$$

This result is also time translation invariant, in that

$$\Pr\{S_h \leq S + s \mid X_s = v_i, X'_s = v_j\} > \delta\epsilon$$

for any $s > 0$.

We now consider what happens in q consecutive time intervals each of length S. The probability that X_s and X'_s do not hit in any of the intervals is at most $(1 - \delta\varepsilon)^q$, and so

$$\Pr\{S_c \leq Sq \mid X_0 = v_i, X'_0 = v_j\} \geq 1 - (1 - \delta\varepsilon)^q \to 1 \text{ as } q \to \infty.$$

Thus, with probability 1, the random walks X_s and X'_s hit and so coalesce. Applying this result to all n random walks in the system $\{(X_s, D_s)\}$, we have that with probability 1, the system $\{(X_s, D_s)\}$ coalesces into a single random walk. $\square$

Note that Theorem 2.3 trivially implies that $\{D_s\}_{s\geq 0}$ also coalesces, *i.e.*, that $D_s(v_i) = D_s(v_j)$ for all $s \geq S_c$. This is true because $D_s(v_i) = D_s(v_j)$ whenever $X_s(v_i) = X_s(v_j)$.

Since the system coalesces, we will now drop the vertex name and only write X_s, D_s, and $\widehat{X}_m$. We wish to calculate $\lim_{s\to\infty} \Pr\{Z_0(X_s) = 1 \mid Z_0\}$; what we will actually calculate is the more detailed $\lim_{s\to\infty} \Pr\{X_s \in K(v_i) \mid Z_0\}$ for vertex v_i. Note that because of our setup it would be very difficult to calculate $\lim_{s\to\infty} \Pr\{X_s = v_i \mid Z_0\}$ as is done in the original voter model.

To analyze the limiting probability distribution of X_s we create a new semi-Markov process $\widetilde{X}_s$ that captures the movement of X_s between different trusted components. Since this only makes sense when there are different trusted components, we henceforth assume that the initial set of opinions Z_0 is not uniform.

Definition 2.4. The continuous-time process $\{\widetilde{X}_s\}_{s\geq 0}$ is defined on the state space of vertices $v_1, \ldots, v_n$ of G, where we assume that Z_0 is not uniform. Given that $X_s \in K(v_i)$, let $\widetilde{X}_s$ be the vertex v_j in $K(v_i)$ that X_s first visited when entering $K(v_i)$. Formally, let s' be the least time such that $X_{s''} \in K(v_i)$ for all $s' \leq s'' \leq s$. Then let $\widetilde{X}_s = X_{s'}$.

Since X_s is dependent only on its immediate future that is spent in the current trusted component before leaving, $\widetilde{X}_s$ is a semi-Markov process. The transition probability p_{ik} from v_i to v_k is the probability that X_s first exits $K(v_i)$ to v_k given that $X_0 = v_i$. Note that

$p_{ik} = 0$ if $v_k \in K(v_i)$ or if v_k is not a neighbor to any vertex in $K(v_i)$. The holding time H_i of $\widetilde{X}_s$ at v_i is the time X_s takes to leave $K(v_i)$, given that $X_0 = v_i$. [2] Note also that the definition of C_0 as a random variable representing a typical value of C_t allows us to form the semi-Markov process without needing to condition on the initial condition C_0.

Lemma 2.5. *The process* $\{\widetilde{X}_s\}_{s\geq 0}$ *is dual to* $\{X_s\}_{s\geq 0}$ *in the sense that*

$$\Pr\{X_s \in K(v_i) \mid Z_0\} = \Pr\{\widetilde{X}_s \in K(v_i) \mid Z_0\}, \textit{ and so}$$
$$\lim_{s\to\infty} \Pr\{X_s \in K(v_i) \mid Z_0\} = \lim_{s\to\infty} \Pr\{\widetilde{X}_s \in K(v_i) \mid Z_0\},$$

if at least one of the limits exists.

Proof. Note that if $X_s \in K(v_i)$, then $\widetilde{X}_s = v_j$ for some $v_j \in K(v_i)$. Then

$$\begin{aligned}\Pr\{X_s \in K(v_i) \mid Z_0\} &= \Pr \bigcup_{v_j\in K(v_i)} \{X_s \in K(v_i) \text{ and } \widetilde{X}_s = v_j \mid Z_0\} \\ &= \Pr \bigcup_{v_j\in K(v_i)} \{\widetilde{X}_s = v_j \mid Z_0\} \qquad \text{since } \widetilde{X}_s = v_j \text{ implies } X_s \in K(v_i), \\ &= \Pr\{\widetilde{X}_s \in K(v_i) \mid Z_0\}.\end{aligned}$$

The last statement of the proposition follows by taking the limit as $s \to \infty$ of both sides of the above equation. □

If we know the limiting probability of $\widetilde{X}_s$, then summing these probabilities over all vertices with initial opinion 1, we have

$$\begin{aligned}\sum_{v_i:Z_0(v_i)=1} \lim_{s\to\infty} \Pr\{\widetilde{X}_s = v_i \mid Z_0\} &= \sum_{v_i:Z_0(v_i)=1} \lim_{s\to\infty} \Pr\{X_s \in K(v_i) \mid Z_0\} \\ &= \lim_{s\to\infty} \Pr\{Z_0(X_s) = 1 \mid Z_0\}\end{aligned}$$

where the last expression is the quantity we are interested in. The theory of semi-Markov processes enables us to calculate the limiting probability distribution of $\widetilde{X}_s$, if the expected holding time $\mathrm{E}[H_i]$ is finite for each i (see [13]). Let $\{\widehat{\widetilde{X}}_r\}_{r=0}^{\infty}$ be the embedded Markov

---

[2] Since X_s is independent of its past, the transition probability p_{ik} and the holding time H_i are the same if the definitions condition on $X_{s'} = v_i$ for some specific time s'. However, for convenience, we simply shift time so that $s' = 0$.

chain for $\widetilde{X}_s$. Then

$$\lim_{s\to\infty} \Pr\{\widetilde{X}_s \in K(v_i) \mid Z_0\} = \frac{\displaystyle\sum_{v_j \in K(v_i)} \left(\mathrm{E}[H_j] \lim_{r\to\infty} \Pr\{\widehat{\widetilde{X}}_r = v_j\} \right)}{\displaystyle\sum_{k=1}^{n} \mathrm{E}[H_k]}. \tag{2.9}$$

Note that $\{\widehat{\widetilde{X}}_r\}$ is defined by the transition probabilities p_{ik}, which in turn can be determined from the Markov chain $\{\widehat{X}_m\}$. Since $\{\widehat{X}_m\}$ ignores waiting times, $\{\widehat{X}_m\}$ and hence $\{\widehat{\widetilde{X}}_r\}$ is independent of the choice of the waiting function f. To explicitly calculate p_{ik}, let A denote the transition matrix of $\widehat{X}_m$ on G, and let A_H denote the restriction of A to a subgraph H of G. We denote the complement in G of a subgraph H by $\overline{H}$. Let $A_{H\times\overline{H}}$ denote the restriction of A to transitions from H to $\overline{H}$. Given that $X_0 = v_i$, we wish to compute the probability that X_s exits $K(v_i)$ to vertex $v_k \in \overline{K(v_i)}$. This probability is exactly the probability that, given that $\widehat{X}_0 = v_i$, $\widehat{X}_m$ leaves $K(v_i)$ by going to vertex v_k. Using Markov chain theory (see, for example, chapter 5 of [26]), the matrix

$$B = (I - A_{K(v_i)})^{-1} A_{K(v_i)\times\overline{K(v_i)}}$$

gives the exit probabilities of $\widehat{X}_m$ from $K(v_i)$. That is, if $B = [b_{jk}]$, then

$$b_{jk} = \Pr\{\widehat{X}_m \text{ exits } K(v_i) \text{ to } v_k \in \overline{K(v_i)} \mid \widehat{X}_0 = v_j \in K(v_i)\}.$$

We also need to calculate the expected holding time $\mathrm{E}[H_i]$, which is just the expected time s when X_s exits $K(v_i)$, given that $X_0 = v_i$. Recall that the ik^{th} entry of $A^d_{K(v_i)} A_{K(v_i)\times\overline{K(v_i)}}$ gives the probability that $\widehat{X}_m$ exits $K(v_i)$ to $v_k \in \overline{K(v_i)}$ in exactly $d+1$ steps, given that $\widehat{X}_0 = v_i$. Thus,

$$\begin{aligned}
&\mathrm{E}[\text{time for } X_s \text{ to exit } K(v_i) \text{ to } v_k \mid X_0 = v_i \text{ and } X_s \text{ exits to } v_k] \\
&= \sum_{d=0}^{\infty} \Big(\mathrm{E}[\text{time for } X_s \text{ to exit } K(v_i) \mid \widehat{X}_m \text{ exits to } v_k \text{ in } d+1 \text{ steps and } \widehat{X}_0 = v_i] \times \\
&\qquad \Pr\{\widehat{X}_m \text{ exits to } v_k \text{ in } d+1 \text{ steps} \mid \widehat{X}_0 = v_i\} \Big)
\end{aligned}$$

The value of D_s determines the distribution of the holding time of X_s at v_i: D_s is the number of steps that will be taken in $K(v_i)$ before exiting $K(v_i)$. Let $v_i = u_0, u_1, \ldots, u_{d+1} = v_k$ be the path taken by $\widehat{X}_m$ in $K(v_i)$ before exiting $K(v_i)$ to v_k. The holding time of X_s at v_i is thus independent of v_i, v_k, and the sequence $u_1, \ldots, u_d$ of steps taken within $K(v_i)$ from v_i to v_k. Thus,

$$
\begin{aligned}
&\mathrm{E}[\text{time for } X_s \text{ to exit } K(v_i) \mid \widehat{X}_m \text{ exits in } d+1 \text{ steps to } v_k \text{ and } \widehat{X}_0 = v_i] \\
&= \mathrm{E}[\text{time for } X_s \text{ to exit } K(v_i) \mid \widehat{X}_m \text{ exits in } d+1 \text{ steps}] \\
&= \sum_{q=0}^{d} \mathrm{E}[\text{holding time for } X_s \text{ at } u_q \mid \widehat{X}_m \text{ exits } K(v_i) \text{ in } d+1-q \text{ steps}] \\
&\qquad\qquad\qquad\qquad\qquad\qquad (\textit{i.e.}, \text{ conditioned on } D_s = d+1-q) \\
&= f(0) + [f(1) - f(0)] + [f(2) - f(1)] + \ldots + [f(d) - f(d-1)] \\
&= f(d),
\end{aligned}
$$

where the second to last equality follows from how the confidence level and f affects the time between updates. Thus, the expected time until X_s exits $K(v_i)$, given that $X_0 = v_i$ and that $X_s = v_k$, is

$$
\begin{aligned}
&\mathrm{E}[\text{time for } X_s \text{ to exit } K(v_i) \mid \widehat{X}_m \text{ exits to } v_k \text{ and } \widehat{X}_0 = v_i] \\
&= \sum_{d=0}^{\infty} \mathrm{E}[\text{time for } X_s \text{ to exit } K(v_i) \mid \widehat{X}_m \text{ exits in } d+1 \text{ steps to } v_k \text{ and } \widehat{X}_0 = v_i] \times \\
&\qquad\qquad \Pr[X_s \text{ exits } K(v_i) \text{ in } d+1 \text{ steps} \mid \widehat{X}_m \text{ exits to } v_k \text{ and } \widehat{X}_0 = v_i] \\
&= \sum_{d=0}^{\infty} f(d) \frac{\Pr[X_s \text{ exits } K(v_i) \text{ in } d+1 \text{ steps to } v_k \mid \widehat{X}_0 = v_i]}{\Pr[\widehat{X}_m \text{ exits to } v_k \mid \widehat{X}_0 = v_i]} \\
&= \frac{1}{b_{ik}} \left(ik^{\text{th}} \text{ entry of } \sum_{d=0}^{\infty} f(d) A^d_{K(v_i)} A_{K(v_i) \times \overline{K(v_i)}} \right).
\end{aligned}
$$

If the series is convergent, then the expected time is finite, and formula (2.9) for the limiting distribution of $\widetilde{X}_s$ holds. In the next section, we will consider specific choices for the waiting function f and examine how the process $\{(X_s, D_s)\}$ behaves with each choice.

2.4 Specific Choices for the Waiting Function

We first need a standard result about power series of matrices.

Lemma 2.6. *Let* $\sum_{d=0}^{\infty} f(d)x^d$ *be a power series in the complex variable* x *with radius of convergence* R. *Then* $\sum_{d=0}^{\infty} f(d)M^d$ *is a convergent power series in the matrix* M *if all of* *the eigenvalues of M have modulus less than R. Furthermore, if $r(x)$ is a rational function in x such that $r(x) = \sum_{d=0}^{\infty} f(d)x^d$ for all $|x| < R$, then $r(M) = \sum_{d=0}^{\infty} f(d)M^d$ if all of the eigenvalues of M have modulus less than R.*

This result follows from the fact that a power series formula r clearly holds for diagonalizable matrices, and the diagonalizable matrices are dense in the space of all matrices.

Definition 2.7. Let G^1 denote the subgraph of G induced by vertices v_i whose initial opinions $Z_0(v_i)$ are 1, and similarly define G^0. Let A denote the transition matrix of $\widehat{X}_m$ on G, and let A_H denote the restriction of A to a subgraph H of G. For each connected component K of G^1 or G^0, let $e(K)$ denote the largest modulus of eigenvalues of A_K. Define $\mu(G, Z_0)$ to be the maximum $e(K)$ taken over all components K of G^1 and G^0. If every component that attains the maximum $\mu(G, Z_0)$ has the same initial opinion, then define $\mu'(G, Z_0)$ to be the maximum $e(K)$ taken over all components K that have the opposite opinion.

Lemma 2.8. *For any non-uniform set of initial opinions Z_0, $\mu(G, Z_0) < 1$.*

Proof. This statement follows immediately from the fact that $\widehat{X}_m$ will leave each trusted component with probability 1. ☐

We now consider several specific waiting functions, and show how the voter model with confidence levels behaves with these choices. To make the examples concrete, we calculate $\lim_{s\to\infty} \Pr\{Z_t(v_i) = 1 \mid Z_0\}$ for the graph G shown in Figure 2.2 and for various parameters of the waiting functions. In the figure, vertices with initial opinion $Z_0(v_i) = 1$ are marked with black dots, while vertices with initial opinion 0 are marked with hollow dots. We will refer to the three trusted components as $K(v_1)$, $K(v_2)$, and $K(v_6)$. The transition matrix

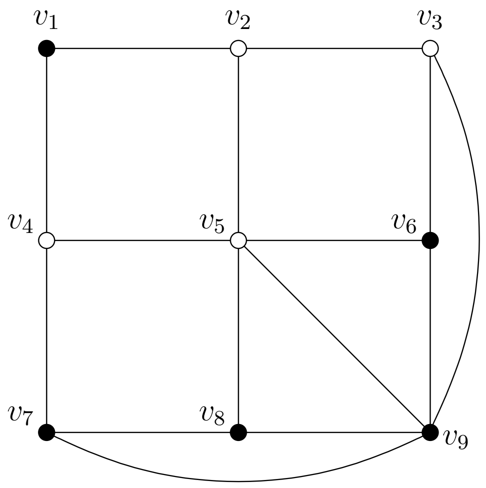

Figure 2.2: Initial opinions Z_0 on G. Vertices with initial opinion $Z_0(v_i) = 1$ are marked with black dots, while vertices with initial opinion 0 are marked with hollow dots. We refer to the three trusted components as $K(v_1)$, $K(v_2)$, and $K(v_6)$.

A of $\widehat{X}_m$ on G is

$$A = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 0 & 1/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 \\ 0 & 1/5 & 0 & 1/5 & 0 & 1/5 & 0 & 1/5 & 1/5 \\ 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 1/3 & 1/3 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 & 0 \end{bmatrix}.$$

It is interesting to note the behavior of the classic voter model, where

$$\lim_{t\to\infty} \Pr\{Z_t(v_i) = 1 \mid Z_0\} = \frac{16}{30} \approx .5333$$

from equation (2.2).

Lemma 2.6 enables us to calculate the limiting probability for any polynomial waiting function. We present one specific class of polynomials here.

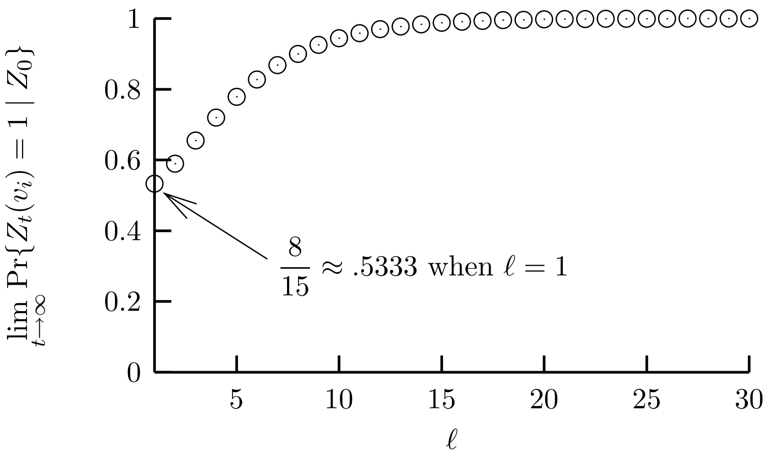

Figure 2.3: $\Pr\{Z_t$ enters the uniform 1 opinion state $\mid Z_0\}$ as a function of ℓ.

Example 2.9. Let f be the polynomial $f(d) = \binom{d+\ell}{\ell}$, where ℓ is a positive integer. Note that

$$\sum_{d=0}^{\infty} \binom{d+\ell}{\ell} x^d = \frac{1}{(1-x)^{\ell+1}}$$

by integrating the power series ℓ times. This series is convergent for $|x| < 1$, and so $\sum_{d=0}^{\infty} \binom{d+\ell}{\ell} A_{K(v_i)}^d = (I - A_{K(v_i)})^{-(\ell+1)}$ holds for any component $K(v_i)$. Figure 2.3 shows the limiting distribution for $\ell = 1, \ldots, 30$ for our example graph G. Note that when $\ell = 1$, $f(d) = d + 1$, and the voter model with confidence levels reduces to the special case of the classic voter model. Thus, at $\ell = 1$, $\lim_{s\to\infty} \Pr\{Z_t(v_i) = 1 \mid Z_0\} = 8/15$.

Example 2.10. Let f be the exponential function $f(d) = \lambda^d$, where $\lambda > 1$. If $\lambda < 1/\mu(G, Z_0)$, then all of the eigenvalues of $\lambda A_{K(v_i)}$ are less than 1 in modulus for any component $K(v_i)$. Thus, $\sum_{d=0}^{\infty} \lambda^d A_{K(v_i)}^d$ is a convergent power series, and, using the formula for a convergent geometric series, is equal to $(I - \lambda A_{K(v_i)})^{-1} A_{K(v_i)\times\overline{K(v_i)}}$. For $\lambda \geq 1/\mu(G, Z_0)$, it is not clear that the limiting probability exists, since the theory of semi-Markov processes in general does not deal with infinite expected holding times. It is an open problem to determine the behavior beyond the threshold $1/\mu(G, Z_0)$.

When $\lambda = 1$, $f(d)$ is a constant function. Thus, the expected time for X_s to leave a trusted component is 1, regardless of the number of steps taken. The limiting probability reduces to the stationary distribution of the embedded Markov chain $\widehat{\widehat{X}}_r$ of the semi-Markov

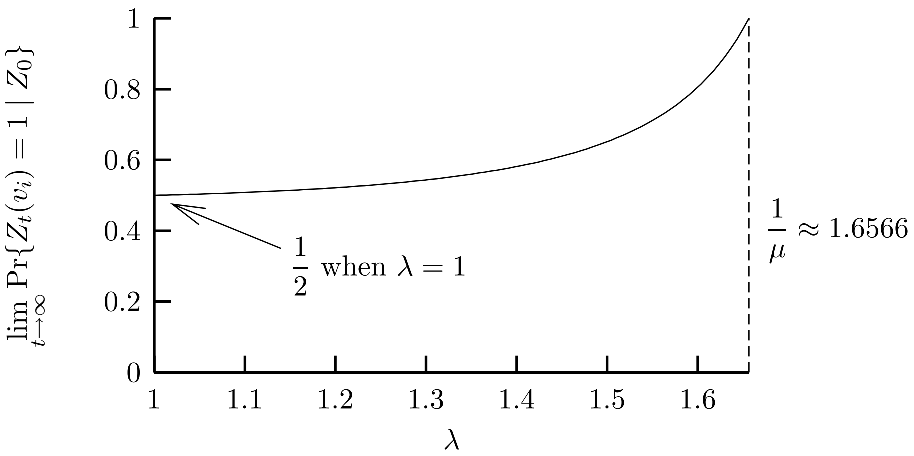

Figure 2.4: $\Pr\{Z_t$ enters the uniform 1 opinion state $\mid Z_0\}$ as a function of λ.

process $\widetilde{X}_s$. Since there are two opinions, the underlying graph of $\widehat{\widetilde{X}}_r$ is always a bipartite graph, and hence $\lim_{t\to\infty} \Pr\{Z_t(v_i) = 1 \mid Z_0\} = 1/2$.

For our example graph G, the largest eigenvalue in modulus of $A_{K(v_1)}$, $A_{K(v_2)}$, and $A_{K(v_6)}$ is $\mu(G, Z_0) \approx .6037$ from $A_{K(v_6)}$. We can thus calculate $\lim_{s\to\infty} \Pr\{Z_t(v_i) = 1 \mid Z_0\}$ for $1 < \lambda < 1/\mu(G, Z_0) \approx 1.6566$ using the geometric series formula.

We can also create a waiting function that is increasing with d yet bounded.

Example 2.11. Let f be the function $f(d) = 2 - \theta^d$, for $0 < \theta < 1$. Here 2 is an arbitrary constant chosen for concreteness. Since

$$\sum_{d=0}^{\infty}(2 - \theta^d)x^d = \frac{2}{1-x} - \frac{1}{1-\theta x}$$

for $|x| < 1$, $\sum_{d=0}^{\infty}(2 - \theta^d)A_{K(v_i)}^d$ is convergent for any component $K(v_i)$. The limiting probability as a function of θ is shown in Figure 2.5. When θ is 0 or 1, $f(d)$ is constant, and, as seen above, the limiting probability is 1/2. Interestingly, the limiting probability is less than 1/2 for $0 < \theta < 1$, attaining a minimum at $\theta \approx .6430$. However, because the waiting function is bounded, high confidence levels do not significantly affect the limiting probability, and the limiting probability varies little for θ in the range $0 < \theta < 1$.

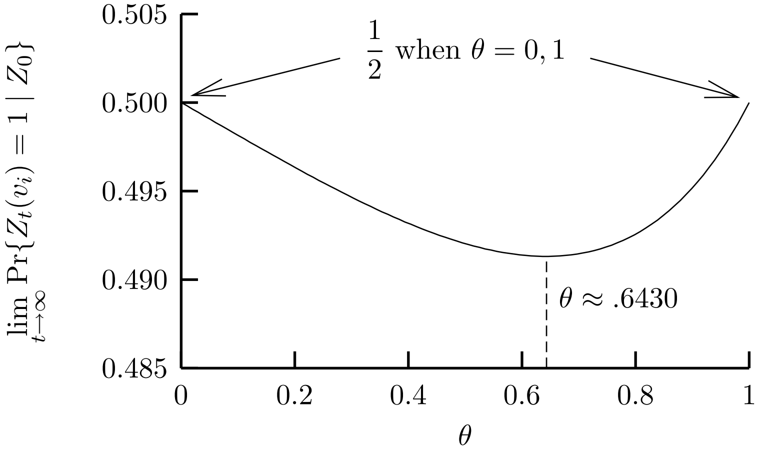

Figure 2.5: $\Pr\{Z_t$ enters the uniform 1 opinion state $\mid Z_0\}$ as a function of θ.

2.5 Conclusion

As with the classic voter model, the results we have presented on the voter model with confidence levels hold in slightly more general situations. More opinions than 0 and 1 are allowed, and the influence of a neighbor can be more general. If we allow G to be a digraph, have loops, and have weights on the edges, then we might modify the model so that when v_i updates, v_i randomly picks a neighbor v_j from $N_G^{\text{in}}(v_i)$ with probability proportional to the weight of the directed edge (v_j, v_i). In the digraph case, the digraph must be strongly connected for coalescence to occur, and the proof of Theorem 2.3 also needs to be strengthened in a straightforward way to handle larger periods in Markov chains. Also, trusted components can be any connected induced subgraphs as long as the initial opinions within the trusted components are uniform. Unfortunately, if certain features of the model are weakened, then the type of arguments we have made no longer work. The independence of the confidence level and the opinion held is absolutely crucial for the dual to be tractable. Similarly, the 1 added to the confidence level when adopting a trusted associate's opinion must remain constant for all vertices and trusted associates within a trusted component so that the expected time for X_s to leave a trusted component can be calculated. Distinguishing between trusted associates and non-trusted associates allows the process to reset itself in a Markov-like property. When treating all neighbors the same, there is no control over how fast the confidence levels are increasing.

Our model of confidence levels presented here was chosen primarily for tractability. However, there are other possibilities both for how the confidence level is updated and for how the confidence level affects the time period until a vertex updates again. Perhaps a more natural choice for how the confidence level updates is that the confidence level increases by one if v_i adopts v_j's opinion and v_j's opinion is the same as v_i's. If v_i's and v_j's opinions are different, then v_i's confidence level decreases by 1. A different model would be to reset v_i's confidence level to 0 in the latter case. Confidence levels could also be bounded above and stay at the upper bound once reached until they decrease. The confidence level could also increase or decrease by amounts determined by which neighbor the opinion was adopted from.

The choice of exponential distributions for the time periods between updates is common in stochastic models. In our model they are crucial for being able to define the holding time at each vertex for the dual random walk X_s. However, if other techniques can be used, then other distributions could be considered. The use of the differences $f(c) - f(c-1)$ is purely for notational ease and does not restrict the waiting function f. The main restriction on f from the point of view of analysis of the model is that the limiting probability of the semi-Markov process must exist. One necessary condition is that the expected holding times of $\widetilde{X}_s$ are finite. However, it seems that little is known about the existence of the limiting probability when the expected holding times are infinite. This remains an interesting open question.

One reason our modification to the voter model seems appealing is that it is a stochastic system with nontrivial dependence of the limiting behavior on the structure of G but which still remains tractable. It would be very interesting to construct other modifications of the voter model that also have this property, or to construct similar modifications of other stochastic models such as the contact process or the exclusion process (see [20]). As mentioned in the introduction, confidence levels were introduced originally in some simplistic deterministic models of opinion formulation. It would be interesting to incorporate confidence levels into other deterministic models, such as the majority process or the k-threshold process (see [8]).

Since we were motivated to consider the voter model by our interest in epidemiological models, it would be interesting to consider other modifications that are more biologically motivated. Confidence levels can be thought of as modeling an individual's resistance to infection, but perhaps there are other ways of modeling resistance. More complicating factors such as mutations and length of infection would also be interesting to consider.

References

[1] D. J. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, available from `http://www.stat.berkeley.edu/users/aldous`.

[2] S. J. Brams and P. C. Fishburn, *Approval Voting*, Birkhauser, Boston, 1983.

[3] P. Clifford and A. Sudbury, A Model for Spatial Conflict, *Biometrika*, 60 (1973), 581-588.

[4] N. C. Dalkey, The Delphi Method: An Experimental Study of Group Opinion, Memorandum RM-5888-PR, RAND Corporation, Santa Monica, June 1969.

[5] M. H. DeGroot, Reaching a Consensus, *Journal of the American Statistical Association*, 69 (1974), 167-182.

[6] J.-P. Doignon and J.C. Falmagne, Well-graded Families of Relations, *Discrete Mathematics*, 173 (1997), 35-44.

[7] P. Donnelly and D. Welsh, Finite Particle Systems and Infection Models, *Math. Proc. Comb. Phil. Soc.*, 94 (1983), 167-182.

[8] P. Dreyer, Applications and Variations of Domination in Graphs, Ph.D. dissertation, Rutgers University, 2000.

[9] R. Durrett, *Lecture Notes on Particle Systems and Percolation*, Wadsworth and Brooks/Cole, Pacific Grove, California, 1988.

[10] J.-C. Falmagne, Stochastic Token Theory, *Journal of Mathematical Psychology*, 41 (1997), 129-143.

[11] J.-C. Falmagne and M. Regenwetter, A Random Utility Model for Approval Voting, *Journal of Mathematical Psychology*, 40 (1996), 152-159.

[12] J. R. P. French, A Formal Theory of Social Power, *Psychology Review*, 63 (1956), 181-194.

[13] R. G. Gallager, *Discrete Stochastic Processes*, Kluwer Academic Publishers, Boston, 1996, 174-178.

[14] D. Griffeath, *Additive and Cancellative Interacting Particle Systems*, Lecture Notes in Mathematics, vol. 724, Springer-Verlag, Berlin, 1979.

[15] F. Hoffman, presentation at the Twenty-Sixth Southeastern International Conference on Combinatorics, Graph Theory, and Computing. Held at Florida Atlantic University in Boca Raton, Florida, March 6-10, 1995.

[16] R. Holley and T. M. Liggett, Ergodic Theorems for Weakly Interacting Systems and the Voter Model, *Ann. Probab.*, 3 (1975), no. 4, 643-663.

[17] B. Latané, The Emergence of Clustering and Correlation from Social Interaction, in: R. Hegeselmann and H. O. Peitgen, eds., *Models of Social Dynamics: Order, Chaos, and Complexity*, Holder-Pichler-Tempsky, Vienna, 1996, 79-104.

[18] B. Latané and E. Fink, Symposium: Dynamical Social Impact Theory and Communication, *Journal of Communication*, 46, no. 4 (1996), 4-77.

[19] B. Latané and A. Nowak, Attitudes as Catastrophes: From Dimensions to Categories with Increasing Involvement, in: R. R. Vallacher and A. Nowak, eds., *Dynamical Systems in Social Psychology*, Academic Press, New York, 1994, 219-249.

[20] T. M. Liggett, *Interacting Particle Systems*, Grundlehren der mathematicshen Wissenschaften, no. 276, Springer-Verlag, New York, 1985.

[21] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Grundlehren der mathematicshen Wissenschaften, no. 324, Springer-Verlag, Berlin, 1999.

[22] H. A. Linstone and M. Turoff, eds., *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, MA, 1975.

[23] L. Lovász, Random Walks on Graphs: A Survey, in: *Combinatorics, Paul Erdős is Eighty*, vol. 2, Bolyai Soc. Math. Stud., 2, János Bolyai Math. Soc., Budapest, 1996.

[24] S. Merrill, *A Unified Theory of Voting: Directional and Proximity Spatial Models*, Cambridge University Press, New York, 1999.

[25] S. Poljak and M. Sûra, On Periodical Behavior in Society with Symmetric Influences, *Combinatorica*, 3 (1983), 119-121.

[26] F. S. Roberts, *Discrete Mathematical Models, with Applications to Social, Biological, and Environmental Problems*, Prentice-Hall, Upper Saddle River, New Jersey, 1976, 258-363.

[27] F. S. Roberts, unpublished notes, 1999.