Chapter 7
Voter Models, Coalescing RWs

7.1 On $\mathbb{Z}^d$, on graphs, and mean-field theory

The voter model was introduced independently by Clifford and Sudbury (1973) and Holley and Liggett (1975) on the $d$-dimensional integer lattice. It is a very simple model for the spread of an opinion and has been investigated in great detail, see Liggett’s (1999) book for a survey.

As was the case for the contact process, we construct the voter model using a graphical representation which is built using independent Poisson processes, one for each site $x$. At the times $T^n_x$, $n \geq 1$ of a rate 1 Poisson process $x$ decides to change its opinion. To do this we pick a neighbor $y^n_x$ at random, and at time $t = T^n_x$ we set $\xi_t(x) = \xi_t(y^n_x)$. This construction allows us to define a dual process that works backwards in time to determine the source of the opinion at $x$ at time $t$. To define this process we place a dot at $x$ at time $T^n_x$ and draw an arrow from $(x, T^n_x)$ to $(y^n_x, T^n_x)$. To define $\zeta_{x,t}^s$ we start with $\zeta_{x,t}^0 = x$. The process stays at $x$ until the first time $s$ that there is a dot at $x$. If this occurs at time $t - s = T^n_x$ then $\zeta_{x,t}^s = y^n_x$ and we continue to work our way down until we encounter a dot. This definition guarantees that

$$\xi_t(x) = \xi_{t-s}(\zeta_{s,t}^x) \quad (7.1.1)$$

or in words, the opinion of $x$ at time $t$ is the same as that of $\zeta_{s,t}^x$ at time $t - s$.

For fixed $x$ and $t$, $\zeta_{s,t}^x$ is a random walk that jumps at rate 1 and to a neighbor chosen at random. It should be clear from the definition that if $\zeta_{s,t}^x = \zeta_{s,t}^y$ for some $s$ then the two random walks will stay together at later times. For these reason the $\zeta_{s,t}^x$ are called coalescing random walks. The almost sure equality in (7.1.1) is convenient for establishing the equation but it is useful to (i) rewrite the equality without the superscript $t$ and (ii) as a relationship between set-valued processes. If $t < t'$ then the joint distribution of the $\zeta_{s,t}^x$, $s \leq t$ with $x \in \mathbb{Z}^d$ is the same as that of $\zeta_{s,t'}^x$, with $x \in \mathbb{Z}^d$ when $s \leq t'$, so using the Kolmogorov extension theorem there is a family of processes $\zeta_s^x$ whose joint distributions agree with $\zeta_{s,t}^x$ on $s \leq t$.

If $A = \{x : \xi_0(x) = 1\}$ then we let $\xi_t^A$ be the set of sites in state 1 at time $t$. If we start with coalescing random walk particles on sites $B$ in the dual then we let $\zeta_t^B = \cup_{x \in B} \zeta_{s,t}^x$. In
this notation the duality becomes

\[ P(\xi^A \cap B \neq \emptyset) = P(A \cap \zeta^B \neq \emptyset) \]  \hspace{2cm} (7.1.2)

Harris (1976) introduced a family of processes that he called **additive processes** which have this property. Griffeath (1978) greatly advanced the theory of these processes. In the next chapter we will see a second important example, the contact process.

If we consider the voter model on \( \mathbb{Z}^d \) with the usual nearest neighbors then as Holley and Liggett (1975) have shown the recurrence of random walks in \( d \leq 2 \) and transience in \( d > 3 \) implies

**Theorem 7.1.1.** In \( d \leq 2 \) the voter model approaches complete consensus, i.e., \( P(\xi_t(x) = \xi_t(y)) \to 1 \). In \( d \geq 3 \) if we start from product measure with density \( p \) (i.e., we assign opinions 1 and 0 independently to sites with probabilities \( p \) and \( 1 - p \)) then as \( t \to \infty \), \( \xi_t^p \) converges in distribution to \( \xi_\infty^p \), a stationary distribution with a fraction \( p \) of the sites have opinion 1.

**Proof.** In \( d \geq 2 \) recurrence of random walk implies that \( P(\zeta_t^x \neq \zeta_t^y) \to 0 \) so \( P(\xi_t(x) \neq \xi_t(y)) \to 0 \). If we let \( \xi_t^p \) be the set of sites occupied by 1’s at time \( t \) then the duality equation (7.1.2) implies that

\[ P(\xi_t^p \cap B = \emptyset) = P(\eta_0^p \cap \zeta_t^B = \emptyset) = E(1 - p)^{|\zeta_t^B|} \]

Since size of dual never increases \( |\zeta_t^B| \downarrow \) a limit.

\[ P(\xi_\infty^p \cap B = \emptyset) = \lim_{t \to \infty} E(1 - p)^{|\zeta_t^B|} \text{ exists} \]
Knowing the probabilities $P(\eta^{\infty}_{\infty} \cap B \neq \emptyset)$ determines the distribution of $\eta^{\infty}_{\infty}$. Since $\eta^{\infty}_{\infty}$ is the limit in distribution of $\eta^{p}_{t}$ it must be a stationary distribution for the voter model.

7.1. On graphs

There are two versions of the voter model, which coincide on a regular graph but can have much different behavior in general. In either each site $x$ has an opinion $\xi_t(x)$. The first description matches the one we have given above.

**Vertex voter model.** At the times $T^n_x$, $n \geq 1$ of a rate 1 Poisson process $x$ decides to change its opinion. To do this it picks a neighbor $y^n_x$ at random, and at time $t = T^n_x$ we set $\xi_t(x) = \xi_t(y^n_x)$.

**Edge voter model.** At the times $T^n_x$, $n \geq 1$ of a Poisson process rate $d(x)$, $x$ decides to change its opinion. To do this it picks a neighbor $y^n_x$ at random, and at time $t = T^n_x$ we set $\xi_t(x) = \xi_t(y^n_x)$. We call this the edge voter model because it can also be constructed by having a rate 1 Poisson process $T^{(x,y)}_n$ for each oriented edge, and set $\xi_t(x) = \xi_t(y)$.

In the first case the dual random walk jumps from $x$ to $y$ at rate $p(x, y) = 1/d(x)$ so if we let $a(x, y)$ be the adjacency matrix of the graph then

$$d(x)p(x, y) = a(x, y) = a(y, x) = d(y)p(y, x)$$

so the measure that assigns mass $d(x)$ is a reversible stationary measure. If the graph is finite we can convert it to a stationary distribution by dividing by the sum of the degrees: $\pi(x) = d(x)/D^*$ where $D^* = \sum_x d(x)$.

In the second case the rate of jumps from $x$ to $y$ and from $y$ to $x$ are equal so the dual random walk is reversible with respect to the uniform stationary distribution. In the words of Suchecki, Eguíluz and Miguel (2005): “conservation of the global magnetization.” In terms more familiar to probabilists, the number of voters with a given opinion is a time change of simple random walk and hence is a martingale. To see this note that at each edge between a 0 and 1 the 0 changes to 1 at the same rate that the 1 changes to 0.

If we consider the biased voter model in which changes from 0 to 1 are always accepted but changes from 1 to 0 occur with probability $\lambda < 1$, then the last argument shows that the number of voters with a given opinion is a time change of a biased simple random walk and hence the fixation probability for a single 1 introduced in a sea of 0’s does not depend on the structure of the graph. This is an old result of Maruyama (1970) and Slatkin (1981), which has been rediscovered by Lieberman, Hauert, and Nowak (2005).

7.1.2. Kingman’s coalescent

If we start the coalescing random walk with $k$ walkers on the complete graph each jumping at rate one then the time until the first coalescence is approximately exponential with rate $k(k-1)/n$. To see this note that our $k$ particles jump at rate $k$ and when one does it hits another particle with probability $(k-1)/n$. 

Let $T_m$ be the first time the coalescing random walk has only $m$ particles. By the calculation in the previous paragraph

$$E(T_{m-1} - T_m) = n/(m(m-1))$$

If we follow the practice in population genetics and let $\tau_m$ be the time $T_m$ measure time in units of $n$ generation then

$$E(\tau_{m-1} - \tau_m) = 1/(m(m-1)) = \frac{1}{m-1} - \frac{1}{m}$$

so telescoping the series

$$E\tau_1 = \sum_{m=2}^{n} \frac{1}{m-1} - \frac{1}{m} = 1 - 1/n$$

The interest in this process in population genetics is that it gives the genealogy of the sample of $m$ individuals in the Wright-Fisher model. As we work backwards in discrete time each individual chooses their parent at random, so two lineages will coalesce with probability $1/n$. If $k$ stays constant as $n \to \infty$ the genealogy of the sample converges to the continuous time coalescent. For more on this see Kingman (1982)

The coalescent on the complete graph is sometimes called the mean-field version of the system because each site interacts equally with all the other. Our interest in the coalescent here stems from the fact that on graphs where the random walks are transient then the genealogy shows mean-field behavior, i.e., all pairs of particles have an equal chance of being the next coalesce.
7.2 Coalescing random walk on the torus

On a finite set the voter model will eventually reach "consensus" i.e., enter an absorbing state in which all voters have the same opinion. Cox (1989) studied the nearest neighbor voter model on a finite torus \((\mathbb{Z} \mod N)^d\). Using techniques he developed in work with David Griffeath (1986), he proved that mean-field behavior of kingman's coalescent holds for coalescing random walk on the \(d\)-dimensional torus, i.e., all pairs of particles have equal probability to be the next coalesce. He proved the result in dimensions \(d \geq 2\). The \(d = 2\) case is perhaps the most interesting, but after we state the first result we restrict our attention her to the case \(d \geq 3\).

**Theorem 7.2.1.** Let \(\xi^p_t\) denote the voter model starting from product measure with density \(p \in (0, 1)\). The time to reach consensus \(\tau_N\) satisfies

\[
\tau_N = O(s_N) \quad \text{where} \quad s_N = \begin{cases} 
N^2 & d = 1 \\
N^2 \log N & d = 2 \\
N^d & d \geq 3
\end{cases}
\]

and \(E\tau_N \sim c_d[-p \log p - (1 - p) \log(1 - p)]s_N\), where \(c_d\) is a constant that depends on the dimension. In \(d \geq 3\) the finite system looks like the stationary distribution for the infinite system at times that are large but \(o(s_N)\).

It is this result we want to generalize to our random graphs. For simplicity we will consider the initial condition in which all sites have different opinions. In this case \(\tau_N\) has the same distribution as the time for the coalescing random walk starting from all sites occupied to be reduced to one particle.

Let \(S_n\) be the random walk that jumps to nearest neighbors with equal probability, let \(p_n\) be its transition probability, and let

\[
G = \sum_{n=0}^{\infty} p_n(0, 0) = 1/\beta_d
\]

where \(\beta_d = P(S_n \neq 0 \text{ for all } n \geq 1)\). To check this note that the number of visits to 0 (including the one at time 0) has a geometric distribution with "success probability" \(\beta_d\). We have put success probability in quotes since success is failing to return to 0.

Our first step is to consider the coalescence time of two particles.

**Lemma 7.2.2.** Assume \(d \geq 3\), and let \(a_N = \log N\). Run two independent random walks on the torus. Let \(\tau^2\) be the first time they hit. Then uniformly in starting from points \(x, y\) separated by distance \(d(x, y) \geq a_N\)

\[
P(\tau^2 > tN^d) \to \exp(-2t/G)
\]
Since the torus is translation invariant, time for two walkers to hit has the same distribution as 1/2 the time $T_0$ for one walker $S_t$ to hit the origin. The first thing to do is to show:

if $b_N = B \log N$ then $P_x(T_0 \leq b_N N^2) \to 0$ uniformly for $x$ with $d(0, x) \geq a_N$.

Proof. Run the random walk until the first time it comes within distance $a_N$ of the origin. If $0 < \alpha < d-2$ then the function $1/|x|^{\alpha}$ is superharmonic for the random walk when $|S_t| \geq K_\alpha$, see e.g., Example 5.3 in Varadhan’s (2001) book. From this we see the probability $S_t$ hits 0 before going a distance $N^{1-\epsilon}$ away from it is $\leq Ca^{-\alpha}N$. Let $\epsilon < 0$.

The local central limit theorem implies that $\max_x P(S_{r(N)} = x) \leq CN^{-d(1-2\epsilon)}$

Using the Markov property and translation invariance it follows that this estimate is valid for all $t \geq r(N)$. Since a random walk that hits the origin will spend time at least one unit of time there with probability $e^{-1}$ the desired result follows.

If $B$ is large enough and $t \geq b_N N^2$

$$\max_x \sum_y |P_x(S_t = y) - N^{-d}| \leq N^{-2d} \quad (7.2.1)$$

The result now follows from Theorem 7.3.3.

The next step is to consider $k$ particles.

**Lemma 7.2.3.** Run $k \geq 3$ independent random walks starting from points separated by distances at least $a_N$. (a) If we let $\tau_{i,j}$ be the first time $i$ and $j$ hit and let $\tau^k = \min_{1 \leq i < j \leq k} \tau_{i,j}$ then

$$P(\tau^k > tN^d) \to \exp(-tk(k-1)/G) \quad (7.2.2)$$

(b) The probability that at time $\tau_{i,j}$, that there are random walks (other than $i$th and $j$th) separated by distance $\leq a_N$ tends to 0.

Proof. To prove (b), we note that by the proof of Lemma 7.2.2 the probability of a pair hitting by time $\leq b_N N^2$ tends to 0. Using (7.2.1) we conclude that two pairs of particles $\{i_1, j_1\}$ and $\{i_2, j_2\} \neq \{i_1, j_1\}$ are both within distance $a_N$ at some time $t \in [b_N N^d, N^d \log N]$ is

$$\leq k^4 \cdot N^d \log N \left(\frac{2a_N}{N^d}\right)^2 \to 0$$
To prove (a) now, we need to argue that the $\tau_{ij}$ are approximately independent. Let $\alpha_N = GN^2/2$. Let $H_t(i, j) = \{\tau_{ij} \leq t\alpha_N\}$, $F_t(i, j) = \{\tau = \tau_{ij} \leq t\alpha_N\}$, and $q(t) = P(\tau \leq t\alpha_N)$.

\[
P(H_t(i, j)) = P(F_t(i, j)) + \sum_{\{k, \ell\} \neq \{i, j\}} \int_0^{t\alpha_N} P(\tau = \tau_{k\ell} = s, \tau_{ij} \leq t\alpha_N) \, ds \tag{7.2.3}
\]

where the quantity being integrated on the right is the density of the hitting time $\tau_{k\ell}$. To evaluate the $k, \ell$ term in the sum we break things down according to the locations $X_i^\ell$ and $X_j^\ell$. By (b) we can ignore the possibility that $|X_i^\ell - X_j^\ell| < a_N$. When the distance is $\geq a_N$ the argument for Lemma 7.2.2 shows that the positions will become randomized before they hit so the hitting time will be in the limit as $N \to \infty$ exponentially distributed. Writing $\epsilon_N$ for a quantity that goes to 0 as $N \to \infty$

\[
\int_0^{t\alpha_N} P(\tau = \tau_{k\ell} = s, \tau_{ij} \leq t\alpha_N) \, ds = \int_0^{t\alpha_N} P(\tau = \tau_{k\ell} = s) \left[1 - \exp \left(-\frac{t-s}{\alpha_N}\right)\right] \, du + \epsilon_N
\]

Integrating by parts and then changing variables $u = s/\alpha_N$ and $r = t/\alpha_N$, the above is

\[
= \int_0^{t\alpha_N} \frac{1}{\alpha_N} \exp(-t/s)P(\tau = \tau_{k\ell} \leq s) \, ds + \epsilon_n
= \int_0^r \exp(-(r-u))P(\tau = \tau_{k\ell} \leq u\alpha_N) \, du + \epsilon_N
\]

Using this in our initial decomposition (7.2.3), with the convergence of the hitting time $\tau_{i,j}$ to the exponential distribution, we get

\[
1 - e^{-t} = P(F_t(i, j)) + \sum_{\{k, \ell\} \neq \{i, j\}} e^{-t} \int_0^t e^s P(F_s(k, \ell)) \, ds + \epsilon_N \tag{7.2.4}
\]

Summing over all $\binom{k}{2}$ pairs $\{i, j\}$

\[
\binom{k}{2} (1 - e^{-t}) = q(t) + \left[\binom{k}{2} - 1\right] e^{-t} \int_0^t e^s q(s) \, ds + \epsilon_N
\]

It follows [see page 365 of Cox and Griffeath (1986) for more details] that as $N \to \infty$, $q(t)$ converges to $u(t)$ the solution of

\[
\binom{k}{2} (1 - e^{-t}) = u(t) + \left[\binom{k}{2} - 1\right] e^{-t} \int_0^t e^s u(s) \, ds
\]

Multiplying both sides by $e^t$ and rearranging we have

\[
e^t u(t) - \binom{k}{2} (e^t - 1) = - \left[\binom{k}{2} - 1\right] \int_0^t e^s u(s) \, ds
\]
CHAPTER 7. VOTER MODELS, COALESCING RWS

Differentiating we have

$$e^t u(t) + e^t u'(t) - \left(\begin{array}{c} k \\ 2 \end{array}\right) e^t = - \left[ \left(\begin{array}{c} k \\ 2 \end{array}\right) - 1 \right] e^t u(t)$$

The first term on the left cancels the last term on the right. Dividing by $e^t$ and rearranging gives

$$\left(\begin{array}{c} k \\ 2 \end{array}\right) (1 - u(t)) = u'(t) = - \frac{d}{dt} (1 - u(t))$$

which has solution $1 - u(t) = \exp(-k(k-1)/2)$.

The final detail is to show that all $\binom{k}{2}$ pairs have equal probability to be the next to coalesce. To do this we go back to (7.2.4) and subtract and add the $\{i, j\}$ term in the sum to get

$$1 - e^{-t} = P(F_t(i, j)) - e^{-t} \int_0^t e^s P(F_s(i, j)) ds + e^{-t} \int_0^t e^s q(s) ds + \epsilon_N$$

Recalling $u(t) = 1 - e^{-t}$, it follows that $P(F_t(i, j))$ converges to the solution of

$$v(t) - \frac{2}{G} e^{-2t/G} \int_0^t e^{2s/G} v(s) ds = u(t) - \frac{2}{G} e^{-2t/G} \int_0^t e^{2s/G} u(s) ds$$

Since the limit is independent of $i, j$ we must have $v(t) = u(t)/\binom{k}{2}$, which completes the proof.

**Remark.** To prepare for later arguments, note that the proof of (a) only requires that $\tau_{i,j}/\alpha_N$ converges to a mean one exponential, and that the time to reach equilibrium is $o(\alpha_N)$.

Lemma 7.2.3 implies that starting from $k$ locations separated by $a_n$ the number of particles in the coalescing random walk converges to Kingman’s coalescent. However there is one small technical point remaining: we want to show that as $n \to \infty$ the coalescing random walk converges to Kingman’s coalescent starting from infinitely many particles. We will delay addressing this point until we can use a nice argument of Cooper et al to deal with it.
7.3 Using ideas from Markov chains

When the degree is not constant the random walk does not spend an equal amount of time in all parts of the space. If \( a(i, j) \) is the adjacency matrix of the graph then then the transition probability \( p(i, j) = a(i, j)/d(i) \) where \( d(i) \) is the degree of \( i \) so if we let \( \pi(i) = d(i)/D \) where \( D = \sum_i d(i) \) then the walk is reversible with respect to \( \pi \)

\[
\pi(i)p(i, j) = a(i, j)/D = a(j, i)/D = \pi(j)p(j, i)
\]

so we can take advantage of the results from Chapter 6 on rates of convergence to equilibrium.

The first step in understanding the time to reach consensus in the voter model is to pick two starting points \( x_1 \) and \( x_2 \) at random according to the stationary distribution \( \pi \) for the random walk and investigate the time it takes for their coalescing random walks to hit. Define independent continuous time random walks \( X_1^t \) and \( X_2^t \) with \( X_1^0 = x_1, X_2^0 = x_2 \), that jump at rate 1, and jump from \( i \) to \( j \) with probability \( p(i, j) \). Let \( A = \{(x, x) : x \in G\} \) be the “diagonal” and let \( T_A = \inf\{t \geq 0 : X_1^t = X_2^t\} \) be the first hitting time of \( A \) by \( (X_1^t, X_2^t) \).

### 7.3.1. Asymptotics for \( E_\pi T_A \)

Consider the discrete time version \( X_n \) of the two particle chain in which at each step we pick a particle at random and let it jump. Writing \( P_A \) for \( \pi(X_0 \in A) \), a theorem of Kac (Theorem 6.3.3 in PTE5) implies that

\[
E_A(T_A) = 1/\pi(A)
\]

As we will soon see the following situation exists in a number of examples:

(Clumping Heuristic) there is a time \( t_n = o(E_\pi T_A) \) so that if \( T_A \gg t_n \) then the two particles reach equilibrium before they hit.

In this case the expected value on \( T_A \leq t_n \) makes a negligible contribution to the expected value so the convergence

\[
1/\pi(A) \approx P_A(T_A \gg t_n) E_\pi(T_A)
\]

and we have

\[
E_\pi(T_A) = \frac{1}{\pi(A)} \cdot \frac{1}{P_A(T_A \gg t_n)} \tag{7.3.1} \text{[Kacf]}
\]

To connect with the Aldous’ clumping heuristic, we note that the naive guess for the waiting time is \( 1/\pi(A) \) but this must be corrected for by multiplying by the clump size, i.e., the expected number of hits that occur soon after the first one. In nice cases, e.g. the torus or random regular graphs, the number of hits in a clump is geometric and hence has mean \( 1/P_A(T_A \gg t_n) \). However, a geometric distribution of return times is not necessary for (7.3.1) to hold.

**Example.** \( d \)-dimensional torus, \( d \geq 3 \). \((\mathbb{Z} \mod N)^d \) has \( n = N^d \) sites. \( \pi(x, x) = N^{-2d} \) so \( \pi(A) = N^{-d} \). \( t_n = O(N^2) \). In \( d \geq 3 \), \( P_A(T_A \gg t_n) \rightarrow \beta_d = P_0(S_n \neq 0 \text{ for all } n \geq 1) \), so

\[
E_\pi T_A \sim N^d / \beta_d
\]
Example. Random $r$-regular graphs. Locally these graphs look like a tree in which each vertex has degree $r$. The probability that two random walkers that start from the origin will hit after they separate is the same as the probability that a single random walk will return to the origin, which is $1/(r-1)$. To check this note that $\phi(x) = 1/(r-1)^x$ is a harmonic function for the distance $S_n$ from the root of the tree, i.e., $\phi * S_n$ is a martingale when $S_n > 0$. Since $P_A(T_A \gg t_n) \approx (r-2)/(r-1)$, we should have

$$E_\pi T_A \sim \frac{r - 2}{r - 1} n$$  \hspace{1cm} (7.3.2)  \hspace{1cm} RRGhit

Example. Configuration model graphs. In this case,

$$\pi(A) = \sum_{i=1}^{n} \frac{d_i^2}{D^2} \quad \text{where} \quad D = \sum_{j=1}^{n} d_j$$

If the degree distribution has finite second moment then $\pi(A) = O(1/n)$. When the degree distribution has infinite variance $\pi(A)$ will go to 0 more slowly than $1/n$. Consider random graphs with a power law degree distribution $p_k \sim Ck^{-\alpha}$. When $\alpha > 3$ the distribution has finite variance. For $2 \leq \alpha \leq 3$ results of Sood and Redner (2005) for the consensus time of the voter model, suggest that

$$E_\pi T_A \approx \begin{cases} 
\frac{n}{\log n} & \alpha = 3 \\
\frac{n^{2\alpha-4}/(\alpha-1)}{\log n} & 2 < \alpha < 3 \\
(\log n)^2 & \alpha = 2
\end{cases}$$

where $\approx$ should be read “is of order.” We will later show that when $2 < \alpha \leq 3$, this also gives the asymptotic behavior of the coalescence time.

Lemma 7.3.1. When $2 \leq \alpha \leq 3$ the asymptotic behavior of $1/\pi(A)$ is given by the formulas above.

Proof. When $p_k \sim Ck^{-\alpha}$ we have

$$P(d_i > k) \sim C'k^{-(\alpha-1)} \quad \text{and} \quad P(d_i^2 > k) \sim C''k^{-(\alpha-1)/2}.$$  

To prepare for later, note that the maximum degree is $O(n^{1/(\alpha-1)})$ when $\alpha > 2$. When $\alpha = 3$, the power is $-1$ so if we use what is known about convergence to stable laws (see e.g., Theorem 3.7.2 in PTE) we have

$$a_n = \inf\{x : P(d_i^2 > x) < 1/n\} \sim C_an$$

$$b_n = nE(d_i1_{d_i^2 \leq a_n}) \sim C_bn \log n$$

so we have $(\sum_{i=1}^{n} d_i^2 - b_n)/a_n \Rightarrow \chi$ where $\chi$ is the one-sided stable law with index 1 so

$$\frac{1}{n^2} \sum_{i=1}^{n} d_i^2 \sim \frac{C_b \log n}{n}$$
When 2 < \alpha < 3, d_i^2 is in the domain of attraction of a one-sided stable law with index \((\alpha - 1)/2\) so
\[
\frac{1}{n^2} \sum_{i=1}^{n} d_i^2 \sim n^{2/(\alpha-1)}/n^2 = n^{-(2\alpha - 4)/(\alpha-1)}
\]
When \alpha = 2, the power in the last result vanishes but \(P(d_i > k) \sim Ck^{-1}\) so the argument from the case \(\alpha = 3\) again, we have \(D \sim n \log n\).

If we assume that the minimum degree is 3 then the mixing time \(t_n = O(\log n)\), while if the minimum degree is \(\leq 2\) then \(t_n = O(\log^2 n)\).

**Lemma 7.3.2.** \(P_A(T_A \gg \log n)\) stays bounded way from 0.

**Proof.** If the minimum degree is 3, then the probability the two particles leave the diagonal by different edges is \(\geq 2/3\) and the probability they never return is at least 1/2. To extend this to the general case, note that we always have a positive fraction of vertices with degree 3, and the probability of not returning to root is positive almost surely and has positive expected value. \(\square\)

### 7.3.2. Asymptotic exponential distribution

Proposition 23 of Aldous and Fill (2002) implies
\[
\sup_t |P_\pi(T_A > t) - \exp(-t/E_\pi T_A)| \leq \tau_2/E_\pi T_A \quad (7.3.3) \quad \text{exptime}
\]
where \(\tau_2\) is the relaxation time, which they define to be 1 over the spectral gap. In the random regular case if \(r \geq 3\) then \(\tau_2 \leq C \log n\) and as we have seen
\[
E_\pi T_A \sim \frac{r-1}{r-2} n \quad (7.3.4) \quad \text{htmeanasy}
\]
so the hitting time is approximately exponential.

The proof of (7.3.3) is based on a result of Mark Brown (1983) for IMRL (increasing mean residual life) distributions. If one is willing to give up on the explicit error bound, it is fairly easy to give a proof based on the idea that since convergence to equilibrium occurs much faster than the two particles hitting, then subsequential limits of \(T_A/E_\pi T_A\) must have the lack of memory property, and hence the sequence converges to a mean 1 exponential.

**Theorem 7.3.3.** If the mixing time \(t_n = o(1/\pi(A))\), and \(E_\pi T_A \leq C/\pi(A)\) then under \(P_\pi\), \(T_A/E_\pi T_A\) converges weakly to an exponential with mean 1.

**Proof.** Since \(t_n \ll 1/\pi(A) \leq E_\pi T_A\), we can find \(\epsilon_n \to 0\) with \((\epsilon_n E_\pi T_A)/t_n \to \infty\). The expected time that \(X_t^1 = X_t^2\) in \([r E_\pi T_A, (r + \epsilon_n) E_\pi T_A + 1] = \pi(A)(\epsilon_n E_\pi T_A + 1)\). If \(X^1_s = X^2_s\) then with probability \(\geq e^{-2}\) neither jumps before time \(s + 1\) and hence they agree on \([s, s+1]\). From this we conclude that
\[
P_\pi(T_A/E_\pi T_A \in [r, (r + \epsilon_n)]) \leq \delta_n = \pi(A)(\epsilon_n E_\pi T_A + 1) \to 0 \quad (7.3.5) \quad \text{intbd}
\]
CHAPTER 7. VOTER MODELS, COALESCING RWS

Notice that $\delta_n$ does not depend on $r$. Using this result with $r = s + t$ and writing $X_t = (X^1_t, X^2_t)$, $x = (x_1, x_2)$

$$P_\pi(T_A/E_\pi T_A > (s + t)) = P_\pi(T_A/E_\pi T_A > (s + t + \epsilon_n)) + \delta_n$$

$$= \sum_{x,y} P_\pi(T_A/E_\pi T_A > s, X(sE_\pi T_A) = x) P_x(X(\epsilon_n E_\pi T_A) = y) P_y(T_A/E_\pi T_A > t)$$

Subtracting $\pi(y)$ from $P_x(X(\epsilon_n E_\pi T_A) = y)$ and adding $\pi(y)$ gives two terms. The second is

$$\sum_{x,y} P_\pi(T_A/E_\pi T_A > s, X(sE_\pi T_A) = x) \pi(y) P_y(T_A/E_\pi T_A > t)$$

$$= P_\pi(T_A/E_\pi T_A > sn) P_\pi(T_A/E_\pi T_A > t)$$

The absolute value of the first term is bounded by

$$P_\pi(T_A/E_\pi T_A > s) \sup_x \sum_y |P_x(X(\epsilon_n E_\pi T_A) = y) - \pi(y)| \to 0 \quad (7.3.6)$$

due to the definition of the mixing time.

The sequence of random variables $T_A/E_\pi T_A$ has mean one, so it is tight. Let $F$ denote a subsequential limit. From the calculation above we see that if $s$, $t$, and $s + t$ are continuity points of $F$ then

$$1 - F(s + t) = (1 - F(s))(1 - F(t))$$

$F$ can have at most countably many discontinuity points, so there is a $\theta > 0$ so that $F$ is continuous at all points $m/(\theta 2^n)$ where $m$ and $n$ are positive integers. Define $\lambda$ by $e^{-\lambda} = 1 - F(1/\theta)$. It follows from the equation that if $t = m/(\theta 2^n)$ then $1 - F(t) = e^{-M}$. To conclude that $\lambda$ is independent of the subsequential limit, note that for all $n$, $P_\pi(T_A/E_\pi T_A > 2) \leq 1/2$, so using the remark about $\delta_n$ and (7.3.6), we see that if $\gamma > 0$ and $n$ is large

$$P(T_A/E_\pi T_A > 2(k + 1)) \leq P(T_A > 2k)(\gamma + P(T_A > 2)).$$

This gives an exponential bound on the tail of the distribution, which enables us to conclude that every subsequential limit has mean 1.
7.4 Cooper’s bound

To be precise I should say Cooper, Elsässer, Ono, and Radzik’s (2012) bound but that is a little much for a section title. These authors use lazy discrete time simple random walks, but the result is easier to prove and more closely related to the voter model if the walks have continuous time.

**Theorem 7.4.1.** Let $G$ be a connected graph with $n$ vertices, average vertex degree $\bar{d}$ and maximum degree $\Delta = O(n^{1-\epsilon})$. Let $\nu = (\sum_{v \in V} d^2(v))/(d^2 n)$. Let $\mathcal{C}(n)$ be the expected coalescence time for a system of particles making a continuous time random walk starting from each vertex that jumps at rate 1. Then

$$\mathcal{C}(n) = O\left(\frac{n}{\nu(1 - \lambda_1)}\right)$$

where $\lambda_1$ is the second largest eigenvalue.

**Remark.** Noting that $\nu/n = \pi(A)$, we see that if the minimum degree is 3 and hence $1 - \lambda_1$ is bounded away from 0, this result gives an upper bound for the coalescence time which is accurate in the case of power law degree distributions $p_k \sim C k^{-\alpha}$ with $\alpha > 2$. In the case of minimum degree $\leq 2$, $1 - \lambda_1 = O(1/\log n)$ so this is bigger than Sood and Redner’s correct answer by a factor of $\log n$.

7.4.1 Random walk properties

Let $G = (V, E)$ be a connected graph with $|V| = n$ and let $d(v)$ be the degree of the vertex $v$. Let $W_u(t)$ denote a continuous time random walk that starts from $u$ jumps at rate 1 and jumps from $v$ to each of its neighbors with equal probability. Let $P_{tu}(x) = P(W_u(t) = x)$. It follows from Theorem 7.4.1 that

$$|P_{tu}(x) - \pi_x| \leq (\pi_x / \nu) \frac{1}{2} e^{-(1 - \lambda_1)t} \quad (7.4.1)$$

Define the time to reach equilibrium $T_G$ so that for $t \geq T_G$

$$\max_{u, x} |P_{tu}(x) - \pi_x| = o(1/n^2) \quad (7.4.2)$$

Let $E_\pi(H_w)$ denote the expected hitting time of $w$ starting from the stationary distribution.

**Lemma 7.4.2.** Let $F$ be a graph with eigenvalue gap $1 - \lambda_1$. Then

$$E_\pi H_w \leq \frac{1}{1 - \lambda_1} \cdot \frac{1}{\pi_v}$$
CHAPTER 7. VOTER MODELS, COALESCING RWS


$$E_\pi(H_v) = Z_{v,v}/\pi_v$$ (7.4.3) \[E_{\text{EpiHv}}\]

where $Z$ is the recurrent potential kernel

$$Z_{v,v} = \int_{t=0}^{\infty} P^t_v(v) - \pi_v dt$$

They prove this result in discrete time in Lemma 11, but as they observe in Section 2.3 the result is the same in continuous time. Using (7.4.1) with $x = u = v$

$$|P^t_v(v) - \pi_v| \leq e^{-(1-\lambda_1)t}$$

Integrating gives the desired result. \[\square\]

Let $A_v(t; u)$ be the event that $W_u$ does not visit vertex $v$ in $[0, t]$.

**Lemma 7.4.3.** $P(A_v(t; u)) \leq \exp(-|t/(T_G + 3E_\pi H_v)|)$.

Proof. To simplify formulas write $T_G$ as $T$. Let $\rho = P_T^u$ be the distribution of $W_u$ after $T$ steps. Then (7.4.2) and the fact that $\pi_x \geq 1/n^2$ for any connected graph (since $1 \leq d_y \leq n-1$ for all $y \in V$) imply

$$E_\rho(H_v) = (1 + o(1))E_\pi(H_v)$$

and it follows that

$$E_u(H_v) \leq T_G + (1 + o(1))E_\pi(H_v)$$ (7.4.4) \[Erho\]

If we let $H_v(\rho)$ be the hitting time of $v$ starting from $\rho$ and $\tau = T + 3E_\pi(H_v)$ then

$$P(A_v(\tau; u)) = P(A_v(T; u), H_v(\rho) \geq 3E_\pi H_v)$$

$$\leq P(H_v(\rho) \geq eE_\pi H_v) \leq e^{-1}$$

Iterating this bound $|t/(T_G + 3E_\pi H_v)|$ times gives the desired result. \[\square\]

### 7.4.2 Multiple random walks

To study the coalescence of $k \geq 2$ walks on a graph $G = (V_G, E_G)$, we replace the $k$ walks by one walk on a new graph $Q_k$ with vertex set $V^k$. Since we work in continuous time where only one particle jumps at a time, our edge set is much simpler than the one in Cooper et al: vertices $v, w \in Q_k$ adjacent if there is a $j$ so that $v_i = w_i$ for $i \neq j$ and $\{v_j, w_j\} \in E_G$.

Although we are interested in coalescence, our $k$ random walks will be independent. For any starting positions $u = (u_1, \ldots, u_k)$ for the walks, let $M_k$ be the time of the first meeting in $G$. Let

$$S_k = \{(v_1, v_2, \ldots, v_k) : v_i = v_j \text{ for some } 1 \leq i < j \leq n\}.$$
To use results from the previous section it is convenient to contract the set $S_k$ to be a single vertex $\gamma$ making a new graph $\Gamma_k$. On contraction all edges including those that have become loops or parallel edges are retained.

Let $\pi_k$ and $\hat{\pi}_k$ be the stationary distributions on $Q_k$ and $\Gamma_k$ respectively. For any vertex $v = (v_1, v_2, \ldots, v_k) \in Q_k$, $\pi_k(v) = \pi(v_1) \cdots \pi(v_k)$. If $v \notin S_k$, $\hat{\pi}_v = \pi_v$, while $\hat{\pi}_\gamma = \sum_{x \in S} \pi_x$.

The next order of business is to find a lower bound for $\hat{\pi}_\gamma$.

\[ \text{Lemma 7.4.4.} \] There is a constant $c_k > 0$ so that if $k \leq \log^4 n$

\[ \hat{\pi}_\gamma \geq \frac{c_k k^2 \nu}{2n} \]

\textbf{Proof.} For $1 \leq x < y \leq k$ let $S_{x,y} = \{(v_1, \ldots, v_k) : v_x = v_y\}$ and note $S = \cup_{1 \leq x < y \leq k} S_{x,y}$.

\[ \pi(S_{k-1,k}) = \sum_{v_1, \ldots, v_{k-1}} \pi(v_1) \cdots \pi(v_{k-2}) \pi^2(v_{k-1}) = \sum_v \left( \frac{d(v)}{nd} \right)^2 = \frac{\nu}{n} \]

The inclusion exclusion formula implies

\[ \pi(S_k) \geq \sum_{x,y} \pi(S_{x,y}) - \sum_{(x,y) \neq (p,q)} \pi(S_{x,y} \cap S_{p,q}) \]

Taking the cases $\{x, y, p, q\} = 3, 4$ separately (the second one does not exist if $k = 3$)

\[ \pi(S_{k-2,k-1} \cap S_{k-1,k}) = \sum_v \pi(v)^3 \leq \Delta \sum_v \frac{d(v)^2}{n^3 d^3} \leq \frac{\Delta \nu}{n^2 d} \]

\[ \pi(S_{1,2} \cap S_{k-1,k}) = \sum_{u,v} \pi(u)^2 \pi(v)^2 \leq \frac{\nu^2}{n^2} \]

Combining the last two results

\[ \pi(S) \geq \left( \frac{k}{2} \right) \frac{\nu}{n} - 3 \left( \frac{k}{3} \right) \frac{\Delta \nu}{n^2 d} - 3 \left( \frac{k}{4} \right) \frac{\nu^2}{n^2} \geq \left( \frac{k}{2} \right) \frac{\nu}{n} \left[ 1 - \frac{k \Delta \nu}{nd} - \frac{k^2 \nu}{n} \right] \]

We have assumed $\Delta = O(n^{1-\epsilon})$ and $k \leq \log^4 n$, so the middle term inside the square brackets is $o(1)$. Noting that

\[ \nu \leq \Delta \sum_v \frac{d(v)}{d^2 n} = \frac{\Delta}{d} \]

the last term is also $o(1)$ and the proof is complete. 

\[ \Box \]

It follows from (7.4.4) that if $T_\Gamma$ is the time to reach equilibrium satisfying (7.4.2) then

\[ E(M_k) \leq T_\Gamma + (1 + o(1)) E_\bar{\pi}(H_\gamma) \]

To bound the first quantity we use
Lemma 7.4.5. For the random walks on $G$, $Q$, and $\Gamma$ there are mixing times

$$T_G \leq \frac{C \log n}{1 - \lambda_1(G)} \quad T_Q = O(kT_G) \quad T_\Gamma = O(kT_G)$$

Proof. The bound on $T_G$ follows from (7.4.1). In the jargon of Markov processes the random walk on $Q_k$ is the tensor product chain. The eigenvectors are the $k$-wise products of those of $G$, so $\lambda_1(Q_k) = \lambda_1(G)$. See page 168 of Levin, Peres, and Wilmer’s book for more details.

In the notation of Aldous and Fill, Chapter 3, the random walk on $\Gamma_k$ is the random walk on $Q_k$ with $S_k$ collapsed to $\gamma_k$. Their Corollary 27 implies that if a subset $A$ of vertices is collapsed to one then the second eigenvalue is not increased. We get the factor of $k$ in the bounds on the mixing times because we need to have $|P_T(x) - \pi_x| = o(1/n^2k)$ and $\pi_x/\pi_u \leq n^{2k}$ by the trivial bound we used in the proof of Lemma 7.4.3.

Lemma 7.4.6. If $k \leq \log^4 n$ then

$$E_u(M_k) = O\left(\frac{1}{1 - \lambda_1(G)} \left( k \log n + \frac{n}{\nu k^2} \right) \right)$$

Proof. Using Lemmas 7.4.2 and 7.4.4

$$E_H(H_\gamma) \leq \frac{1}{\pi(\gamma)} \cdot \frac{1}{1 - \lambda_1(\Gamma)} \leq \frac{1}{c_k k^2} \cdot \frac{n}{\nu(1 - \lambda_1(G))} \quad (7.4.6)$$

Using (7.4.4) we have $E_u(M_k) \leq O(kT_G) + (1 + o(1))E_HH_\gamma$, so the desired result follows from Lemma 7.4.5.

Summing over $k \leq \log^4 n$

$$\sum_{k=2}^{\log^4 n} E(M_k) = O\left( \frac{n}{\nu(1 - \lambda_1(G))} \right) + O\left( \frac{\log^7 n}{1 - \lambda_1(G)} \right)$$

From (7.4.5) we have $\nu \leq \Delta/d$, the first term is dominant as long as $\Delta = O(n^{1-\epsilon})$.

7.4.3 The Big Bang

The last thing to do is to estimate the time for the coalescing random walk to be reduced $n$ particles to $\log^3 n$. We prove that with high probability there cannot be a set of $k_* = \log^4 n$ particles that has not had a meeting by time $t^* = k_*^{3/2}(T_\Gamma + 3E_HH_\gamma)$ Suppose that the starting points are $v = (v_1, \ldots, v_k)$. The probability that the particles do not meet by time $t$ is the same as the probability that the random walk in $\Gamma_{k_*}$ starting from $v$ does not visit $\gamma$ by time $t^*$. By Lemma 7.4.3 this probability is

$$\leq \exp(-|t^*/(T_\Gamma + 3E_HH_\gamma)|) \leq \exp(-k_*/2) = e^{-\log^6 n}$$
The probability no set of size \( k \) does this is
\[
\leq \binom{n}{k} e^{-\log^6 n} \leq n^k e^{-\log^6 n} \leq \exp(-\log^6 n + \log^5 n)
\]

To complete the result of Theorem 7.4.1 now, we note that Lemma 7.4.5 and (7.4.6) imply
\[
t^\ast = O\left(\frac{1}{1 - \lambda_1(G)} \left( k^{5/2} \log n + \frac{n}{\nu} \right) \right)
\]
and \( k^{5/2} \log n = o(n/\nu) \) as long as \( \Delta = O(n^{1-\epsilon}) \).
7.5 Random regular graphs

Cooper, Freize and Razdik (2010) studied the behavior of multiple random walks on random regular graphs with a variety of interaction rules: independent, talkative, coalescing, annihilating, and predator-prey. Their main result for coalescing systems is the following

**Theorem 7.5.1.** Let $r \geq 3$ and $\theta_r = (r - 1)/(r - 2)$. Let $G_r$ be a random $r$-regular random graph, and start with one coalescing random walk particles at each vertex. Then the expected time $C(n)$ to coalesce to one particle

$$C(n) \sim 2\theta_r n$$

This result is false when $r = 2$ since in that case the random regular graph consists of a collection of circles. We will prove their result for continuous time walks with 2 replaced by 1. In addition, we will prove that when time is scaled correctly, the number of coalescing random walks converges to Kingman’s coalescent.

The first step is to show that locally a random $r$-regular graph looks like a tree. Part (i) of this lemma can be found in many places, see e.g., Lemma 2.1 in Lubetsky and Sly (2010). The version included here is a generalization Lemma 2.1 in Chatterjee and Durrett (2013).

**Lemma 7.5.2.** Suppose $r \geq 3$ and $R_\alpha = \lfloor \alpha \log_{r-1} n \rfloor$. (i) If $\alpha < 1/2$ For any $x$ the probability that the collection of vertices within distance $R_\alpha$ of $x$ differs from an $r$-regular tree is $\leq 10n^{-(1-2\alpha)}$ for large $n$. (ii) If $\alpha < 1/4$ then the probability there is some vertex in the graph where the collection of vertices within distance $R_\alpha$ of $x$ differs from an $r$-regular tree by more than one edge is $\leq 100n^{-(1-4\alpha)}$ for large $n$.

**Proof.** To construct an $r$-regular graph we attach $r$ half-edges to each vertex. To grow the graph starting from $x$ we pick $r$ half edges at random from the $rn$ possible options. Let $x_1, \ldots, x_{r}$ be the neighbors of $x$. For each of these we will pick $(r - 1)$ half-edges to determine their neighbors. To generate the graph out to distance $R$ we will make

$$M = r \left[ 1 + (r - 1) + \cdots + (r - 1)^{R-1} \right]$$

$$\leq r(r - 1)^{R-1} \sum_{k=0}^{\infty} (1-r)^{-k}$$

$$= r(r - 1)^{R-1} \cdot \frac{1}{1 - 1/(r - 1)} \leq \frac{rn^{\alpha}}{r - 2}$$

choices. The probability that at some we pick a vertex that has already been chosen is

$$\leq M \cdot \frac{M}{n - M} \leq \left( \frac{r}{r - 2} \right)^{2} \frac{n^{2\alpha}}{n - rn^{\alpha}}$$

which proves (i). The probability that on at least two different instances pick a vertex that has already been chosen is

$$\leq \left( \frac{M}{2} \right) \cdot \left( \frac{M}{n - M} \right)^{2} \leq \left( \frac{r}{r - 2} \right)^{4} \frac{n^{4\alpha}}{(n - rn^{\alpha})^{2}}$$
which proves (ii) \hfill \Box

To assess the impact of the misconnected edge, we recall the definition of the exploration process. There are three sets:

- \( R_t \), the number of removed sites with known connections,
- \( U_t \), the number of unexplored sets which are not yet part of the growing cluster,
- \( A_t \) the number of active sites that are part of the growing cluster but are not in \( R_t \)

We start with \( R_0 = \emptyset \), \( U_0 = \{2, 3, \ldots, n\} \), and \( A_0 = \{1\} \). At time \( \tau = \inf\{t : A_t = \emptyset\} \) the process stops. To explore the vertices in the cluster we use a breadth-first search. If \( A_t \neq \emptyset \), pick \( i_t \) from the vertices in \( A_t \) with the smallest distance \( d_t \) from 1 and let

\[
\begin{align*}
R_{t+1} &= R_t \cup \{i_t\} \\
A_{t+1} &= A_t - \{i_t\} \cup \{y \in U_t : \eta_{i_t,y} = 1\} \\
U_{t+1} &= U_t - \{y \in U_t : \eta_{i_t,y} = 1\}
\end{align*}
\]

By the definition of \( d_t \) all of the sites in \( A_t \) have distance \( \geq d_t \). Since all of the sites in \( A_t \) are neighbors of a point in \( R_t \) they all have distance \( \leq d_t + 1 \). We say a collision occurs if \( i_t \) has a neighbor in \( A_t \). In this case the worst thing that can happen is that \( i_t \) connects to a vertex at distance \( d_t \) rather than \( d_t + 1 \), so a jump from this vertex gets closer to \( x \) with probability \( 1/r \), stays the same distance with probability \( 1/r \) and goes further away with probability \( (r-2)/r \). In view of this small change in the walk we will do our computations as if the graph up to distance \( R \) is exactly a tree and we will leave it to the reader to deal with the annoying details of dealing with the one bad vertex.

### 7.5.1 Convergence to equilibrium

Following our approach to coalescing random walks on the torus, the first thing to do is to estimate the time to converge to equilibrium.

**Lemma 7.5.3.** There is a constant \( C_{\text{reg}} \) so that if \( t \geq C_{\text{reg}} \log N \) then

\[
\max_x \sum_y |H_t(x,y) - \pi(y)| \leq \frac{1}{n^2}
\]

**Proof.** Theorem 6.3.1 implies that Cheeger’s constant \( h \geq \alpha_0 \). Theorem 6.2.1 implies that the spectral gap \( 1 - \lambda_1 \geq h^2/2 \). Theorem 6.1.2 implies

\[
\left| \frac{H_t(x,y)}{\pi(y)} - 1 \right| \leq \frac{e^{-(1-\lambda)t}}{\pi_{\min}} \leq ne^{-\gamma t}
\]

where \( \gamma = \alpha_0^2/2 \). Multiplying on each side by \( \pi(y) \) and summing, then taking \( t = C \log n \) where \( C \) is sufficiently large gives the desired result. \hfill \Box

A much more precise result has been proved by Lubetsky and Sly (2010)
Chapter 7. Voter Models, Coalescing RWS

Theorem 7.5.4. The simple random walk on the random \( r \)-regular graph \( G_r \) exhibits cutoff at time \( t_{r,n} = r(r-2)\log_{r-1} n \) with a window of order \( \sqrt{\log n} \)

Cutoff at time \( t_{n,r} \) means that the total variation distance between the the distribution starting from a fixed vertex \( x \) and the stationary distribution, which is uniform,

\[ \to 1 \quad \text{at time} \quad (1 - \epsilon)t_{r,n} \]
\[ \to 0 \quad \text{at time} \quad (1 + \epsilon)t_{r,n} \]

The window size is the width of the region over which the total variation distance goes from \( 1 - \epsilon \) to \( \epsilon \). For more on this notion see Chapter 18 of Levin and Peres (2017).

To explain the result suppose that when viewed from \( x \), \( G_r \) is exactly a tree. There are \( r \) edges connected to \( x \) but after that there are \( r-1 \) edges that lead away from \( x \). Thus the height of the tree is \( \approx \log_{r-1} n \). When the random walk is away from the root of tree the probability of moving one step further away from \( x \) is \( (r-1)/r \) while one step closer has probability \( 1/(r-1) \) so the drift is \( (r-2)/r \). Thus the time to reach the top of the tree is \( \sim t_{r,n} \).

Lemma 7.5.2 shows that up to height \((1/3)t_{r,n}\) the graph seen from \( x \) is with high probability exactly a tree. Berestycki and Durrett (2008) showed that while the distance is \( \leq (1 - \epsilon)t_{r,n} \) the graph looks enough like the tree so that the computation of the drift is correct. To oversimplify more than a little, Lubetsky and Sly (2010) have strengthened the connection between the random graph and the tree, and showed that when the distance from \( x \) exceeds \( t_{r,n} \) the distribution is close to uniform. The distance of the random walk on the tree from \( x \) satisfies the central limit theorem, which gives the width of the window in cutoff.

7.5.2 Hitting times

To prove Theorem 7.5.1 we will follow the general approach we took on the torus. However, now the graph is not translation invariant so we cannot reduce the coalescence of two particles to the hitting time of a point. We will ignore our particles when they are at a distance larger than \( R_{1/3} \). For this reason it will be useful to have an estimate for the motion of a random walk on an \( r \)-regular tree. We assume there is a root at \( 0 \) with degree \( r \), and let \( X_n \) be the height of the particle at time \( n \) in the discrete time walk. Let \( T_a = \min\{n \geq 0 : X_n = a\} \)

Lemma 7.5.5. Suppose \( a < m < b \)

\[
P_m(T_a < T_b) = \frac{(r-1)^{-m} - (r-1)^{-b}}{(r-1)^{-a} - (r-1)^{-b}}
\]
\[
P_m(T_a > T_b) = \frac{(r-1)^{-a} - (r-1)^{-m}}{(r-1)^{-a} - (r-1)^{-b}}
\]

Proof. It is easy to check that \( (r-1)^{-X_n} \) is a martingale until \( X_n \) hits 0. Applying the optional stopping theorem at the exit time from \((a,b)\) gives

\[
(r-1)^{-m} = (r-1)^{-a}P_m(T_a < T_b) + (r-1)^{-b}[1 - P_m(T_a < T_b)].
\]
Rearranging gives the first equality, and the second one follows immediately. □

Since the degree is constant the stationary distribution is uniform \( \pi(x) = 1/n \). To be able to use Kac’s theorem Theorem 6.3.3 in PTE5 let \( \bar{X}_m \) be the discrete time chain on \( G_r \times G_r \) in which on each step one of the two walks is chosen to jump. Let \( \pi_2 \) denote the initial condition in which the two particles are at sites independently chosen at random from the graph. Let \( A = \{(x, x) : x \in G_r \} \). Let \( \bar{T}_A = \inf\{m : \bar{X}_n \in A\} \)

**Lemma 7.5.6.** \( E_{\pi_2} \bar{T}_A \sim \theta_r n \) where \( \theta_r = (r - 1)/(r - 2) \).

*Proof.* Writing \( P_A \) for \( P_{\pi_2}(\cdot | \bar{X}_0 \in A) \), a theorem of Kac implies that

\[
E_A(\bar{T}_A) = \frac{1}{\pi_2(A)}
\]

Let \( t_N = C_{reg} \log N \). Starting from the diagonal \( A \), the chain may return to it in a time that is \( O(t_N) \). The expected value on this event makes a contribution that is \( o(N) \) to the expected value. When the two particles don’t hit by time \( O(t_N) \), Lemma 7.5.3 implies that the chain is close to equilibrium, so

\[
1/\pi_2(A) \approx P_A(\bar{T}_A \geq t_N)E_{\pi_2}(\bar{T}_A)
\]

and we have

\[
E_{\pi_2}(\bar{T}_A) = \frac{1}{\pi_2(A)} \cdot \frac{1}{P_A(\bar{T}_A \geq t_N)} \quad (7.5.4)
\]

Let \( x \in G_r \) and \( R = [(1/3) \log_{r-1} n] \). Lemma 7.5.2 implies that with probability \( \geq 1 - 10n^{-1/3} \) a neighborhood of size \( R \) around \( x \) looks exactly like a tree in which all vertices have degree \( r \). Suppose both particles start at \( x \) at time 0, and let \( D_m = d(\bar{X}_m^1, \bar{X}_m^2) \). \( D_1 = 1 \). Using Lemma 7.5.5 we see that on the infinite \( r \)-regular tree

\[
P_1(T_0 < T_\infty) = 1 - (r - 1)^{-1} = \frac{r - 2}{r - 1}
\]

and hence if \( a_n \to \infty \), \( P_1(T_0 < T_{a_n}) \to (r - 2)/(r - 1) \). In the proof of Theorem 7.5.8 if two particles are separated by \( a_n \) then it is unlikely they will hit by \( t_n \), which completes the proof. □

**Theorem 7.5.7.** Let \( T_A \) be the hitting time for the continuous time chain. \( E_{\pi_2}T_A \sim \theta_r N/2 \), and when the initial condition is \( \pi_2 T_A/n \to \text{exponential}(2/\theta_r) \)

*Proof.* The formula for the mean follows from the fact that the continuous time chain moves twice as fast as \( \bar{X}_n \). The fact that the limit is exponential follows from the proof of Theorem 3 in the previous chapter. □

The next step is to consider separated particles

**Theorem 7.5.8.** Let \( a_N \to \infty \) and suppose \( d(x, y) \geq a_n \). Let \( T_{x,y} \) be the first time the random walks starting from \( x \) and \( y \) hit. Then \( T_{x,y}/n \) converges in distribution to \( \text{exponential}(2/\theta_r) \)
CHAPTER 7. VOTER MODELS, COALESCING RWS

Proof. Step 1. Suppose first that

\[ a_N \leq d(x, y) \leq R_1 = \lfloor 0.1 \log_{r-1} n \rfloor \]

If the particles are further apart than \( R_1 \) we will first let them wander until they come to distance \( R_1 \). Lemma 7.5.2 implies that the set of vertices within distance \( R_2 = \lfloor 0.2 \log_{r-1} n \rfloor \) of \( X \) differs from the \( r \)-regular tree by at most one edge. As announced earlier, we will prove the result under the assumption that suppose that what we see up to distance \( R_2 \) is exactly an \( r \)-regular tree and leave the annoying details of dealing with the rogue edge to the reader.

Let \( z_0 = x, z_1, \ldots, z_{d(x,y)} = y \) be the path from \( x \) to \( y \) that minimizes the distance. Suppose for simplicity that \( d(x, y) \) is even and let \( w = z_{d(x,y)/2} \) be the midpoint. Let \( B_z \) for bad be the event that the random walk starting at \( z \) hits \( w \) before it goes a distance \( R(2) = R_2 \) away from \( w \). Using Lemma 7.5.

\[
P(B_x) \leq P_{d(x,y)/2}(T_0 < T_{R(2)}) = \frac{(r-1)^{-d(x,y)/2} - (r-1)^{-R(2)}}{1 - (r-1)^{-R(2)}} \leq 2(r-1)^{-d(x,y)/2}
\]

when \( n \) is large. This implies that the probability that \( x \) and \( y \) hit before \( T = \) the first time one of the two moves more than \( R_2 \) away from \( w \) is \( \leq 4(r-1)^{-d(x,y)/2} \). Note that on \( B_x \cap B_y \) they are separated by \( R_2 \) at time \( T \).

Induction step. If the two particles are a distance \( > R_1 \) then run them until they come to distance \( R_1 \). Repeating the last argument

\[
P(B_x) \leq P_{R_1/2}(T_0 < T_{R_2}) = \frac{(r-1)^{-R_1/2} - (r-1)^{-R_2}}{1 - (r-1)^{-R(2)}} \leq 2N^{-0.05}
\]

The number of jumps required to go from distance \( R_1/2 \) to \( R_2 \) is \( \geq \lfloor 0.15 \log_{r-1} N \rfloor \), so with high probability the total time is \( \geq \lfloor 0.15 \log_{r-1} N \rfloor / 2 \). After a finite number of iterations the time will be \( \geq C_{reg} \log N \) and the particles will be in equilibrium. Using Theorem 7.5.7 now gives the desired result. \( \square \)

7.5.3 Coalescing random walks

The next step is to consider \( k \) particles.

Lemma 7.5.9. Run \( k \geq 3 \) independent random walks starting from points separated by distances at least \( a_N \). (a) If we let \( \tau_{i,j} \) be the first time \( i \) and \( j \) hit and let \( \tau^k = \min_{1 \leq i < j \leq k} \tau_{i,j} \) then

\[
P(\tau^k > \theta_r n t) \to \exp(-tk(k-1)) \tag{7.5.5}
\]

(b) The probability that at time \( \tau_{i,j} \), that there are random walks (other than \( i \)th and \( j \)th) separated by distance \( \leq a_n \) tends to 0.

Proof. The proof of Theorem 7.5.8 implies that the probability that two particles hit by time \( t_n = C_{reg} \log n \) tends to 0. Lemma 7.5.3 implies that at this time the particles are
in equilibrium. Using the last two facts we can conclude that the probability two pairs of particles \( \{i_1, j_1\} \) and \( \{i_2, j_2\} \neq \{i_1, j_1\} \) are both within distance \( a_n = (1/10) \log_{r-1} n \) at some time \( t \in [C_{reg} \log n, n \log n] \) is

\[
\leq k^4 \cdot n \log n \left( \frac{2n^{1/10}}{n} \right)^2 \to 0
\]

The proof of (a) now follows from the proof of Lemma 7 in the previous chapter. \( \square \)

Let \( q_{j,k}(t) \) be the transition probability of Kingman’s coalescent. Since coalescence occurs at rate \( O(k^2) \), \( q_{\infty,k}(t) = \lim_{n \to \infty} q_{n,k}(t) \) exists.

**Theorem 7.5.10.** Assume \( d \geq 3 \), \( T > 0 \), the initial number of particles \( n \geq 2 \) and initial conditions \( \zeta_0^{N,n} = \{x_1, \ldots, x_n\} \) with \( |x_i - x_j| \geq a_N \) for \( i \neq j \)

\[
P(|\zeta_{tN}^{N,n}| = k) \to q_{n,k}(2t/\theta_r)
\]  

(7.5.6)

If we start with \( \zeta_0^{n,\infty} = G_r \) then the conclusion holds for \( n = \infty \).

**Proof.** The result for fixed \( n \) follows from Lemma 7.5.9. The extension to \( k = \infty \) can be done using Lemmas 7.4.6 and ??.
7.6 Oliveira’s result

On the torus and random regular graphs we proved convergence of the rescaled coalescence time to Kingman’s coalescent using detailed calculation. The result in this section shows that it is a general result. To state that result

Let $X_t$ be a continuous time Markov chain with generator $Q$

Let $\pi$ be the stationary distribution of $Q$

let $m(Q)$ be the expected meeting time of two independent copies of the chain $Q$, each starting from the stationary distribution,

let $t^{\text{mix}}(\alpha)$ be the $\alpha$-mixing time, i.e., the smallest value of $t$ so for all initial states $x$

$$\|P_x(X_t \in \cdot) - \pi(\cdot)\| \leq \alpha$$

The specific value $t^{\text{mix}}(1/4)$ is called the mixing time and denoted by $t^{\text{mix}}$. The constant $1/4$ is chosen so that for all $\epsilon < 1/4$,

$$t^{\text{mix}}(\epsilon) \leq C \ln(1/\epsilon)t^{\text{mix}}$$

To define the limiting value of Kingman’s coalescent let $Z_i$, $i \geq 2$ be independent with

$$P(Z_i > t) = \exp(-t \binom{i}{2}).$$

Finally to define the sense in which convergence occurs define the Wasserstein distance between the distributions of two random variables with finite means by

$$d_W(X, Y) = \int |P(X > x) - P(Y > x)| dx$$

As is well known

$$d_W(X, Y) = \sup\{|Ef(X) - Ef(Y)| : f \text{is 1-Lipshitz}\}$$

The first results is for the transitive case: i.e., given $x$ and $y$ one can find a permutation of the state space that makes $x$ to $y$, and in addition assumes reversibility i.e., the stationary distribution satisfies detailed balance $\pi(x)Q(x, y) = \pi(y)Q(y, x)$.

**Theorem 7.6.1. (Mean-field limit for transitive reversible Markov chains).** Start with one copy of the Markov chain at each site and let $C$ be the time needed for all the particles to coalesce to 1. If $\rho(Q) = t^{\text{mix}}/m(Q)$ then

$$d_W \left( \frac{C}{m(Q)}, \sum_{i \geq 2} Z_i \right) = O \left( \left[ \rho(Q) \ln(1/\rho(Q)) \right]^{1/6} \right)$$
The second result is for a general chain but assume that the mixing time is small enough relative to the other parameters of the chain.

**Theorem 7.6.2.** Let $V$ be the state space. Let $q_{\text{max}}$ be the maximum transition rate from any $x \in V$ and let $\pi_{\text{max}} = \max\{\pi(v) : v \in V\}$. Let $C$ be the coalescent time when one starts a random walk starting from each vertex of $v$. If $\alpha(Q) = (1 + q_{\text{max}}t_{\text{mix}})^{\pi_{\text{max}}}$ then

$$d_W \left( \frac{C}{m(Q)}, \sum_{i=2}^{\infty} Z_i \right) = O \left( \frac{[\alpha(Q) \log(1/\alpha(Q)) \cdot \log^4 |V|]^{1/6}}{} \right)$$
CHAPTER 7. VOTER MODELS, COALESCING RWS

7.7 Asymptotics for CRW densities

7.7.1 On the torus in \( d \geq 2 \)

By duality between the voter model \( \xi_t^0 \) and coalescing random walk \( \zeta_t^1 \)

\[
p_t = P(\xi_t^0 \neq \emptyset) = P(0 \in \zeta_t^1).
\]

In 1980 Bramson and Griffeath (1980), completed the proof of

\[
p_t \sim \begin{cases} 
(\log t)/\pi t & d = 2 \\
1/\beta_d t & d \geq 3
\end{cases}
\tag{7.7.1} \text{CRWa2}
\]

where \( \beta_d \) is the probability a \( d \)-dimensional simple random walk starting at 0 never returns there.

There is a simple heuristic that gives the answer. If lattice sites on \( \mathbb{Z}^d \) are independent and occupied with probability \( u \), then the instantaneous rate at which collisions occur is \( u^2 \). If we assume that the particles are always randomly scattered on the lattice then their density would satisfy the differential equation:

\[
\frac{du}{dt} = -u^2 \quad u(0) = 1
\]

The solution is \( u(t) = 1/(1+t) \). If our random walks do not coalesce when they meet then in \( d \geq 3 \), two that hit once would hit a geometric number of times with mean \( 1/\beta_d \). Removing the over counting from the previous equation

\[
\frac{du}{dt} = -\beta_d u^2
\]

so \( u(t) \sim 1/(\beta_d t) \). To prepare for the next calculation, note that we have multiplied the naive asymptotics, \( 1/u \), by the expected number of collisions given that one occurred.

In \( d = 2 \), this reasoning needs to be modified since random walks are recurrent and hence they hit infinitely many times. In \( d = 2 \) the random walk \( S_t \) that jumps to a randomly chosen nearest neighbor at rate 2 has mean 0 and covariance matrix \( I \) where the \( I \) is the identity matrix, so using the (local) central limit theorem

\[
P(S_t = 0) \sim \frac{1}{2\pi t}
\]

When \( S_t \) hits 0, it stays there for an amount of time that is exponential with mean \( 1/2 \). The expected number of times two random walks hit in \([0, t]\) is thus

\[
\sim 2 \int_1^t \frac{1}{2\pi s} ds \sim \frac{\log t}{\pi}.
\]

Multiplying the naive asymptotic \( \sim 1/t \) by this gives the result in \( d = 2 \).
The rigorous derivation of (7.7.1) involves a number of clever observations. We follow the summary in Bramson and Griffeath (1980). Unfortunately in that paper the voter model is $\zeta_t$ and the coalescing random walk is $\xi_t$. We will write things in our notation. Consider the voter model starting from all sites different and let

$$N_t = |\{x : \xi_t(x) = \xi_t(0)\}|$$

be the size of the patch of particles in the voter model at time $t$ that is the same type as the particle at the origin. Sudbury (1976) observed that

Lemma 7.7.1. If $R_t$ is the number of points visited by our rate 2 random walk by time $t$, i.e., $R_t = |\{x : x = S_s \text{ for some } s \leq t\}|$ then

$$EN_t = ER_t \tag{7.7.2}$$

Proof. In order for $\xi_t(0) = \xi_t(x)$, we must have $\zeta_s^{0,t} = \zeta_s^{x,t}$ for some $s \leq t$, which has the same probability as $\zeta_s^{0,t} - \zeta_s^{x,t}$ hitting 0 by time $t$, which is the same as the probability $S_r$ hits $x$ by time $t$ when it starts from 0. \(\square\)

Combining (7.7.2) with results of Dvoretsky and Erdős for the range of random walk we have

$$EN_t \sim \begin{cases} 2^{\beta_d t} & d = 2 \\ 2\beta_d t & d \geq 3 \end{cases} \tag{7.7.3}$$

The reasoning is the same as in our heuristic proof. The rate 2 random walk makes $\sim 2t$ jumps by time $t$. To find the range we divide this by the average number of times a site is visited given that it is visited at least once.

Kelly (1977) noticed

Lemma 7.7.2. Let $n_t = |\xi_t^0|$.

$$P(N_t = k) = kP(n_t = k) \tag{7.7.4}$$

Proof. Suppose someone shows us a large $d$-dimensional cube of side $L$ from the configuration from time $t$, not telling us where the origin is located, and then we pick a site at random to be the origin. An opinion with $k$ representatives will be chosen with probability $k/L^d$. Thus when we pick a cluster containing a given site, we get a size biased pick from the original distribution. Note that since $En_t = 1$, the right-hand side is a probability distribution. \(\square\)

Combining (7.7.4) with Jensen’s inequality and (7.7.3) we have

$$p_t = E(N_t^{-1}) \geq (EN_t)^{-1}$$

which gives an asymptotic lower bound of $1/2$ the correct answer in (7.7.1). To close the gap we use the following result of Sawyer (1979).
**Lemma 7.7.3.** For \( d \geq 2 \), and \( k = 1, 2, \ldots \)

\[
\lim_{t \to \infty} E\left(\frac{N_t}{EN_t}\right)^k = \frac{(k + 1)!}{2^k}
\]

and hence

\[
\lim_{t \to \infty} P\left(\frac{N_t}{EN_t} \leq x\right) = \int_0^x 4ye^{-2y} \, dy
\]

In words \( N_t/EN_t \) converges to a limit \( Y \) with gamma(2,2) distribution, the sum of 2 independent exponentials with rate 2. Less formally

\[ N_t \sim Y \cdot EN_t \]

Ignoring the fact that \( x^{-1} \) is unbounded near 0, and hence small probabilities can have a large impact on the expected value, we have

\[ p_t = EN_t^{-1} \sim \frac{EY^{-1}}{EN_t} \sim \frac{2}{EN_t} \]

and we have the quoted result.

Bramson and Griffeath (1980) make this rigorous by showing \( p_t = O((\log t)/t) \) in \( d = 2 \) and \( p_t = O(1/t) \). Our description makes it seem that all they did was cross the t and dot the i in asymptotics. To complete the proof they had to use an ingenious argument to show that the coalescing random walk had a self-correcting mechanism, i.e., if \( p_t \) is larger than it should be then coalescence will be more rapid. A major technical difficult is that knowing the density does not give one much of an idea about how they are distributed in space.

**Open question.** It would be nice to have a less devious proof for the asymptotics in (7.7.1). van den Berg and Kesten (2000) have given a much more direct proof but unfortunately it only works in dimensions \( d \geq 6 \). They start with the equality

\[
\frac{d}{dt}p_t = -2P(0 \text{ and } e_1 \text{ are occupied at time } t)
\]

Intuitively by reversing time the probability that two random walks are sitting next to each other and have not coalesced is the probability two random walks moving forward never hit. Their paper is 49 pages, not because the new proof is difficult but because they prove a result about a system in which coalescence occurs at rate \( p_j \) when there are \( j \) particles on one site, assuming \( p_1 > 0 \) and \( j \to p_j \) is increasing in \( j \).
7.7.2 On graphs

Herman et al (2021) study CRW on a finite graph $G = (V, E)$ starting from one particle at each vertex and in which particles perform independent edge simple random walks i.e., for each edge incident to the current location jumps across that edge occur at rate 1. They also consider the more general situation in which the jump rate $r_{x,y} = r_{y,x}$ depends on the edge. This implies that the Markov chain is reversible with respect to counting measure.

They denote by $\xi_t$ the set of elements of $V$ that are occupied at time $t$ and define the coalescence time

$$\tau_{coal} = \inf\{t : |\xi_t| = 1\}$$

Interest focuses on the rate of decay of

$$p_t(x) = P(x \in \xi_t)$$

the probability $x$ is occupied at time $t$.

They are interested in this problem for two collections of graphs

(1) Graphs generated by the configuration model with minimum degree 3. In this case it is natural to guess that $p_t = O(1/t)$ or even $\lim_{t \to \infty} tp_t = c$.

(2) A family of vertex transitive graphs $G = (V, E)$ satisfying some general ‘transience-like’ and ‘spectral’ conditions.

To formally state the results we need some definitions.

Given two independent walks $X_t$ and $Y_t$ we define the meeting time by

$$t_{meet} = \inf\{t : X_t = Y_t\}$$

For each $x \in V$ we define the neighborhood distribution $\nu_x$

In case (1) to be the uniform distribution on the neighbors of $x$.

In case (2) to be $\nu_x(y) = r_{x,y}/r(x)$ where $r(x) = \sum_y r_{x,y}$.

A moments thought shows that the definition in case (1) is a special case of that in case (2) so in general we can define

$$\alpha_t(x) = r(x)P_{x,\nu_x}(t_{meet} > t)$$

where the subscript gives the distributions of $X$ and $Y$ in the definition of the meeting time. In words we start $X$ at $x$ and $Y$ at a neighbor chosen at random. We define

$$\alpha_t = \frac{1}{|V|} \sum_{x \in V} \alpha_t(x)$$

In terms of the newly introduced notation the two main results can be stated as

(A1) $P_t \approx 1/t \alpha_t$

(A2) $P_t \approx 2t_{meet}/nt$
Configuration model

Assume that the degree distribution \( D \) is supported on \([3, M]\) for some \( M < \infty \). Let \( G_n \) be the graph on \( n \) vertices generated by the configuration model \( \mathbb{CM}_n(D) \). The lower bound \( D \geq 3 \) implies that the probability \( G \) is connected tends to 1 (Theorem 4.15 in [vdH20]).

In addition to \( \mathbb{CM}_n(D) \) they also consider that local weak limit of \( \mathbb{CM}_n(D) \) which is a unimodular Galton-Watson tree \( \mathcal{G} \sim \mathbb{UGT}(D) \), where the root has offspring distribution \( D \) and later generations have offspring distribution \( D^* \) which is the size biased version of \( D \).

Let \( o \) be the root of \( \mathcal{G} \), let \( \alpha(D) = E(\alpha_\infty(o)) \) where \( \alpha_\infty(x) = \lim_{t \to \infty} \alpha_t(x) \) with \( \alpha_t(x) \) defined in (7.7.5). Theorem 1 in Hermon et al is

**Theorem 7.7.4.** Let \( \mathcal{G}_n \) be sampled from \( \mathbb{CM}_n(D) \). For any sequence of times so that \( 1 \ll t_n \ll n \) we have

\[
\lim_{n \to \infty} t_n P_{t_n}(\mathcal{G}_n) = 1/\alpha \\
\lim_{n \to \infty} \frac{nt_n}{2t_{meet}(\mathcal{G}_n)} P_{t_n}(\mathcal{G}_n) = 1
\]
References


Hermon, J., Li, S, Yao, D., and Zhang, L. (2021) Mean-field behavior during the Big Bang for coalescing random walk. arxiv:2105.11585


