# Chapter 5

# Small Worlds

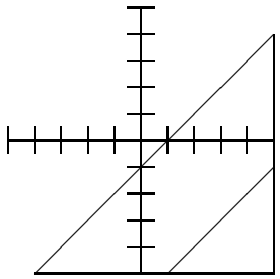
## 5.1 Watts and Strogatz model

As explained in more detail in Section 1.3, our next model was inspired by the popular concept of "six degrees of separation," which is based on the notion that every one in the world is connected to everyone else through a chain of at most six mutual acquaintances. Now an Erdös-Renyi random graph for $n = 6$ billion people in which each individual has an average of $\mu = 42.62$ friends would have average pairwise distance $(\log n)/(\log \mu) = 6$, but would have very few triangles, while in social networks if $A$ and $B$ are friends and $A$ and $C$ are friends, then it is fairly likely that $B$ and $C$ are also friends.

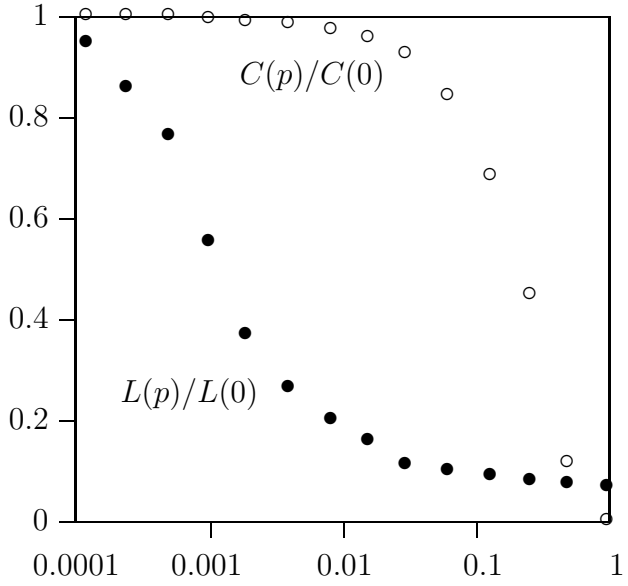
To construct a network with small diameter and a positive density of triangles, Watts and Strogatz (1998) started from a ring lattice with $n$ vertices and $k$ edges per vertex, and then rewired each edge with probability $p$, connecting one end to a vertex chosen at random. This construction interpolates between regularity ($p = 0$) and disorder ($p = 1$). The disordered graph is not quite an Erdös-Rényi graph, since the degree of a node is the sum of a Binomial($k$,1/2) and an independent Poisson($k/2$). Let $L(p)$ be the distance between two randomly chosen vertices. Define the clustering coefficient $C(p)$ to be the fraction of connections that exist between the $\binom{k}{2}$ neighbors of a site.

Suppose that $n \gg k \gg \log n \gg 1$. Extrapolating from the results for Erdös-Renyi graphs, we know that the middle condition implies that the graph will be connected with high probability when $p = 1$ and the diameter will be asymptotically $(\log n)/(\log k)$. Considering the first two steps in the cluster growth branching process tells us that the clustering coefficient $C(1) \sim k/n$. At the other extreme of perfect order, since we can move distance $k/2$ in one step and the maximum distance is $n/2$, $L(0) \sim n/k$.

Our next step is to show $C(0) \to 3/4$. Suppose $k = 2j$. The pairs of points $-j \le y < x \le j$ form a triangle with vertices $(j, j-1)$, $(-(j-1), -j)$, and $(j, -j)$. The points below the line $x - y > j$ are not neighbors, and this is asymptotically 1/4 of the triangle. The next figure shows the situation when $j = 5$.



The next illustration, which is a copy of Figure 2 in Watts and Strogatz (1998), considers $n = 1000$ vertices and $k = 10$ neighbors, and shows that there is a broad interval of $p$ over which $L(p)$ is almost as small as $L(1)$, yet $C(p)$ is far from 0. To see the reason for this, note that when a fraction $p = 0.01$ of the edges have been rewired, $C(p)$ has not changed by much, but the short cuts have dramatically decreased the distance between sites.




To look for the small world phenomenon in real graphs, Watts and Strogatz (1998) computed $L$ and $C$ for three examples: the collaboration graph of actors in feature films, the electrical power grid of the Western United States, and the neural network of the nematode worm *C. elegans*. Results are given in the next table and are compared to the values $L_r$ and $C_r$ for random graphs with the same number of vertices and average number of edges per vertex. As these results show the distances are similar to the random graphs in the first two cases, but 50% larger in the third. However, the clustering coefficients in the real graphs

are significantly larger than in the random graphs, especially when the number of vertices is large in the case of film actors.

| | $L$ | $L_r$ | $C$ | $C_r$ |
|---|---|---|---|---|
| *C. elegans* | 2.65 | 2.25 | 0.28 | 0.05 |
| Film actors | 3.65 | 2.99 | 0.79 | 0.00027 |
| Power grid | 18.7 | 12.4 | 0.08 | 0.005 |

**Bollobás and Chung small world.** Watts and Strogatz (1998) were not the first to notice that random long distance connections could drastically reduce the diameter. Bollobás and Chung (1988) added a random matching to a ring of $n$ vertices with nearest neighbor connections and showed that the resulting graph had diameter $\sim \log_2 n$. This graph, which we will call the BC small world, is not a good model of a social network because (a) every individual has exactly three friends including one long range neighbor, and (b) does not have any triangles, so it is locally tree like. These weaknesses, particularly the second, make it easier to study, so we will have a preference for this case throughout most of the chapter. In the section on epidemics and the final section on the contact process, we will include nonnearest neighbor connections. There, as in the other models considered in this chapter, the qualitative behavior is the same but the proofs are more difficult.

## 5.2 Path lengths

In this section we are concerned with estimating the average path length between two randomly chosen sites in the small world, $\ell(n,p)$ as a function of the number of nodes $n$, the fraction of shortcuts $p$, and the range of interaction $k$. For this problem and the others we will consider below, we will consider Newman and Watts (1999) version of the model in which no edges are removed but one adds a Poisson number of shortcuts with mean $pkn/2$ and attaches then to randomly chosen sites.

To quote Albert and Barabási (2002), "it is now widely accepted that the characteristic path length obeys the general scaling form

$$\ell(n,p) \sim \frac{n}{k} f(pkn)$$

where $f(u)$ is a universal scaling function that obeys

$$f(u) = \begin{cases} 1/4 & \text{if } u \ll 1 \\ \ln(u)/u & \text{if } u \gg 1\text{"} \end{cases}$$

Newman, Moore, and Watts (2000) have taken a "mean-field approach" to computing $\ell(n,p)$. They write differential equations for the number of sites within distance $r$ of a fixed point and the number of clusters of occupied sites, assuming that gaps between clusters have the sizes of a randomly broken stick, i.e., the result of putting that many i.i.d. uniforms in the unit interval. They conclude that

$$f(u) = \frac{1}{2\sqrt{u^2+2u}} \tanh^{-1}\left(\frac{u}{\sqrt{u^2+4u}}\right) \tag{5.2.1}$$

Simulations show that this formula agrees with simulations for small $u$ or large $u$, but "as expected, there is some disagreement when $u \approx 1$. See Figure 3 in Newman (2000). Using the identity

$$\tanh^{-1} y = \frac{1}{2} \log\left(\frac{1+y}{1-y}\right)$$

we have

$$\tanh^{-1}\left(\frac{u}{\sqrt{u^2+4u}}\right) = \frac{1}{2}\log\left(\frac{1+u/\sqrt{u^2+2u}}{1-u/\sqrt{u^2+2u}}\right)$$

Inside the logarithm the numerator $\to 2$ as $u \to \infty$. The denominator

$$= 1 - \frac{1}{\sqrt{1+2/u}} \approx 1 - \frac{1}{1+1/u} \approx 1/u$$

combining our calculations

$$f(u) \sim \frac{\log(2u)}{4u}$$

which matches (21) in Newman, Moore, and Watts (2000).

Barbour and Reinert (2001) have done a rigorous analysis of the average distance between points in a continuum model in which there is a circle of circumference $L$ and a Poisson mean $L\rho/2$ number of random chords. The chords are the short cuts and have length 0. The first step in their analysis is to consider an upper bound model that ignores intersections of growing arcs and that assumes each arc sees independent Poisson processes of shortcut endpoints. Let $S(t)$ be size, i.e., the Lebesgue measure, of the set of points within distance $t$ of a chosen point and let $M(t)$ be the number of intervals. Under our assumptions

$$S'(t) = 2M(t)$$

while $M(t)$ is a branching process in which there are no deaths and births occur at rate $2\rho$.

$M(t)$ is a Yule process with births at rate $2\rho$ so $EM(t) = e^{2\rho t}$ and $M(t)$ has a geometric distribution

$$P(M(t) = k) = (1 - e^{-2\rho t})^{k-1} e^{-2\rho t} \tag{5.2.2}$$

Being a branching process $e^{-2\rho t} M(t) \to W$ almost surely. It follows from (5.2.2) that $W$ has an exponential distribution with mean 1. Integrating gives

$$ES(t) = \int_0^t 2e^{2\rho s}\, ds = \frac{1}{\rho}(e^{2\rho t} - 1)$$

At time $t = (2\rho)^{-1}(1/2)\log(L\rho)$, $ES(t) = (L/\rho)^{1/2} - 1$. Ignoring the $-1$ we see that if we have two independent clusters run for this time then the expected number of connections between them is

$$\sqrt{\frac{L}{\rho}} \cdot \rho \cdot \frac{\sqrt{L/\rho}}{L} = 1$$

since the middle factor gives the expected number of shortcuts per unit distance and the last one is the probability a short cut will hit the second cluster. The precise result is:

**Theorem 5.2.1.** *Suppose $L\rho \to \infty$. Let $O$ be a fixed point of the circle, choose $P$ at random, and let $D$ be the distance from $O$ to $P$. Then*

$$P\left[D > \frac{1}{\rho}\left(\frac{1}{2}\log(L\rho) + x\right)\right] \to \int_0^\infty \frac{e^{-y}}{1 + 2e^{2x}y}\, dy$$

Thus as in Theorem 3.4.1 the fluctuations are $O(1)$. To make a connection between the two results we note that the proof will show that the right-hand side is $E\exp(-2e^{2x}WW')$ where $W$ and $W'$ are independent exponentials.

*Proof.* To prove this we begin with a Poisson approximation result of Arratia, Goldstein and Gordon (1990). Suppose $X_\alpha$, $\alpha \in I$ are Bernoulli random variables with $P(X_\alpha = 1) = p_\alpha$. Let $V = \sum_\alpha X_\alpha$, $\lambda = EV$, and $Z$ be Poisson with mean $\lambda$. We are interested in conditions that imply $V$ and $Z$ are close in distribution. For each $\alpha \in I$ let $B_\alpha \subset I$ be a set that

contains $\alpha$. Intuitively, $B_\alpha$ is the neighborhood of dependence of $X_\alpha$. Variables outside the neighborhood will be almost independent of $X_\alpha$. Define

$$\begin{aligned} b_1 &= \sum_{\alpha\in I}\sum_{\beta\in B_\alpha} p_\alpha p_\beta \\ b_2 &= \sum_{\alpha\in I}\sum_{\beta\in B_\alpha,\beta\neq\alpha} E(X_\alpha X_\beta) \\ b_3 &= \sum_{\alpha\in I} E|E(X_\alpha - p_\alpha|X_\gamma,\gamma\notin B_\alpha)| \end{aligned}$$

**Theorem 5.2.2.** *Let $\mathcal{L}(V)$ be distribution of $V$. The total variation distance*

$$\begin{aligned} \|\mathcal{L}(V)-\mathcal{L}(Z)\| &\le 2\left((b_1+b_2)\left(\frac{1-e^{-\lambda}}{\lambda}\right)+b_3(1\wedge 1.4\lambda^{-1/2})\right) \\ &\le 2(b_1+b_2+b_3) \end{aligned}$$

To apply this suppose that we have intervals $I_1,\ldots,I_m$ with lengths $s_1,\ldots,s_m$ and intervals $J_1,\ldots,J_n$ with lengths $u_1,\ldots,u_n$ that are scattered independently and uniformly on a circle of circumference $L$. Let $X_{ij}$ be the indicator of the event $I_i\cap J_j\neq\emptyset$ and $V=\sum_{i=1}^m\sum_{j=1}^n X_{ij}$.

$$p_{i,j}\equiv P(I_i\cap J_j\neq\emptyset)=(s_i+u_j)/L$$

so if we let $s=s_1+\cdots+s_m$ and $u=u_1+\cdots+u_n$ then

$$\lambda\equiv EV=\sum_{i=1}^m\sum_{j=1}^n(s_i+u_j)/L=(ns+mu)/L$$

We define $B_{i,j}=\{(i,k):k\neq j\}\cup\{(\ell,j):\ell\neq i\}$ so that if $(k,\ell)\notin B_{i,j}$ and $(k,\ell)\neq(i,j)$ then $X_{i,j}$ and $X_{k,\ell}$ are independent and hence $b_3=0$. If we let

$$Z_{i,j}=\sum_{(k,\ell)\in B_{i,j}} X_{k,\ell}$$

then we have

$$\begin{aligned} b_1 &= \sum_{i,j} p_{i,j}EZ_{i,j}+\sum_{i,j}p_{i,j}^2 \\ b_2 &= \sum_{i,j} E(X_{i,j}Z_{i,j})=\sum_{i,j}p_{i,j}EZ_{i,j} \end{aligned}$$

since the $X_{i,j}$ are pairwise independent. To see this note that if $i\neq k$ and $j\neq\ell$ then $X_{i,j}$ and $X_{k,\ell}$ are clearly independent. To complete the proof now, it suffices to consider the case $i=k$ and $j\neq\ell$. However, in this situation even if we condition on the location of $I_i$ the two events $I_i\cap J_j\neq\emptyset$ and $I_i\cap J_\ell\neq\emptyset$ are independent.

Let $r = \max_i s_i + \max_j u_j = L \max_{ij} p_{i,j}$. $EX_{i,j} \le r/L$ so $EZ_{i,j} \le (m+n-2)r/L$ and

$$\sum_{i,j} p_{i,j} EZ_{i,j} \le \lambda(m+n-2)r/L$$

The final term $\sum_{i,j} p_{i,j}^2 \le (r/L)\sum_{i,j} p_{i,j} = r\lambda/L$ so $b_1 + b_2 \le 2\lambda(m+n)r/L$. Using the first inequality in Theorem 5.2.2 with $1 - e^{-\lambda} \le 1$, it follows that if $Z$ is Poisson($\lambda$)

$$\|\mathcal{L}(V) - \mathcal{L}(Z)\| \le 4(m+n)r/L \tag{5.2.3}$$

Let $\tau_x = (2\rho)^{-1}\{(1/2)\log(L\rho) + x\}$. Consider two independent copies of the upper bound model starting from $O$ and $P$ and run until time $\tau_x$. Let $M_x$ and $N_x$ be the number of intervals in the two processes, let $s_x$ and $u_x$ be the Lebesgue measure of the sets of points, and let $\hat{V}_x$ be the number of intersections. From (5.2.3) and $r \le 4\tau_x$ it is immediate that

$$|P(\hat{V}_x = 0 | M_x, N_x, s_x, u_x) - \exp(-L^{-1}(N_x s_x + M_x u_x))| \le 16(M_x + N_x)\tau_x/L \tag{5.2.4}$$

Taking expected value, then putting the expected value inside the absolute value

$$|P(\hat{V}_x = 0) - E\exp(-L^{-1}(N_x s_x + M_x u_x))| \le \frac{16\tau_x}{L} E(M_x + N_x) \tag{5.2.5}$$

Our next step is to estimate the number of collisions between the growing intervals in the upper bound process starting from $O$. Number the intervals $I_j$ in the order in which they were created. Let $Y_{i,j} = 1\{I_i \cap I_j \neq \emptyset\}$ and $G_i = \{Y_{i,j} = 0 \text{ for all } j < i\}$. Each interval $I_i$ with $i > 1$ has a parent, $P(i)$, which was the source of the chord that started it. Let $H_1 = 0$ and

$$H_i = \begin{cases} 0 & \text{on } G_i \cap \{H_{P(i)} = 0\} \\ 1 & \text{otherwise} \end{cases}$$

$H_i = 1$ indicates an interval that is bad due to experiencing a collision or being a descendant of a bad interval.

**Lemma 5.2.3.** *If $P(Y_{i,j} = 1) \le p$ for all $i, j$ then $P(H_i = 1) \le 2(i-1)p$.*

*Proof.* We prove the result by induction on $i$. The conclusion is clear for $i = 1$. $H_i = 1$ can occur for two reasons. The first is $P(G_i^c) \le (i-1)p$. The second is that $k$ is an ancestor of $i$ and $H_k = 0$. Now since the intervals are numbered in order of their creation, their lengths are a decreasing function of their indices, and hence the probability $j$ is the parent of $i$ is a decreasing function on $1, \ldots i-1$. Iterating we see that if we follow the ancestry of $i$ back until we first reach an interval $j \le k$ then the probability we will end up at $k$ is $\le 1/k$. Using induction now

$$\sum_{k=1}^{i-1} P(H_k = 0, k \text{ is an ancestor of } i) \le \sum_{k=1}^{i-1} (k-1)p/k \le (i-1)p$$

which completes the proof of the lemma. □

Let $V_x$ be the number of intersections in the real process in which intervals stop growing when they run into each other.

**Lemma 5.2.4.** *With the notations above we have*

$$P(\hat{V}_x \neq V_x) \leq \frac{32\tau_x^2}{L^2} E(M_x(M_x-1)N_x)$$

*Proof.* Define $Y'_{i,j} = 1\{J_i \cap J_j \neq \emptyset\}$ and $H'_i$ for the process starting at $P$ as in Lemma 5.2.3 and recall that $X_{i,j} = 1\{I_i \cap J_j \neq \emptyset\}$. Since $V_x \leq \hat{V}_x$ are integer valued

$$P(\hat{V}_x \neq V_x) \leq E(\hat{V}_x - V_x) \leq E\left(\sum_{i=1}^{M_x}\sum_{j=1}^{N_x} X_{i,j}(1\{H_i=1\} + 1\{H'_j=1\})\right)$$

Conditioning on $M_x = m$, $N_x = n$ and on the lengths of the intervals, using the trivial observation that all intervals have length $\leq 2\tau_x$ and applying Lemma 5.2.3 we conclude

$$\begin{aligned} &E(X_{i,j}1\{H_i=1\}|M_x=m, N_x=n, s_1,\ldots s_m, u_1,\ldots u_n)\\ &\quad\leq (4\tau_x/L)P(H_i=1|M_x=m, N_x=n, s_1,\ldots s_m, u_1,\ldots u_n)\\ &\quad\leq (4\tau_x/L)\cdot 2(i-1)(4\tau_x/L)\end{aligned}$$

Noting $\sum_{i=1}^{k} 2(i-1) = k(k-1)$, combining this with a similar bound for $X_{i,j}1\{H'_j=1\}$ and using $E(N_x(N_x-1)M_x) = E(M_x(M_x-1)N_x)$ gives the desired result. □

Theorem 5.2.1 follows easily by combining (5.2.5) and Lemma 5.2.4. To do this we need to recall that if $G$ has a geometric distribution with success probability $q$ then $EG = 1/q$ and $E(G(G-1)) = (1-q)/q^2 \leq 1/q^2$. From this and the definition of $\tau_x = (2\rho)^{-1}\{(1/2)\log(L\rho)+x\}$ we have $EM_x \sim (L\rho)^{1/2}e^x$ and $EM_x(M_x-1) \leq (L\rho)e^{2x}$. Using this with the cited results, writing $\ell_x = (1/2)\log(L\rho)+x$ and recalling $\tau_x = \ell_x/(2\rho)$ we have

$$\begin{aligned} &|P(V_x=0) - E\exp(-L^{-1}(N_x s_x + M_x u_x))|\\ &\quad\leq \frac{16\tau_x}{L}2(L\rho)^{1/2}e^x + \frac{32\tau_x^2}{L^2}(L\rho)^{3/2}e^{3x}\\ &\quad\leq \frac{16\ell_x}{(L\rho)^{1/2}}e^x + \frac{8\ell_x}{(L\rho)^{1/2}}e^{3x}\end{aligned}$$

which $\to 0$ if $x \leq (1/7)\log(L\rho)$.

To complete the proof we have to evaluate the expected value. Noting that

$$\frac{S(t)}{M(t)} = \int_0^t \frac{M(r)}{M(t)}\,dr \to \int_0^\infty e^{-\rho s}\,ds = \rho^{-1}$$

and recalling $s_x$ and $u_x$ are the total lengths of the intervals, we have

$$L^{-1}(N_x s_x + M_x u_x) \sim 2(L\rho)^{-1}M_x N_x \to -2e^{2x}WW'$$

where $W$ and $W'$ are independent exponential mean 1. The bounded convergence theorem implies

$$E\exp(-L^{-1}(N_x s_x + M_x u_x)) \to E\exp(-2e^{2x}WW')$$

If we condition on the value of $W$ and use the formula for the Laplace transform of the exponential

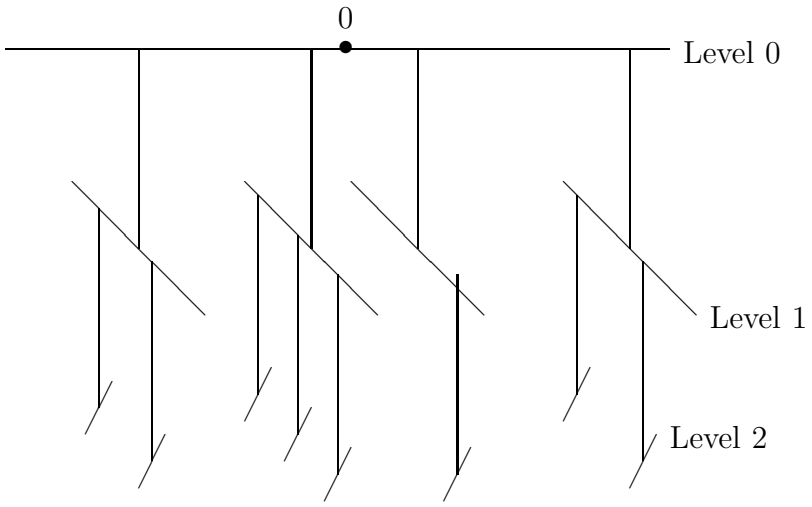
$$E\exp(-2e^{2x}WW') = E\left(\frac{1}{1+2e^{2x}W}\right) = \int_0^\infty \frac{1}{1+2e^{2x}y}e^{-y}\,dy$$

which completes the proof of the theorem. □

## 5.3 Epidemics

In this section we will follow Moore and Newman (2000) and consider epidemic models on the small world, which are essentially percolation processes. There are two extremes: in the first all individuals are susceptible and there is a probability $p$ that an infected individual will transmit the infection to a neighbor, in the second only a fraction $p$ of individuals are susceptible but the disease is so contagious that if an individual gets infected all of their susceptible neighbors will become infected. In percolation terms, the first model is bond percolation while the second is site percolation. The qualitative properties of the model are similar. We will concentrate on the site percolation version since in that case it is possible to do the computations more explicitly.

To give our first nonrigorous derivation of the answer, we will introduce an infinite graph associated with the small world, that we call the "Big World." We begin with a copy of the integers, $\mathbb{Z}$. To each integer we attach a Poisson mean $\rho$ long range bonds that lead to a new copy of $\mathbb{Z}$ on which we repeat the previous construction. The first copy of $\mathbb{Z}$ we call level zero. The levels of other copies are equal to the number of long range bonds we need to traverse to get to them. This structure appeared in an implicit way in the previous section: if we look at how the set of sites within distance $n$ of 0 in the Big World grows then there are no collisions and each interval encounters an independent set of long range connections.




**Site percolation.** To analyze the growth of the process, let $p_0(n)$ be the probability 0 is connected to $n$ sites on Level 0. $p_0(0) = 1 - p$. The number of sites to the right of zero that can reached has a geometric distribution with success probability $(1-p)^k$, since it takes $k$ consecutive closed sites to stop the percolation, and every time we can reach a new open site we can forget about the states of sites behind it. (This is false for bond percolation when $k > 1$ and makes the calculations in that case much more difficult.)

Letting $q = (1-p)^k$, the probability of reaching $j \geq 0$ sites on the right is $(1-q)^j q$. Adding the sites reached on the left and noting that 0 has to be open to get the process started we have

$$p_0(n) = \sum_{j=0}^{n-1} p(1-q)^j q \cdot (1-q)^{n-1-j} q = np(1-q)^{n-1}q^2$$

Noting that the geometrics start at 0, so their means are $1/q - 1$, the mean number of sites reached on level 0 is

$$\nu \equiv \sum_n n p_0(n) = p \cdot \frac{2-q}{q}$$

Conditional on $N = n$ sites being reached on level 0 the number of long range bonds $M$ to level one copies will be Poisson with mean $n\rho$. Thus $E(M|N) = \rho N$ and $EM = \rho\nu$. Each level one copy reached starts an independent version of the original process. Thus if we let $Z_k$ be the number of level $k$ copies reached then $Z_k$ is a branching process. The critical value for percolation occurs when $\rho\nu = 1$.

**Bond percolation, $k = 1$.** This time 0 does not have to be open so

$$p_0(n) = \sum_{j=0}^{n-1} p^j(1-p) \cdot p^{n-1-j}(1-p) = np^{n-1}(1-p)^2$$

and the mean number of sites reached on level 0 is

$$\nu = 1 + 2p/(1-p) = (1+p)/(1-p)$$

This time the edges have to be open in order to reach the next level so $E(M|N) = p\rho N$ and $EM = p\rho\nu$. The critical value for percolation occurs when $p_c\rho\nu = 1$ or

$$\rho p_c \frac{1+p_c}{1-p_c} = 1$$

Solving we have $\rho p_c^2 + (\rho+1)p_c - 1 = 0$ or

$$p_c = \frac{-(\rho+1) + \sqrt{(\rho+1)^2 + 4\rho}}{2\rho}$$

In order to check this result and to prepare for developments in the next section, we will now give another (nonrigorous) derivation of the bond percolation critical value based on the fact that, seen from a fixed vertex, the NW small world is locally tree like. Color vertices blue if they are reached by a long range edge and red if they are reached by a short range edge. Ignoring collisions the growth of the cluster is a two-type branching process with mean matrix

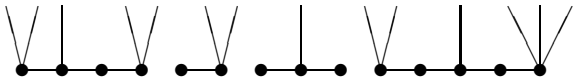
| | $B$ | $R$ |
|---|---|---|
| $B$ | $\rho$ | 2 |
| $R$ | $\rho$ | 1 |

The growth rate of this system is dictated by the largest eigenvalue of this matrix, which solves

$$0 = (\rho - \lambda)(1 - \lambda) - 2\rho = \lambda^2 - (\rho + 1)\lambda - \rho$$

Comparing with the previous quadratic equation we see that $p_c = 1/\lambda$. This is exactly what we should have expected since particles in generation $n$ of the branching process are infected in the epidemic with probability $p^n$.

**Rigorous proof of critical values.** Rather than take our usual approach of showing that the branching process accurately models the growth of the cluster, we will prove the result by reducing to a model with a fixed degree distribution. The reduction is based on the following picture



If we only use the connections around the ring then we get connected components that have a geometric distribution with success probability $r$ where $r = q$ for site percolation and $r = (1 - p)$ for bond percolation. In the case of site percolation we are ignoring closed sites, which are components of size 0. The distribution is a single geometric rather than the sum of two since we scan from left to right to find the successive components.

Now each site in the cluster is connected to a Poisson mean $\lambda = \rho p$ number of edges. In bond percolation this comes from the fact that the edge has to be open to count. In the case of site percolation this comes from the fact that of the Poisson mean $\rho$ edges, a fraction $(1 - p)$ are connected to sites that are closed. Collapsing the components to single vertices, they have degree $S_N = X_1 + \cdots X_N$ where the $X_i$ are i.i.d. Poisson($\lambda$) and $N$ is geometric with success probability $r$. Standard formulas for random sums tell us

$$\begin{aligned} ES_N &= EXEN = \lambda/r \\ \text{var}\,(S_N) &= EN \cdot \text{var}\,(X) + (EX)^2 \cdot \text{var}\,(N) \\ &= \frac{1}{r} \cdot \lambda + \lambda^2 \cdot \frac{1-r}{r^2} \end{aligned}$$

To check the form of the two terms, consider the two cases $N$ is constant and $X$ is constant. Using this we have

$$\begin{aligned} E(S_N(S_N - 1)) &= \text{var}\,(S_N) + (ES_N)^2 - ES_N \\ &= \frac{1}{r} \cdot \lambda + \lambda^2 \cdot \frac{1-r}{r^2} + \frac{\lambda^2}{r^2} - \frac{\lambda}{r} = \frac{\lambda^2}{r^2}(2 - r) \end{aligned}$$

so the mean of the size biased distribution

$$\frac{E(S_N(S_N - 1))}{ES_N} = \lambda \frac{2 - r}{r}$$

It follows that the conditions for a giant component are

$$\begin{aligned} \rho p(2-q)/q > 1 \qquad & \text{for site percolation} \\ \rho p(1+p)/(1-p) > 1 \qquad & \text{for bond percolation} \end{aligned}$$

in agreement with our previous calculations. The main point of this calculation is that using results in Section 3.2 for phase transitions in a graph with a fixed degree ditribution, it leads to a rigorous proof. However, I find it comforting that the same critical values emerge from two much different computations.

**Critical Exponents.** For the rest of their paper, Moore and Newman (2000) are concerned with the values of various critical exponents associated with the percolation process. Their computations are for the Big World graph where cluster growth is a branching process, so they apply equally well to the Erdös-Rényi random graph.

Abstracting the calcluation to simplify it, suppose we have a one parameter family of branching processes indexed by their mean $\mu$. If $\mu < 1$ then the total cluster size $\sum_{m=0}^{\infty} Z_m$ has expected value

$$E|\mathcal{C}| = E(\sum_{m=0}^{\infty} Z_m) = \sum_{m=0}^{\infty} \mu^m = \frac{1}{1-\mu}$$

Thus as $\mu \uparrow 1$, $E|\mathcal{C}| \sim (\mu_c - \mu)^{-1}$ and the critical exponent associated with the divergence of the mean cluster size is $\gamma = 1$.

Suppose now that $\mu > 1$ and consider the probability of no percolation $\rho$ which is the solution $< 1$ of $g(x) = x$. When $\mu$ is close to 1, $\rho \approx 1$. Setting $\rho = 1 - a$ and expanding the generating function to second order:

$$g(1-a) = g(1) - ag'(1) + \frac{a^2}{2} g''(b) \qquad \text{for some } b \in [a, 1]$$

Recalling $g(1) = 1$ and $g'(1) = \mu$, we see that if $g(1-a) = 1 - a$ then

$$1 - a = 1 - \mu a + \frac{a^2}{2} g''(b)$$

or $a = 2(\mu - 1)/g''(b)$. As $\mu \downarrow 1$, $g''(b) \to \mu_2 = \sum_k k(k-1)p_k$, so the critical exponent associated with the survival probability is $\beta = 1$.

Moore and Newman also compute the critical exponent for the asymptotic behavior of the cluster size distribution when $\mu = 1$, but this is the same as the calculation at the end of Section 3.1. The result is

$$P(|\mathcal{C}| = k) \sim bk^{-3/2}$$

To understand this probabilistically, we allow only one individual to reproduce in the branching process at each time, reducing the process to a mean zero random walk in which the time to hit 0 has the same distribution as $|\mathcal{C}|$.

The values we have computed are the "mean-field critical values" which hold for percolation on $\mathbb{Z}^d$ when $d$ is large enough, i.e., $d > 6$. Their appearance here indicates that the long

range connections, make the small world very big. Indeed the fact that the diameter grows like $\log n$ compared to $n^{1/d}$ in $d$-dimensional space, implies that the big and small worlds are essentially infinite dimensional.

**What have we just done?** Our computations are rigorous results for the branching process. In the random graph, the exponents only appear when we first let $n \to \infty$ and then let $\mu$ approach 1, or set $\mu = 1$ and let $k \to \infty$. For finite $n$, the expect values of the average cluster size and the fraction of vertices in the largest component are smooth, and the power law for $P(|\mathcal{C}| = k)$ will have an expoential cutoff for $k = O(n^{2/3})$, see (2.7.4) for the Erdös-Rényi case.

## 5.4 Ising and Potts models

The results in this section were inspired by Häggström (2000), but for the details we mostly follow Häggström (1998). In the Potts model, each vertex is assigned a spin $\sigma_x$ which may take one of $q$ values. Given a finite graph $G$ with vertices $V$ and edges $E$, e.g., the small world, the energy of a configuration is

$$H(\sigma) = 2 \sum_{x,y \in V, x \sim y} 1\{\sigma(x) \neq \sigma(y)\}$$

where $x \sim y$ means $x$ is adjacent to $y$. Configurations are assigned probabilities $\exp(-\beta H(\sigma))$ where $\beta$ is a variable inversely proportional to temperature, and we define a probability measure on $\{1, 2, \ldots q\}^V$ by

$$\nu(\sigma) = Z^{-1} \exp(-\beta H(\sigma))$$

where $Z$ is a normalizing constant that makes the $\nu(\sigma)$ sum to 1. When $q = 2$ this is the Ising model, though in that case it is customary to replace $\{1, 2\}$ by $\{-1, 1\}$, and write the energy as

$$H_2(\sigma) = - \sum_{x,y \in V, x \sim y} \sigma(x)\sigma(y)$$

This leads to the same definition of $\nu$ since every pair with $\sigma(x) \neq \sigma(y)$ increases $H_2$ by 2 from its minimum value in which all the spins are equal, so $H - H_2$ is constant and after normalization the measures are equal.

To study the Potts model on the small world we will use the random-cluster model. This was introduced by Fortuin and Kastelyn (1972) but Aizenman, Chayes, Chayes, and Newman (1988) were the first use this connection to prove rigorous results. See Grimmett (1995) for a nice survey. The random-cluster model is a $\{0, 1\}$-valued process $\eta$ on the edges $E$ of the graph:

$$\mu(\eta) = Z^{-1} \left\{ \prod_{e \in E} p^{\eta(e)} (1-p)^{1-\eta(e)} \right\} q^{\chi(\eta)}$$

where $\chi(\eta)$ is the number of connected components of $\eta$ when we interpret 1-bonds as occupied and 0-bonds as vacant, and $Z$ is another normalizing constant. When $q = 1$ this is just product measure

To relate the two models we introduce the following coupling on $\{1, 2, \ldots, q\}^V \times \{0, 1\}^E$

$$P(\sigma, \eta) = Z^{-1} \left\{ \prod_{e \in E} p^{\eta(e)} (1-p)^{1-\eta(e)} \right\} \prod_{x \sim y} 1\{(\sigma(x) - \sigma(y))\eta(x, y) = 0\}$$

In words, if $\eta(x, y) = 1$ then the spins at $x$ and $y$ must agree. It is easy to check, see Theorem 2.1 in Häggström (1998) for detailed proofs of this and the next three results, that

**Lemma 5.4.1.** *If $p = 1 - e^{-2\beta}$ then the projection onto the $\{1, 2, \ldots, q\}^V$ is $\nu$ and onto $\{0, 1\}^E$ is $\mu$.*

As a corollary of the coupling we see that

**Lemma 5.4.2.** *If we pick a random edge configuration according to $\mu$ and then assign random values to each connected component of edges the result is $\nu$. Conversely if we generate $\sigma \in \{1, 2, \ldots, q\}^G$ and then independently assign each edge $(x, y)$ the value 1 with probability $p$ if $\sigma(x) = \sigma(y)$, and probability 0 if $\sigma(x) \neq \sigma(y)$ then the result is $\mu$.*

To begin to analyze the Potts model, we need the following result that follows immediately from the definition.

**Lemma 5.4.3.** *Fix an edge $e = (x, y)$ and let $\eta^e$ be the values on $E - \{e\}$*

$$\mu(\eta(e) = 1 | \eta^e) = \begin{cases} p & \text{if } x \text{ and } y \text{ are connected in } \eta^e \\ \frac{p}{p+q(1-p)} & \text{otherwise} \end{cases}$$

The next ingredient is a result of Holley (1974), which we consider for the special case of $\{0, 1\}^E$. We say $f : \{0, 1\}^E \to \mathbb{R}$ is increasing if $f(\eta) \leq f(\zeta)$ whenever $\eta \leq \zeta$, i.e., $\eta(e) \leq \zeta(e)$ for all $e \in E$. Given two probability measures on $\{0, 1\}^E$, we say that $\mu_1 \leq \mu_2$ if $\int f \, d\mu_1 \leq \int f \, d\mu_2$ for all increasing $f$.

**Lemma 5.4.4.** *Let $\mu_1$ and $\mu_2$ be two measures on $\{0, 1\}^E$. Suppose that for every $e \in E$ and every $\eta, \zeta$ with $\eta^e \leq \zeta^e$*

$$\mu_1(\eta(e) = 1 | \eta^e) \leq \mu_2(\zeta(e) = 1 | \zeta^e)$$

*then $\mu_1 \leq \mu_2$,*

For a proof see Theorem 3.2 in Häggström (1998).

Introducing the parameters of $\mu$ as subscripts, it follows from Lemmas 5.4.3 and 5.4.4 that if $q > 1$

$$\mu_{p,1} \geq \mu_{p,q}$$

and that if $p' \geq p$ and $p'/(p' + q(1 - p') \geq p$ then

$$\mu_{p,1} \leq \mu_{p',q}$$

**Theorem 5.4.5.** *Let $p_c$ be the critical value for the existence of components of $O(n)$ for percolation on the graph. If*

$$p' > p_I = \frac{qp_c}{1 + (q - 1)p_c} \tag{5.4.1}$$

*then $\mu_{p',q}$ also has large components.*

Using Lemma 5.4.1 we see that if $\beta > \beta_I$ where $\beta = -(1/2)\log(1 - p_I)$ then in $\nu_{\beta,q}$ there is a large clusters of spins all of which have the same value. Turning to concrete examples:

**BC small world.** The BC small world looks locally like the tree in which each vertex has degree 3. Thinking about the growth of a cluster from a fixed vertex it is easy to see that

the critical value for percolation is $p_c = 1/2$. Using Lemma 5.4.5, $p_I = 2/3$. Lemma 5.4.1 gives $\beta_I = (1/2)\log 3 = 0.5493$. It is interesting to note that $1/\beta_I = 1.820$ while simulations of Hong, Kim, and Choi (2002) give $T_c \approx 1.82$. To lead into the next topic which will start to explain this, we begin with some arithmetic that is general:

$$\tanh(\beta_I) = \frac{1 - e^{-2\beta_I}}{1 + e^{-2\beta_I}} = \frac{p_I}{2 - p_I} = p_c$$

since $p_I = 2p_c/(1 + p_c)$.

To explain the significance of this simple calculation, consider the Ising model on a tree with forward branching number $b \geq 2$. $b$ is the degree of vertices $-1$. Define $\beta_I$, the critical value for the Ising model, by $\tanh(\beta_I) = 1/b$. This is the critical value for the onset of "spontaneous magnetization." When $\beta > \beta_I$ if we impose $+1$ boundary conditions at sites a distance $n$ from the root and let $n \to \infty$ then in the resulting limit spins $\sigma(x)$ have positive expected value. When $0 \leq \beta \leq \beta_I$ there is a unique limiting Gibbs state independent of the boundary conditions. See e.g., Preston (1974).

**NW small world, $k = 1$.** At the end of the 1980s, Russ Lyons, using an idea of Furstenburg, defined a notion of branching number $b$ for trees that are not regular. We will not give the general definition, since for us it is enough that in the case of a tree generated by a multitype branching process, the branching number $b = \lambda$ the growth rate for the process, see Lyons (1990). Lyons (1989) showed that the critical value for percolation $p_c = 1/b$ while if we convert his notation to ours by writing $\beta = J/kT$ where $T$ is the temperature and $k$ is Boltzmann's constant then $\tanh(\beta_I) = 1/b$. See Lyons (2000) for a more recent survey.

The first result gives another derivation of the conclusion $p_c = 1/\lambda$ from the previous section. The second allows us to prove of the upper bound in the next result. The Ising model on small worlds has been studied by physicists, see Barrat and Weigt (2000), Gitterman (2000), Pekalski (2001), and Hong, Kim, and Choi (2002). However, no one seems to have noticed this simple exact result.

**Theorem 5.4.6.** *For the BC small world or the nearest neighbor NW small world, the critical value for the Ising model has* $\tanh(\beta_I) = p_c$.

*Proof.* The calculations above show that for $\beta > \beta_I$ there is long range order in the Ising model in the sense that there are clusters of spins of equal value of size $O(n)$. To prove a result in the other direction we note that if $\beta < \beta_I$ then the Gibbs state on the tree is unique. There is a $c > 0$ so that for most sites $x$ in the graph if we look at the graph in a neighborhood of radius $c \log n$ around $x$, we see a tree. If we put $+1$'s on the boundary of this tree then what we see inside is larger than the Gibbs state on the small world, but if $n$ is large $P(\sigma(x) = 1) \approx 1/2$. □

**Spin glass transition.** Consider the Ising model on the tree with forward branching number $b$. Define $\beta_c^{SG}$, where the superscript $SG$ is for spin glass, by $\tanh \beta_c^{SG} = 1/\sqrt{b}$. The second transition concerns the behavior with free boundary conditions, i.e., we truncate the tree

at distance $n$ and throw away the sites outside. Bleher, Ruiz, and Zagrebnov (1995) and Ioffe (1996) showed that the limiting state is ergodic when $\beta_c^F < \beta \le \beta_c^{SG}$, but not when $\beta > \beta_c^{SG}$. Here the phrase "spin-glass" refers a model on the tree analyzed by Chayes, Chayes, Sethna, and Thouless (1986) and Carlson, Chayes, Chayes, Sethna, and Thouless (1989) which provides the key ideas for the proofs in the two papers previously cited.

To see what this second phase transition means, we consider a model of "Broadcasting on trees" considered by Evans, Kenyon, Peres, and Schulman (2000). Starting at the root, which has some value, say $+1$, each vertex receives the state of its parent with probability $1-2\epsilon$ and a randomly chosen state $\in \{-1, 1\}$ with probability $2\epsilon$. This description is supposed to remind the reader of (5.4.3), which with (5.4.1) gives

$$1 - 2\epsilon = \frac{p}{2-p} = \frac{1 - e^{-2\beta}}{1 + e^{-2\beta}} = \tanh(\beta)$$

EKPS show that the probability of correctly reconstructing the spin at the root tends to a limit $> 1/2$ if $1 - 2\epsilon > k^{-1/2}$ and to $1/2$ if $1 - 2\epsilon < k^{-1/2}$.

Q. Does this transition have any meaning for the Ising model on the BC small world?

## 5.5 Contact process

Durrett and Jung (2005) have considered the contact process (SIS epidemic) on a multi-dimensional generalization of the BC small world. To make those results fit more easily into the scheme of this chapter, we will for simplicity restrict our attention to the $d = 1$ case. Based on results for the Ising model in the previous section, we should expect that the contact process on the small world should behave like the contact process on a tree, so we begin with an account of those results.

In the contact process on any graph infected sites become healthy at rate 1, and become infected at rate $\lambda$ times the number of infected neighbors. Let $T$ be a tree in which each vertex has degree $d > 2$ and let 0 be a distinguished vertex (the origin) of the tree. Let $A_t^0$ be the set of infected sites at time $t$ on the tree starting from 0 occupied. We define two critical values:

$$\begin{aligned} \lambda_1 &= \inf\{\lambda : \mathbb{P}(|A_t^0| = 0 \text{ eventually}) < 1\} \\ \lambda_2 &= \inf\{\lambda : \liminf_{t\to\infty} \mathbb{P}(0 \in A_t^0) > 0\}. \end{aligned} \tag{5.5.1}$$

We call $\lambda_1$ the weak survival critical value and $\lambda_2$ the strong survival critical value. Pemantle (1992) showed that for trees with $d \geq 4$, $\lambda_1 < \lambda_2$. He and Liggett (1996) who extended the result to trees with $d = 3$, did this by finding numerical bounds on the two critical values which showed they were different. Later Stacey (1996) found a proof that did not rely on numerical bounds.

To explain the reason for the two critical values, consider branching random walk, which is a contact process without the restriction of one particle per site. In this process each particle dies at rate 1, and for each neighbor gives birth at rate $\lambda$ to a new particle at that site. If $Z_t^0$ is the number of particles at time $t$ starting from a single particle at time 0 then $EZ_t^0 = e^{(\lambda d-1)t}$ so $\lambda_1 = 1/d$. If we let $S_t$ be a random walk on the tree that jumps at rate $\lambda d$ to a randomly chosen neighbor then the expected number of particles at 0 at time $t$ has

$$EZ_t^0(0) = e^{(\lambda d-1)t} P(S_t = 0)$$

For a detailed proof of a similar fact see (4.4.1).

Since the distance from the origin on the tree is a random walk that steps $+1$ with probability $(d-1)/d$ and $-1$ with probability (except at 0 where all steps are $+1$) it is not hard to show that

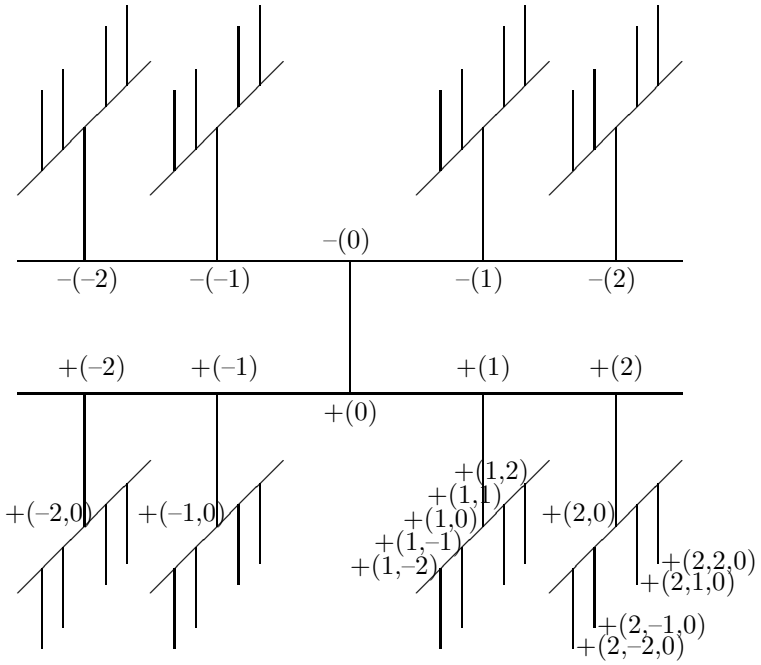
$$(1/t) \log P(S_t = 0) \to -\rho_d < 0$$

so $\lambda_c d - 1 - \rho_d = 0$ or $\lambda_2 = (1 + \rho_d)/d$. For more on the two phase transitions in branching random walk see Madras and Schinazi (1992).

Our version of the BC small world, which we will call $BC_m$, will be as follows. We start with a ring $\mathbb{Z} \bmod L$ and connect each vertex to all other vertices within distance $m$. We require $L$ to be even so that we can partition the $L$ vertices into $L/2$ pairs. Consider all such partitions and then pick one at random. A new edge is then drawn between each pair of vertices in the chosen partition. The reason for insisting that all individuals have exactly

one long-range neighbor is that we can define an associated "big world" graph $\mathcal{B}_m$ that is non-random. Algebraically, $\mathcal{B}_m$ consists of all vectors $\pm(z_1, \ldots, z_n)$ with $n \geq 1$ components with $z_j \in \mathbb{Z}$ and $z_j \neq 0$ for $j < n$. Neighbors in the positive half-space are defined as follows: a point $+(z_1, \ldots, z_n)$ is adjacent to $+(z_1, \ldots, z_n + y)$ for all $y$ with $0 < |y| \leq m$ (these are the short-range neighbors of $+(z_1, \ldots, z_n)$). The long-range neighbor is

$$\begin{array}{ll} +(z_1, \ldots, z_n, 0) & \text{if } z_n \neq 0 \\ +(z_1, \ldots, z_{n-1}) & \text{if } z_n = 0, n > 1 \\ -(0) & \text{if } z_n = 0, n = 1. \end{array}$$



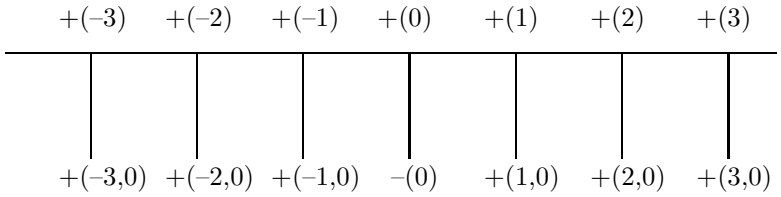

We will consider the discrete-time contact process. On either the small world or the big world, an infected individual lives for one unit of time. During its infected period it will infect some of its neighbors. All infection events are independent, and each site that receives at least one infection is occupied with an infected individual at the next time. A site infects itself or its short-range neighbors with probability $\alpha/(2m+1)$. It infects its long-range neighbor with probability $\beta$. To have a one parameter family of models we think of fixing $r = \alpha/\beta$ and varying $\lambda = \alpha + \beta$.

We will use $B_t$ to denote the contact process on the big world and $\xi_t$ for the contact process on the small world. It is easy to see that if $\alpha + \beta < 1$ the infection on the big world will die out. Our first result shows that this trivial necessary condition becomes exact when the range $m$ is large.

**Theorem 5.5.1.** *If $\alpha + \beta > 1$ then the contact process on the big world survives for large $m$.*

The proofs of this and the other results are somewhat lengthy, so we refer the reader to Durrett and Jung (2005) for details.

To obtain a lower bound $\lambda_2$, we use the fact that strong survival of the contact process on $\mathcal{B}_m$ implies strong survival of the branching random walk on $\mathcal{B}_m$. Let $\lambda_2^{brw}(m)$ be the strong survival critical value of the branching random walk. To compute the limit of $\lambda_2^{brw}(m)$, we define the "comb" of degree $m$, $\mathcal{C}_m$, by restricting $\mathcal{B}_m$ to vertices of the form $\{+(z), +(z,0), -(0)\}$ and all edges between any of these vertices. As before, $+(z)$ and $+(z,0)$ are long-range neighbors as are $+(0)$ and $-(0)$. The short-range neighbors of $+(z)$ are $+(z+y)$ for $0 < |y| \le m$. The vertices $+(z,0)$ and $-(0)$ have no short-range neighbors. To see the reason for the name look at the picture




Viewing particles on the top row as type 1 and those on the bottom row as type 2, we have a two type branching process with mean matrix:

$$\begin{pmatrix} \alpha & \beta \\ \beta & 0 \end{pmatrix}$$

Results for multitype branching processes imply that the branching random walk on the comb survives if the largest eigenvalue of the matrix is larger than 1. Solving the quadratic equation $(\alpha - \lambda)(-\lambda) - \beta^2 = 0$ the largest root is

$$\frac{\alpha + \sqrt{\alpha^2 - 4\beta^2}}{2}$$

A little algebra shows that this is larger than 1 exactly when $\alpha^2 - 4\beta^2 > (2-\alpha)^2$ or $\alpha + \beta^2 > 1$. This is an upper bound on the strong survival critical value of the branching process when what we need is a lower bound but it motivates the following:

**Theorem 5.5.2.** *If $\alpha + \beta^2 < 1$ then there is no strong survival in the contact process on the big world for large $m$.*

Comparing the above with Proposition 5.5.1 shows that for any $r = \alpha/\beta$, $\lambda_1 < \lambda_2$ for large $m$. When $m = 1$ and $\alpha = \beta$ the big world is a tree of degree 3 and we have $\lambda_1 < \lambda_2$ in that case as well. It is reasonable to conjecture that for any range $m$ and ratio $r$ we have $\lambda_1 < \lambda_2$ but this seems difficult to prove.

Since the small world is a finite graph, the infection will eventually die out. However, by analogy with results for the $d$-dimensional contact process on a finite set, we expect that if the process does not become extinct quickly, it will survive for a long time. Durrett and Liu (1988) showed that the supercritical contact process on $[0, L)$ survives for an amount of time of order $\exp(cL)$ starting from all ones, while Mountford (1999) showed that the supercritical contact process on $[0, L)^d$ survives for an amount of time of order $\exp(cL^d)$. At the moment we are only able to prove the last conclusion for the following modification of the small world contact process: each infected site infects its short-range neighbors with probability $\alpha/(2m+1)$ and its long-range neighbor with probability $\beta$, but now in addition, it infects a random neighbor (chosen uniformly from the grid) with probability $\gamma > 0$.

From a modeling point of view, this mechanism is reasonable. In addition to long-range connections with friends at school or work, one has random encounters with people one sits next to on airplanes or meets while shopping in stores. In the language of physics, the model with $\gamma = 0$ has a quenched (i.e., fixed) random environment, while the model with $\beta = 0$ has an annealed environment.

Our strategy for establishing prolonged survival is to show that if the number of infected sites drops below $\eta L$, it will with high probability rise to $2\eta L$ before dying out. To do this we use the random connections to spread the particles out so that they can grow independently. Ideally we would use the long-range connections (instead of the random connections) to achieve this; however, we have to deal with unlikely but annoying scenarios such as all infected individuals being long-range neighbors of sites that are respectively short-range neighbors of each other.

**Theorem 5.5.3.** *Consider the modified small world model on $\mathbb{Z} \bmod L$ with random infections at rate $\gamma > 0$. If $\lambda > \lambda_1$ and we start with all infected individuals then there is a constant $c > 0$ so that the probability the infection persists to time $\exp(cL)$ tends to 1 as $L \to \infty$.*

This result shows that prolonged persistence occurs for $\lambda > \lambda_1$. The next describes a change in the qualitative behavior that occurs in the contact process at $\lambda_2$. Let $\tau_{\mathcal{B}} = \min\{t : B_t^0 = \emptyset\}$ be the extinction time of the contact process on the big world. Let $\sigma_{\mathcal{B}} = \min\{t : B_t^0 = \emptyset \text{ or } 0 \in B_t^0\}$ be the first time that the infection either dies out or comes back to the origin starting from one infection there at time 0. Let $\tau_{\mathcal{S}} = \min\{t : \xi_t^0 = \emptyset\}$ and $\sigma_{\mathcal{S}} = \min\{t \geq 1 : \xi_t^0 = \emptyset \text{ or } 0 \in \xi_t^0\}$ be the corresponding times for the contact process on the small world.

**Theorem 5.5.4.** *Writing $\Rightarrow$ for convergence in distribution as $L \to \infty$ we have*
*(a) $\tau_{\mathcal{S}}$ is stochastically bounded above by $\tau_{\mathcal{B}}$ and $\tau_{\mathcal{S}} \Rightarrow \tau_{\mathcal{B}}$*
*(b) $\sigma_{\mathcal{S}}$ is stochastically bounded above by $\sigma_{\mathcal{B}}$ and $\sigma_{\mathcal{S}} \Rightarrow \sigma_{\mathcal{B}}$.*

Intuitively, when $\lambda_1 < \lambda < \lambda_2$, the infection cannot spread without the help of the long range so even if the infection starts at 0 and does not die out globally then it dies out locally and takes a long time to return to 0.

**Open Problem.** Consider the Ising model on $BC_m$. Is $\tanh(\beta_I) = p_c$ the critical value for percolation on the big world?

.