Small worlds

A. D. Barbour
Applied Mathematics
University of Zürich
CH - 8057 Zürich
and
Gesine Reinert
King's College Research Centre
UK - Cambridge CB2 1ST

April 12, 2001

Abstract

Small world models are networks consisting of many local links and fewer long range 'shortcuts'. In this paper, we consider some particular instances, and rigorously investigate the distribution of their inter–point network distances. Our results are framed in terms of approximations, whose accuracy increases with the size of the network. We also give some insight into how the reduction in typical inter–point distances occasioned by the presence of shortcuts is related to the dimension of the underlying space.

1 Introduction

In [10], Watts and Strogatz introduced a mathematical model for "small–world" networks. These networks had achieved popularity in social sciences, modelling the phenomenon of "six degrees of separation". Further examples that have been suggested are the neural network of C.elegans, the power grid of the western United States, and the collaboration graph of film actors. The work of [10] has received considerable attention during the last two years, in particular by physicists; see the Los Alamos server for condensed matter physics (`http://xxx.lanl.gov/archive/cond-mat`). However, a closely related model, the "great circle model", had already been studied by Ball *et al.* [1], in the context of epidemics.

The model proposed in [10] is as follows. Starting from a ring lattice (a 1-dimensional finite lattice with periodic boundary conditions) with L vertices, the k nearest neighbors to a vertex in clockwise direction are connected to the vertex by an undirected edge, resulting in a $2k$–nearest neighbour graph. Next, each edge is rewired at random with probability p. The procedure for this is: a vertex is chosen, and an edge that connects it to its nearest neighbour in a clockwise sense. With probability p, this edge is reconnected to a vertex chosen uniformly at random over the entire ring, with duplicate edges forbidden; otherwise the edge is left in place. The process is repeated by moving clockwise around the ring, considering each vertex in turn until one lap is completed. Next, the edges that connect vertices to their second-nearest neighbours are considered, as before. So the rewiring process stops after k laps. The quantities computed from this graph are the "characteristic path length" $\ell(p)$, defined as the number of edges in the shortest path between two vertices, averaged over all pairs of vertices, and the "clustering coefficient" $C(p)$, defined as follows. Suppose that a vertex v has k_v neighbours. Let C_v be the quotient of the number of edges between these k_v neighbours and the possible number of edges $\binom{k_v}{2}$. Define C to be the average of C_v over all v. These two quantities are computed for real networks in the examples above, as well as for the random (Bernoulli) graph with the same p. The common phenomenon observed is the "small-world phenomenon": $\ell \geq \ell_{random}$ but $C \gg C_{random}$.

Ideally, from a probabilistic view point, one would like to determine the behaviour of $\ell(p)$ and $C(p)$ so as to estimate the parameter p in the network, or so as to be able to assign statistical significance levels, when distinguishing different network models. Physicists seem to be intrigued by the scaling properties of ℓ; see [6], [7], and references therein; also they enjoy studying percolation on this graph; see [5], [6]. Percolation there is also viewed as a model for disease spread.

A closer look reveals that the above model is not easy to analyze. In particular, there is a nonzero probability of having isolated vertices, which makes $\ell(p)$ infinite with positive probability, and hence $\mathbf{E}\ell(p) = \infty$. As a result, it was soon revised by not rewiring edges, but rather adding edges, thus ensuring that the graph stays connected; see, for example, [6]. More precisely, a number of shortcuts are added between randomly chosen pairs of sites with probability ϕ per connection on the underlying lattice, of which there are Lk. Thus, on average, there are $Lk\phi$ shortcuts in the graph.

Recently, Newman, Moore and Watts [6], [7] gave a heuristic computation (the NMW heuristic) of ℓ in this modified graph. They suggest that

$$\mathbf{E}\ell = \frac{L}{k} f(Lk\phi), \tag{1.1}$$

where

$$f(z) = \frac{1}{2\sqrt{z^2+2z}} \tanh^{-1} \sqrt{\frac{z}{z+2}}.$$

In particular,

$$f(z) \sim \begin{cases} \frac{1}{4} & \text{for } z \ll 1 \\ \frac{\log 2z}{4z} & \text{for } z \gg 1. \end{cases}, \tag{1.2}$$

giving

$$\mathbf{E}\ell \sim \begin{cases} \frac{L}{4k} & \text{for } Lk\phi \ll 1 \\ \frac{\log 2Lk\phi}{4k^2\phi} & \text{for } Lk\phi \gg 1. \end{cases}.$$

Their heuristic is based on mean field approximations, replacing random variables by their expectations.

In the context of epidemics in a spatially structured population, Ball *et al.* consider individuals on a large circle in their "great circle model". They allow only nearest–neighbour ($k = 1$) connections, but claim that most of their results can be extended to general k. An SIR epidemic is studied, where each individual has a probability p_L of infecting a neighbour, and probability p_G of infecting any other individual on the circle; typically, $p_L \gg p_G$. In the SIR framework, this model corresponds to individuals having a fixed infectious period of duration 1. The structure of the graph at time $T = \infty$ is their main object of interest. In terms of small worlds, their model broadly corresponds to having an epidemic on a small-world network with parameter ϕ, where $\phi = p_G/p_L$.

In this paper, we analyze a continuous model, introduced in [5], in which a random number of chords, with Poisson distribution $\mathrm{Po}\,(L\rho/2)$, are uniformly and independently superimposed as shortcuts on a circle of circumference L. Distance is measured as usual along the circumference, and chords are deemed to be of length zero. When L is large, this model approximates the k–neighbour model of [6] if $\rho = 2k\phi$, except that distances should also be divided by k, because unit graph distance in the k–neighbour model covers an arc length of $2k$, rather than 2. In the case when the expected number $L\rho/2$ of shortcuts is large, we prove a distributional approximation for the distance between a randomly chosen pair of points P and P', and give a bound on the order of the error, in terms of total variation distance. This distribution differs, in both location and spread, from that suggested by the NMW heuristic, though to the coarsest order $O(\frac{1}{\rho}\log(L\rho))$ agrees with that suggested by (1.2). We also show that analogous results can be proved in higher dimensions by much the same method, when the circle is replaced by a sphere or a torus; here, the reduction in the typical distance between pairs of points occasioned by shortcuts is less dramatic than in one dimension.

2 The continuous circle model: construction and heuristics

In this section, we consider a continuous model consisting of a circle C of circumference L, to which are added a Poisson $\mathrm{Po}\,(L\rho/2)$ number of uniform and independent random chords. We begin with a dynamic realization of the network, which describes, for each $t \geq 0$, the set of points $R(t) \subset C$ that can be reached from a given point P within time t: time corresponds to arc distance, with chords of length zero. Such a realization is also the basis for the NMW heuristic.

Pick Poisson $\mathrm{Po}\,(L\rho)$ points of the circle C uniformly and independently, and call this set Q. The elements $q \in Q$ are called *potential* label 1 end points of chords. To each $q \in Q$, assign a second independent uniform point of C, say $q' = q'(q)$: the label 2 end point. The unordered pairs $\{q, q'\}$ form the *potential* chords. Only a random subset of the potential chords are actually realized. Let $R(t)$ be the union of the $B(t)$ intervals of C, each of which increases with time, growing deterministically at rate 1 at each end point; we start with $R(0) = \{P\}$. Whenever $card(\partial R(t-) \cap Q) = 1$ — that is, whenever the boundary of $R(t-)$ reaches a potential label 1 end of a chord (note that the intersection never contains more than one element with probability 1) — so that $\partial R(t-) \cap Q = \{q\}$, say, accept the chord $\{q, q'\}$ if $q' \notin R(t-)$ (that is, if the chord would reach beyond the cluster $R(t-)$) and take $R(t) = \overline{R(t-)} \cup \{q'\}$; otherwise, take $R(t) = \overline{R(t-)}$. This defines a predictable thinning of the set of potential chords, to obtain the set of actual, accepted chords. The number of intervals increases by 1 whenever a chord is accepted, and decreases by 1 whenever two intervals grow into one another.

The intensity of adding chords is $2\rho B(t)\{1 - r(t)L^{-1}\}$, with the label 2 end points uniformly distributed over $R^c(t)$, where $r(t) = |R(t)|$, the Lebesgue measure of $R(t)$. The integrated intensity is thus

$$\int_0^\infty 2\rho B(t)\{1 - r(t)L^{-1}\}dt = L\rho/2 \quad \text{a.s.},$$

since $\frac{d}{dt} r(t) = 2B(t)$ a.e. with respect to Lebesgue measure, and

$$r(\infty) = r(L/2) = L.$$

Thus we generate $\mathrm{Po}\,(L\rho/2)$ accepted chords. To see that they are uniformly chosen, simply note that this is true of the potential chords, and that each potential chord is accepted with probability $1/2$, independently of the number and positions of all potential chords, according to whether the first of its end points to belong to R had initially been chosen as the label 1 or the label 2 end point. Thus this growth and merge construction indeed results in $\mathrm{Po}\,(L\rho/2)$ chords, uniformly distributed over C.

The NMW heuristic takes the equation

$$\frac{d}{dt}r(t) = 2B(t), \tag{2.1}$$

and adds to it an equation

$$\frac{d}{dt}B(t) = 2\rho B(t)\{1 - r(t)L^{-1}\} - 2B(t)(B(t) - 1)(L - r(t))^{-1}, \tag{2.2}$$

derived by treating the discrete variable $B(t)$ as continuous and the corresponding jump rates as differential rates. The final term in (2.2), describing the rate of merging, follows from the observation that the smallest interval between n points scattered uniformly on a circle of circumference c has an approximately exponential distribution with mean $c/n(n-1)$, and that unreached intervals shrink at rate 2. These equations have an explicit solution $\hat{r}$ and $\hat{B}$, with (for $\rho > 0$)

$$1 - \hat{r}(Lw)/L = \frac{(a+1)^2 - (a-1)^2 e^{2aL\rho w}}{2\{(a+1) + (a-1)e^{2aL\rho w}\}} =: \hat{p}(w), \tag{2.3}$$

where $a = \sqrt{1 + 4/L\rho}$; this is used for

$$0 \le w \le \frac{1}{aL\rho}\log\{(a+1)/(a-1)\} =: w^*,$$

the range in which $\hat{p}(w) \ge 0$. Then, if D is the random variable denoting the distance from P to a randomly chosen point, the NMW heuristic takes

$$\mathbf{P}_{NMW}[D > t] = 1 - \hat{r}(t)/L$$

as an approximation to the true value $\mathbf{P}[D > t] = 1 - \mathbf{E}r(t)/L$, where $r(t)$ denotes the random quantity defined in the growth and merge model. The formula for $\mathbf{E}\ell$ resulting from this heuristic is then $\frac{L}{k}\int_0^{w^*} \hat{p}(w)\,dw$ with $\rho = 2k\phi$, the factor $1/k$ arising from the definition of graph distance in the k–neighbour model, as observed above.

Note that the NMW heuristic always gives $\mathbf{P}_{NMW}[D > Lw^*] = 0$. However, since the probability of having *no* shortcuts is $e^{-L\rho/2}$, it is clear that in fact

$$\mathbf{P}[D > Lw^*] > e^{-L\rho/2}(\tfrac{1}{2} - w^*) > e^{-L\rho/2}\frac{1}{3(1+4/L\rho)},$$

so that their heuristic cannot give accurate results unless $L\rho$ is either very small or very big. If $L\rho$ is very small, then $\mathbf{P}_{NMW}[D > t] = 1 - 2t/L + O(L\rho)$, and the same is true for $\mathbf{P}[D > t]$, reflecting that there are no shortcuts, except for a probability of order $O(L\rho)$. The interesting case is that in which there are many shortcuts, when $L\rho$ is large, and this we investigate rigorously.

For our analysis, it will be convenient also to define a second process $S(t)$, starting from the same P and the same set of potential chords and with the same unit growth rate. The main differences are that *every* potential chord is included, so that no thinning takes place, and, additionally, whenever two intervals intersect, they continue to grow, overlapping one another, and each continues to generate further chords according to a Poisson process of rate ρ. This pure growth process $S(t)$ agrees with the original construction during the initial development with high probability, until S has grown enough that overlap becomes likely; its advantage is that it has a branching structure, and is thus much more easily analysed. We denote its length at time t by $s(t) \geq r(t)$, and the number of intervals by $M(t) \geq B(t)$. Then $M(t)$ is just a Yule process with birth rate 2ρ, so that $\mathbf{E}M(t) = e^{2\rho t}$, and $s(t) = \int_0^t 2M(u)du$, so that $\mathbf{E}s(t) = \rho^{-1}(e^{2\rho t} - 1)$. Furthermore, $e^{-2\rho t}M(t) \to W$ a.s., where $W \sim NE(1)$, and hence $e^{-2\rho t}s(t) \to \rho^{-1}W$ a.s. and $\frac{s(t)}{M(t)} \to \rho^{-1}$ a.s.

Our strategy will be to pick a starting point P and run both constructions up to a time $\tau_x = \frac{1}{2\rho}\left\{\frac{1}{2}\log(L\rho) + x\right\}$, for any fixed $|x| \leq \frac{1}{4}\log(L\rho)$: since we are interested in asymptotics as $L\rho \to \infty$, we may safely assume that $L\rho > 1$. Letting $R_x = R(\tau_x), S_x = S(\tau_x), M_x = M(\tau_x)$, and $s_x = s(\tau_x)$, we have $\mathbf{E}M_x = (L\rho)^{1/2}e^x$ and $\mathbf{E}s_x = \rho^{-1}\left((L\rho)^{1/2}e^x - 1\right)$. Then, independently and uniformly, we pick a second point $P' \in C$, and a second set of potential chords, Q', and we run the constructions for time τ_x once more, yielding $R'_x, S'_x, M'_x =: N_x$ and $s'_x =: u_x$. Then there are about $L\rho e^{2x}$ pairs of intervals, with one in S_x and the other in S'_x, and each is of typical length ρ^{-1}, so that the expected number of intersecting pairs of intervals is about $\frac{2}{L\rho} L\rho e^{2x} = 2e^{2x}$, which, in the chosen range of x, grows from $2(L\rho)^{-1/2}$ to $2(L\rho)^{1/2}$. We show that the number $\widehat{V}_x$ of such intersections approximately has a mixed Poisson distribution, and identify the mixture distribution. For later use, label the intervals in S_x as $I_1, \ldots, I_{M_x}$, and the intervals in S'_x as $J_1, \ldots, J_{N_x}$, where the indices are assigned in chronological order of birth. Let $p_I(l)$ denote the unique index $j < l$ of the interval I_j to which I_l is linked by a chord — the 'parent' of I_l — and define p_J analogously. Then write

$$\widehat{V}_x = \sum_{i=1}^{M_x} \sum_{j=1}^{N_x} X_{ij}, \tag{2.4}$$

where

$$X_{ij} = \mathbf{1}\{I_i \cap J_j \neq \emptyset\}. \tag{2.5}$$

Now the probability that $\widehat{V}_x = 0$ is the same as when the second construction, for S', uses the original set Q of potential chords, because of the independence of Poisson processes on disjoint subsets; the event $\widehat{V}_x = 0$ indicates that the two processes have no intersecting pairs of intervals up to τ_x, and thus use disjoint sets of chords. We then show that the event $\widehat{V}_x = 0$ is with high probability the same as the event $V_x = 0$, where V_x is the number of intersections of R_x and R'_x. Finally, if R_x and R'_x have no intersections, the "small worlds" distance between P and P' is more than $2\tau_x$.

3 The continuous circle model: proofs

The first step in the argument outlined above is to establish a Poisson approximation theorem for the number of pairs of overlapping intervals, one in S_x and the other in S'_x. This is based on the following general result.

Theorem 3.1 *Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be independent random elements, and let $\phi_{ij} := \phi_{ij}(X_i, Y_j)$ be indicator random variables with $p_{ij} := \mathbf{E}\phi_{ij}$ such that*

$$\max\{\max_k \mathbf{E}(\phi_{ij}\phi_{ik}), \max_l \mathbf{E}(\phi_{ij}\phi_{lj})\} \le pp_{ij},$$

for all pairs i, j. Then, if $\Phi := \sum_{i=1}^m \sum_{j=1}^n \phi_{ij}$,

$$d_{TV}(\mathcal{L}(\Phi), \mathrm{Po}\,(\mathbf{E}\Phi)) \le (2(m+n) - 3)p.$$

Proof: Using the local version of the Stein–Chen method ([2], Theorem 1.A), we have

$$d_{TV}(\mathcal{L}(\Phi), \mathrm{Po}\,(\mathbf{E}\Phi)) \le (\mathbf{E}\Phi)^{-1} \sum_{i=1}^m \sum_{j=1}^n (p_{ij}^2 + p_{ij}\mathbf{E}Z_{ij} + \mathbf{E}(\phi_{ij}Z_{ij})),$$

where

$$Z_{ij} := \sum_{\substack{l=1 \\ l \ne j}}^m \phi_{lj} + \sum_{\substack{l=1 \\ l \ne i}}^m \phi_{il}.$$

By assumption, $\mathbf{E}(\phi_{ij}Z_{ij}) \le (m+n-2)pp_{ij}$, and the result follows. $\square$

This theorem, in the context of intervals scattered on the circle, has the following direct consequence.

Corollary 3.2 *Let m intervals $I_1, \ldots, I_m$ with lengths $s_1, \ldots, s_m$ and n intervals $J_1, \ldots, J_n$ with lengths $u_1, \ldots, u_n$ be positioned uniformly and independently on C. Set $V := \sum_{i=1}^m \sum_{j=1}^n X_{ij}$, where $X_{ij} := I[I_i \cap J_j \ne \emptyset]$. Then*

$$d_{TV}(\mathcal{L}(V), \mathrm{Po}\,(\lambda_{(m,n,s,u)})) \le 4(m+n)l_{su}/L,$$

where $\lambda_{(m,n,s,u)} := L^{-1}(ns + mu)$, $s := \sum_{i=1}^m s_i$, $u := \sum_{j=1}^n u_j$ and $l_{su} := \max\{\max_i s_i, \max_j u_j\}$.

Proof: We apply Theorem 3.1 with X_i and Y_j the centres of the intervals I_i and J_j, and with X_{ij} for ϕ_{ij}; the X_{ij} are pairwise independent, and satisfy $\mathbf{E}X_{ij} \le 2l_{su}L^{-1}$. It remains only to note that $\mathbf{E}V = \sum_{i=1}^m \sum_{j=1}^n (s_i + u_j)/L = \lambda_{(m,n,s,u)}$. $\square$

The corollary translates immediately into a useful statement about $\widehat{V}_x$, when P' is chosen uniformly at random, independently of all else.

Corollary 3.3 *For the processes S and S' of the previous section, we have*

$$|\mathbf{P}[\widehat{V}_x = 0 \,|\, M_x = m, N_x = n, s_x = s, u_x = u] - \exp\{-L^{-1}(ns + mu)\}|$$
$$\leq 8(m+n)\tau_x L^{-1}.$$

PROOF: It suffices to note that all the intervals of S_x and S'_x are of length at most $2\tau_x$. []

Remark. If P' is not chosen at random, but is a fixed point of C, the result of Corollary 3.3 remains essentially unchanged, provided that P and P' are more than an arc distance of $2\tau_x$ apart. The only difference is that then $X_{11} = 0$ a.s., and that $ns + mu$ is replaced by $ns + mu - 4\tau_x$. If P and P' are less than $2\tau_x$ apart, then $\mathbf{P}[\widehat{V}_x = 0] = 0$.

The next step is to show that $\mathbf{P}[\widehat{V}_x = 0]$ is close to $\mathbf{P}[V_x = 0]$. We do this by directly comparing the random variables $\widehat{V}_x$ and V_x in the joint construction.

Theorem 3.4 *Let Y_{il}, $1 \leq l < i \leq n$ be possibly dependent indicator random variables, with $\mathbf{P}[Y_{il} = 1] \leq p$ for all $1 \leq l < i \leq n$. Suppose that each index $2 \leq i \leq n$ is randomly assigned a 'parent' index $P(i) \in \{0, 1, \ldots, i-1\}$, independently of everything else, from a distribution satisfying*

$$c \min_{1 \leq l \leq j} \mathbf{P}[P(i) = l] \geq \mathbf{P}[P(i) = j+1], \quad 1 \leq j \leq i-2, \tag{3.1}$$

for some $c \geq 1$. Set $H_1 = 0$, and then define $H_i = 0$ if both $Y_{i1} = Y_{i2} = \cdots = Y_{i,i-1} = 0$ and $H_{P(i)} = 0$, setting $H_i = 1$ otherwise. Then

$$\mathbf{P}[H_i = 1] \leq (1+c)(i-1)p.$$

PROOF: The event $H_i = 1$ can occur because $Y_{il} = 1$ for some $l < i$, or because i has an ancestor l (at some remove) such that $Y_{lr} = 1$ for some $r < l$. Hence

$$\{H_i = 1\} \subset \left\{ \bigcup_{1 \leq l < i} \{Y_{il} = 1\} \right\} \cup \left\{ \bigcup_{1 \leq r < l < i} (\{Y_{lr} = 1\} \cap \{A^l(i) = l\}) \right\},$$

where, considering $\{1, 2, \ldots, l\}$ as the set of ancestors, $A^l(i) \in \{1, 2, \ldots, l\}$ denotes which of them was the ancestor of i. Now, following the line of descent of i backwards in time until it first reaches $\{1, 2, \ldots, l\}$, suppose that the last index larger than l is m. Then the probability that m is connected to l is

$$\mathbf{P}[P(m) = l] \Big/ \sum_{j=1}^{l} \mathbf{P}[P(m) = j] \leq c/l,$$

in view of (3.1), so that

$$\mathbf{P}[A^l(i) = l \,|\, Y_{il},\, 1 \leq l < i \leq n] \leq c/l.$$

Hence

$$\mathbf{P}[H_i = 1] \leq (i-1)p + c \sum_{1 \leq l < i} (l-1)p/l \leq (1+c)(i-1)p,$$

as required. $\Box$

Corollary 3.5 *With notation as above, we have*

$$\mathbf{P}[\widehat{V}_x \neq V_x] \leq 32\tau_x^2 L^{-2} \mathbf{E}\{\tfrac{1}{2} M_x N_x (M_x + N_x - 2)\}.$$

PROOF: Let $Y_{il} = I[I_i \cap I_l \neq \emptyset]$ and define H_i as for Theorem 3.4, with $P(i) = p_I(i)$; note that then (3.1) is satisfied with equality for $c = 1$. Then define $Y'_{jl} = I[J_j \cap J_l \neq \emptyset]$ and define H'_j as for Theorem 3.4, with $P(j) = p_J(j)$. Now

$$\widehat{V}_x \geq V_x \geq \sum_{i=1}^{M_x} \sum_{j=1}^{N_x} X_{ij} I[H_i = H'_j = 0], \tag{3.2}$$

so that

$$\widehat{V}_x - V_x \leq \sum_{i=1}^{M_x} \sum_{j=1}^{N_x} X_{ij}(1 - I[H_i = H'_j = 0]),$$

and hence

$$\begin{aligned} \mathbf{P}[\widehat{V}_x \neq V_x] &\leq \mathbf{E}(\widehat{V}_x - V_x) \\ &\leq \mathbf{E}\left\{ \sum_{i=1}^{M_x} \sum_{j=1}^{N_x} X_{ij}(I[H_i = 1] + I[H'_j = 1]) \right\}. \end{aligned} \tag{3.3}$$

Conditioning on $M_x = m$ and $N_x = n$ and on the lengths $s_1, \ldots, s_m$ of the intervals I_i and $u_1, \ldots, u_n$ of the intervals J_j, we can apply Theorem 3.4 to give

$$\mathbf{E}(X_{ij} I[H_i = 1] \,|\, M_x = m, N_x = n, s_1, \ldots, s_m, u_1, \ldots, u_n)$$

$$\leq \frac{4\tau_x}{L}\mathbf{P}[H_i = 1 \mid M_x = m, N_x = n, s_1, \ldots, s_m, u_1, \ldots, u_n)$$
$$\leq \frac{4\tau_x}{L}\, 2(i-1)\, \frac{4\tau_x}{L}. \tag{3.4}$$

Combining (3.3) and (3.4), and arguing similarly for $\mathbf{E}(X_{ij} I[H_j' = 1])$, it thus follows that

$$\mathbf{P}[\widehat{V}_x \neq V_x] \leq \frac{32\tau_x^2}{L^2}\mathbf{E}\left\{\tfrac{1}{2} M_x N_x (M_x + N_x - 2)\right\},$$

as required. □

To apply Corollaries 3.3 and 3.5, it remains to establish more detailed information about the distributions of M_x and s_x. In particular, we need to bound the first and second moments of M_x, and to approximate the quantity $\mathbf{E}(\exp\{-L^{-1}(N_x s_s + M_x u_x)\})$. As from now, we assume that $L\rho \geq 16$. We begin with the following lemma.

Lemma 3.6 *For any x with $|x| \leq \frac{1}{4}\log(L\rho)$, we have*

$$\mathbf{P}[M_x \geq r] = (1 - q_x)^{r-1}, \quad r \geq 1,$$

where $q_x = e^{-2\rho\tau_x}$. Hence, in particular,

$$\mathbf{E}M_x = e^{2\rho\tau_x} = e^x\sqrt{L\rho}; \quad \tfrac{1}{2}\mathbf{E}\{M_x(M_x+1)\} = e^{2x}L\rho; \tag{3.5}$$

furthermore,

$$\mathbf{E}\{e^{-\theta s_x} | M_x = m\} = e^{-m\theta/\rho} + O\left(\theta\tau_x + m\theta^2\rho^{-2} + m\theta\tau_x e^{-2\rho\tau_x} + \theta\rho^{-1}\right), \tag{3.6}$$

uniformly in $|x| \leq \frac{1}{4}\log(L\rho)$.

Proof: Let $z_t = z_t(\theta, w) := \mathbf{E}\left(e^{-\theta s(t)} w^{M(t)}\right)$; then, by splitting at the first jump, which is exponentially distributed with mean $1/2\rho$, we have

$$z_t = we^{-2\rho t - 2\theta t} + \int_0^t 2\rho e^{-2\rho u}\left\{z_{t-u}^2 e^{-2\theta u}\right\} du, \tag{3.7}$$

implying that

$$\frac{dz}{dt} = -2z(\theta + \rho - \rho z); \quad z_0 = w. \tag{3.8}$$

Solving the differential equation, it follows that

$$z_t(\theta, w) = \frac{w(\theta+\rho)e^{-2(\theta+\rho)t}}{\theta+\rho-\rho w\left(1-e^{-2(\theta+\rho)t}\right)}. \tag{3.9}$$

Putting $\theta = 0$ gives the probability generating function of $M(t)$:

$$\mathbf{E}\left(w^{M(t)}\right) = we^{-2\rho t}\left(1-w(1-e^{-2\rho t})\right)^{-1}, \tag{3.10}$$

so that

$$\mathbf{P}[M(t)=m] = e^{-2\rho t}\left(1-e^{-2\rho t}\right)^{m-1}, \quad m \geq 1. \tag{3.11}$$

Note that, by the geometric series expansion, we have

$$z_t(\theta, w) \quad = \quad we^{-2(\theta+\rho)t}\sum_{m\geq 0} w^m \left(\frac{\rho}{\rho+\theta}\right)^m \left(1-e^{-2(\theta+\rho)t}\right)^m. \tag{3.12}$$

Hence, taking the coefficient of w^m in (3.12),

$$\mathbf{E}\left(e^{-\theta s(t)}|M(t)=m\right) = e^{-2\theta t}\left\{\frac{\rho}{\rho+\theta}\left[\frac{1-e^{-2(\theta+\rho)t}}{1-e^{-2\rho t}}\right]\right\}^{m-1}. \tag{3.13}$$

This completes the proof of all but (3.6). For this last, note that

$$\begin{aligned}
&\left(\frac{\rho}{\rho+\theta}\right)^r \left|\left\{\frac{1-e^{-2(\theta+\rho)\tau_x}}{1-e^{-2\rho\tau_x}}\right\}^r - 1\right| \\
&= \left(\frac{\rho}{\rho+\theta}\right)^r \left|\left\{1+\left(\frac{e^{-2\rho\tau_x}}{1-e^{-2\rho\tau_x}}\right)\left(1-e^{-2\theta\tau_x}\right)\right\}^r - 1\right| \\
&\leq \left(\frac{\rho}{\rho+\theta}\right)\frac{2r\theta\tau_x e^{-2\rho\tau_x}}{1-e^{-2\rho\tau_x}}\left[1+\frac{e^{-2\rho\tau_x}2\theta\tau_x}{1-e^{-2\rho\tau_x}}\right]^{r-1}\left(\frac{\rho}{\rho+\theta}\right)^{r-1} \\
&\leq 4r\theta\tau_x e^{-2\rho\tau_x}\left\{\left(1+\frac{e^{-2\rho\tau_x}2\theta\tau_x}{1-e^{-2\rho\tau_x}}\right)\left(\frac{\rho}{\rho+\theta}\right)\right\}^{r-1}
\end{aligned} \tag{3.14}$$

since $|1-e^{-2\theta\tau_x}| \leq 2\theta\tau_x$ and $L\rho \geq 16$. Then, since $y\frac{e^{-y}}{1-e^{-y}} \leq 1$ for $y \geq 0$, we have

$$\frac{\rho}{\rho+\theta}\left(1+\frac{e^{-2\rho\tau_x}2\theta\tau_x}{1-e^{-2\rho\tau_x}}\right) \leq 1;$$

furthermore,

$$\left|\left(\frac{\rho}{\rho+\theta}\right)^r - e^{-r\frac{\theta}{\rho}}\right| \quad \leq \quad r|\log(1+\theta/\rho)-\theta/\rho| \leq \frac{r\theta^2}{2\rho^2}. \tag{3.15}$$

Combining these inequalities, it follows that

$$\begin{aligned}
&\left|\mathbf{E}\left\{e^{-\theta s_x} | M_x = m\right\} - e^{-m\frac{\theta}{\rho}}\right| \\
&\quad \leq 4m\theta\tau_x e^{-2\rho\tau_x} + 2\theta\tau_x + \tfrac{1}{2}m\theta^2\rho^{-2} + \theta\rho^{-1}, \qquad (3.16)
\end{aligned}$$

as required, if $|x| \leq \frac{1}{4}\log(L\rho)$. []

These estimates can already be applied to Corollaries 3.3 and 3.5.

Corollary 3.7 *We have*

$$\begin{aligned}
&|\mathbf{P}[V_x = 0] - \mathbf{E}\exp\{-L^{-1}(N_x s_x + M_x u_x)\}| \\
&\quad \leq 6e^x(L\rho)^{-1/2}\log(L\rho) + 9e^{3x}(L\rho)^{-1/2}\log^2(L\rho).
\end{aligned}$$

We now estimate the quantity $\mathbf{E}(\exp\{-L^{-1}(N_x s_x + M_x u_x)\})$. Since $e^{-2\rho t}M(t) \to W$ a.s. and $s(t)/M(t) \to \rho^{-1}$ a.s., and since $e^{2\rho\tau_x} = e^x\sqrt{L\rho}$, it is clear that

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

so that $\mathbf{P}[V_x = 0]$ can be approximated in terms of the distribution of the limiting random variable W associated with the Yule process M. Here, we make the precise calculations.

Lemma 3.8

$$\begin{aligned}
&\mathbf{E}\left\{e^{-(N_x s_x + M_x u_x)L^{-1}}\right\} - \int_0^\infty \frac{e^{-y}}{1 + c_x y}dy \\
&\quad = O\left\{(L\rho)^{-\frac{1}{2}}\left(e^{3x} + e^x\log(L\rho) + e^{-x}\right)\right\}
\end{aligned}$$

uniformly in $|x| \leq \frac{1}{4}\log(L\rho)$, *where* $c_x = 2e^{2x}$.

Remark. Recall that $W = \lim_{t\to\infty} e^{-2\rho t}M(t)$ and $W' = \lim_{t\to\infty} e^{-2\rho t}N(t)$. Note that, as W and W' are independent and exponentially distributed with mean 1, we have

$$\int_0^\infty \frac{e^{-y}}{1 + c_x y}dy = \mathbf{E}\left(e^{-2WW'e^{2x}}\right). \qquad (3.17)$$

PROOF: We start by computing

$$\mathbf{E}\left\{e^{-(N_x s_x+M_x u_x)L^{-1}}|M_x=m,N_x=n\right\} \qquad (3.18)$$
$$= \mathbf{E}\left\{e^{-n s_x L^{-1}}|M_x=m\right\}\mathbf{E}\left\{e^{-m u_x L^{-1}}|N_x=n\right\};$$

using (3.16) with $\theta=n/L$ and then with $\theta=m/L$, it follows that

$$\left|\mathbf{E}\left\{e^{-(N_x s_x+M_x u_x)L^{-1}}|M_x=m,N_x=n\right\}-e^{-2\frac{mn}{L\rho}}\right| \qquad (3.19)$$
$$\leq \frac{8mn}{L}\tau_x e^{-2\rho\tau_x}+\frac{2(n+m)}{L}\tau_x+\frac{1}{2}\frac{mn(m+n)}{(L\rho)^2}+\frac{m+n}{L\rho},$$

and thus, from Lemma 3.6,

$$\left|\mathbf{E}\left\{e^{-(N_x s_x+M_x u_x)L^{-1}}\right\}-\mathbf{E}e^{-2\frac{M_x N_x}{L\rho}}\right| \qquad (3.20)$$
$$\leq 3e^x\log(L\rho)(L\rho)^{-\frac{1}{2}}+\tfrac{3}{2}e^x\log(L\rho)(L\rho)^{-\frac{1}{2}}+2e^{3x}(L\rho)^{-\frac{1}{2}}+2e^x(L\rho)^{-1/2}.$$

Then, from Lemma 3.6, since $M_x\sim Ge\left(1-e^{-2\rho\tau_x}\right)$, we have

$$\begin{aligned}\mathbf{E}\left\{e^{-2M_xN_x(L\rho)^{-1}}|N_x=n\right\} &= \frac{e^{-(\frac{2n}{L\rho}+2\rho\tau_x)}}{1-e^{-\frac{2n}{L\rho}}\left(1-e^{-2\rho\tau_x}\right)}\\ &= \left\{e^{2\rho\tau_x}\left(e^{\frac{2n}{L\rho}}-1\right)+1\right\}^{-1}.\end{aligned}$$

Observing that $|(1+z)^{-1}-(1+y)^{-1}|\leq\frac{|z-y|}{(1+z)(1+y)}$ in $z,y\geq 0$, it follows that

$$\left|\left(b\left(e^d-1\right)+1\right)^{-1}-(1+bd)^{-1}\right|$$
$$\leq \frac{1}{2}\frac{bd^2e^d}{\{1+b\left(e^d-1\right)\}\{1+bd\}}\leq b^{-1}(2\vee d),$$

and hence that

$$\left|\mathbf{E}\left\{e^{-2M_xN_x(L\rho)^{-1}}\right\}-\mathbf{E}\left\{\left(1+\frac{2e^{2\rho\tau_x}N_x}{L\rho}\right)^{-1}\right\}\right|$$
$$\leq 2e^{-2\rho\tau_x}\left(1+\frac{\mathbf{E}N_x}{L\rho}\right)\leq\frac{2e^{-x}}{\sqrt{L\rho}}+\frac{2}{L\rho}. \qquad (3.21)$$

Finally,

$$\left|\sum_{r\geq 1}q\frac{(1-q)^{r-1}}{1+ar}-\int_1^\infty\frac{qe^{-q(w-1)}}{1+aw}dw\right|$$

$$
\begin{aligned}
&\leq \quad \frac{q}{1+a} + \left| \sum_{r\geq 2} q\frac{(1-q)^{r-1}}{1+ar} - \int_1^\infty \frac{qe^{-q(w-1)}}{1+aw}dw \right| \\
&\leq \quad \frac{q}{1+a} + \int_1^\infty \frac{qe^{-q(w-1)}}{1+aw}dw \leq \frac{2q}{1+a}
\end{aligned}
$$

and

$$
\begin{aligned}
&\left| \int_1^\infty \frac{qe^{-q(w-1)}}{1+aw}dw - \int_0^\infty \frac{e^{-y}}{1+aq^{-1}y}dy \right| \\
&\quad \leq \quad \int_0^\infty \frac{e^{-y}a}{(1+aq^{-1}y)\left(1+a\left(q^{-1}y+1\right)\right)}dy \leq a,
\end{aligned}
$$

so that

$$
\begin{aligned}
&\left| \mathbf{E}\left\{ \left(1+\frac{2e^{2\rho\tau_x}N_x}{L\rho}\right)^{-1} \right\} - \int_0^\infty \frac{e^{-y}}{1+c_x y}dy \right| \qquad (3.22) \\
&\leq \quad 2e^{-2\rho\tau_x} + \frac{2e^{2\rho\tau_x}}{L\rho} \leq 2e^{-x}(L\rho)^{-\frac{1}{2}} + 2e^{x}(L\rho)^{-\frac{1}{2}}.
\end{aligned}
$$

Combining (3.20) – (3.22), it follows that

$$
\begin{aligned}
&\left| \mathbf{E}\left\{ e^{-(N_x s_x + M_x u_x)L^{-1}} \right\} - \int_0^\infty \frac{e^{-y}}{1+c_x y}dy \right| \qquad (3.23) \\
&\leq \quad \left[\tfrac{17}{2}e^{x}\log(L\rho) + 2e^{3x} + 4e^{-x} + 2(L\rho)^{-\frac{1}{2}} \right](L\rho)^{-\frac{1}{2}}.
\end{aligned}
$$

[]

Theorem 3.9 *If P' is randomly chosen on C, or if P' is any fixed point of C at arc distance more than $(3/4\rho)\log(L\rho)$ from P, then*

$$
\begin{aligned}
&\left| \mathbf{P}\left[D > \frac{1}{\rho}\left(\frac{1}{2}\log(L\rho)+x\right) \right] - \int_0^\infty \frac{e^{-y}dy}{1+c_x y} \right| \\
&\quad = \quad O\left((L\rho)^{-\frac{1}{2}} \left\{ e^{x}(1+e^{2x})\log^2(L\rho) + e^{-x} \right\} \right),
\end{aligned}
$$

uniformly in $|x| \leq \frac{1}{4}\log(L\rho)$, where, as before, D denotes the shortest distance between P and P' on the shortcut graph.

Proof: Since

$$
\{V_x = 0\} = \left\{ D > \frac{1}{\rho}\left(\frac{1}{2}\log(L\rho)+x\right) \right\},
$$

the theorem follows from Corollary 3.7 and Lemma 3.8. []

Corollary 3.10 *If T denotes a random variable with distribution given by*

$$\mathbf{P}[T > x] = \int_0^\infty \frac{e^{-y}\,dy}{1 + 2e^{2x}y}$$

and $D^* = \frac{1}{\rho}\left(\frac{1}{2}\log(L\rho) + T\right)$, then

$$d_{TV}(\mathcal{L}(D), \mathcal{L}(D^*)) = O\left((L\rho)^{-\frac{1}{5}}\log^2(L\rho)\right).$$

Proof: As $c \to 0$, $\int_0^\infty \frac{e^{-y}dy}{1+cy} \sim 1 - c$; as $c \to \infty$, we have

$$\int_0^\infty \frac{e^{-y}\,dy}{1 + cy} \leq \left(1 - e^{-\frac{1}{c}}\right) + \frac{1}{c}\log c + \frac{1}{c} \sim \frac{1}{c}\log c.$$

Then use the bound from Theorem 3.9 for $-\frac{1}{6}\log(L\rho) \leq x \leq \frac{1}{10}\log(L\rho)$, and the tail estimates of $\mathcal{L}(D^*)$ outside this range. $\square$

Note that, in $|x| \leq \frac{1}{4}\log(L\rho)$, the asymptotics of the NMW heuristic give

$$\mathbf{P}_{NMW}\left[D > \frac{1}{\rho}\left(\frac{1}{2}\log(L\rho) + x\right)\right] = \frac{1}{1 + e^{2x}}(1 + O(\{L\rho\}^{-1/2})),$$

agreeing with Theorem 3.9 only at the $\log(L\rho)$ order. Under the NMW heuristic, the asymptotic distribution of $\rho D - \frac{1}{2}\log(L\rho)$ is a logistic distribution with mean zero, whereas the true asymptotic distribution given in Corollary 3.10 has mean ≈ -0.058, and has a much wider spread: see Figure 1.

As in the NMW model, we could instead have taken the number of shortcuts to be fixed at $\frac{1}{2}L\rho$. Using a standard deviation argument, such a process could with high probability be bracketed by two processes with Poisson numbers of shortcuts, but with different shortcut rates

$$\rho_1 = \rho\{1 - \kappa L^{-1}\sqrt{L\rho}\log(L\rho)\} \quad \text{and} \quad \rho_2 = \rho\{1 + \kappa L^{-1}\sqrt{L\rho}\log(L\rho)\},$$

for κ big enough. Expanding $\log(L\rho_1)$ and $\log(L\rho_2)$ around $\log(L\rho)$, we see that the distributional approximation for the shortest path length would remain the same, for a slightly different region and with different error bounds.

4 r-dimensions

Our method of proof can be adapted to many other models. Here, we consider only the generalization of the previous continuous circle model to higher dimensions, taking $\mathrm{Po}\,(L\rho/2)$ shortcuts between random pairs of points in a finite, homogeneous space C in r dimensions, such as a sphere or a torus. We construct the shortcuts by a “growth and merge” process as before, but now with intervals replaced by local neighbourhoods of the form tK for some given convex set K in r dimensions, after growth time t. The basic Poisson approximation of Theorem 3.1 can be applied as before, as can Theorem 3.4; we then need to find an appropriate τ_x, and make the necessary computations for the associated pure growth process. In particular, we shall need to be able to approximate the sums

$$\sum_{i=1}^{m}\sum_{j=1}^{n} p(s_i, u_j), \tag{4.1}$$

where $p(s_i, u_j)$ is the probability that two independently and randomly distributed sets, one an s_iK and the other a u_jK, intersect one another. We assume this probability to be of the form

$$p(s_i, u_j) = L^{-1}\sum_{l=0}^{r} \alpha_l s_i^l u_j^{(r-l)} \qquad \text{for constants } \alpha_l = \alpha_{r-l}, \tag{4.2}$$

where L is the area of C. In one dimension, as in the previous section, $\alpha_0 = \alpha_1 = 2$; for a torus in two dimensions with K a unit square, $\alpha_0 = \alpha_1/2 = \alpha_2 = 4$. For a sphere in two dimensions, it is almost the case, neglecting curvature, that $\alpha_0 = \alpha_1/2 = \alpha_2 = \pi$, and the error in using this approximation is negligible for large L, to our order of approximation.

For the pure growth process with independently and uniformly positioned neighbourhoods, define its neighbourhood size process by the purely atomic measure $(\xi_t,\, t \geq 0)$ on $\mathbf{R}_+$:

$$\xi_t(A) = \#\{\text{neighbourhoods with radii having lengths in } A\}.$$

Then the quantities

$$M_l(t) := \int_{\mathbf{R}_+} x^l \xi_t(dx), \quad l \geq 0,$$

are basic to our analysis: $M_0(t)$ is just the number of neighbourhoods in the pure growth process at time t, corresponding to $M(t)$ in the circle model, and

$$M_l(t) = \sum_{j=1}^{M_0(t)} s_j^l$$

is the sum of the l'th powers of the 'radii' of the neighbourhoods. In particular, in view of (4.1) and (4.2), the analogue of $\lambda_{(m,n,s,u)}$ of Corollary 3.2 is

$$L^{-1} \sum_{l=0}^{r} \alpha_l \sum_{i=1}^{m} s_i^l \sum_{j=1}^{n} u_j^{r-l},$$

easily expressible for two pure growth processes $\mathbf{M}$ and $\mathbf{N}$ at time t as

$$L^{-1} \sum_{l=0}^{r} \alpha_l M_l(t) N_{r-l}(t). \tag{4.3}$$

The quantities $M_l(t)$, $0 \leq l \leq r$, satisfy the following evolution equations:

$$\begin{aligned} \frac{d}{dt} M_i(t) &= i M_{i-1}(t), \quad i \geq 2; \\ \frac{d}{dt} M_1(t) &= M_0(t) \quad \text{for a.e. } t; \\ M_0(t) - \rho r v(K) \int_0^t M_{r-1}(u) du &= X_0(t), \end{aligned} \tag{4.4}$$

where $v(K)$ is the area of K, and where X_0 is a martingale, with $X_0(0) = 1$ and with (centred) Poisson innovations having rate $\rho r v(K) M_{r-1}(t)$ at time t. Properties of their solution are given in the following theorem.

Theorem 4.1 *Let $\mathbf{M}(t)$ denote the $(r+1)$-vector $(M_0(t), \ldots, M_r(t))$. Then, as $t \to \infty$,*

$$\mathbf{E}\mathbf{M}(t) \sim r^{-1} e^{\lambda_0(t)} \mathbf{e}^{(0)}; \quad \mathbf{E} M_0^2(t) = O(e^{2\lambda_0 t}),$$

where

$$\lambda_0 = \lambda_0(\rho) := (r! \rho v(K))^{\frac{1}{r}}; \quad \mathbf{e}^{(0)} = (1, \lambda_0^{-1}, 2\lambda_0^{-2}, \ldots, (r-1)! \lambda_0^{-(r-1)}, r! \lambda_0^{-r})^T.$$

Furthermore,

$$\mathbf{W}(t) \ := \ \mathbf{M}(t) e^{-\lambda_0 t} \to \mathbf{e}^{(0)} W \ \text{a.s.},$$

where

$$W := \frac{1}{r}\left\{1 + \int_0^\infty e^{-\lambda_0 y}\, dX_0(y)\right\}$$

satisfies $\mathbf{P}[W > 0] = 1$, *and*

$$(j!)^{-1}\lambda_0^j \mathbf{E}|W_j(t) - W\mathbf{e}_j^{(0)}| \leq c_r(1 + t^{1/2}1_{\{r=6\}})e^{-\lambda_0(1-\omega_*)t}, \quad 0 \leq j \leq r,$$

for some constant c_r *not depending on* ρ, *where*

$$\omega_* = \max\{0, \cos(2\pi/r)\}.$$

Remark. The fact that $W_0(t) \to W$ a.s. implies that

$$M_0(t)^{-1}\xi_t\{[0,b)\} = 1 - W_0(t-b)e^{-\lambda_0 b}/W_0(t) \to 1 - e^{-\lambda_0 b}$$

for all $b > 0$, so that the form of $\mathbf{e}^{(0)}$ is not surprising.

PROOF: Solving (4.4) for the vector $\mathbf{M}(t)$, we find that

$$M(t) = Ae^{At}\int_0^t e^{-Ay}X_0(y)\varepsilon^{\mathbf{0}}\,dy + X_0(t)\varepsilon^{\mathbf{0}}, \tag{4.5}$$

where $\varepsilon^{\mathbf{0}}$ is the coordinate vector $(1,0,0,\ldots,0)^T$, and

$$A := \begin{pmatrix} 0 & & & & 0 & c & 0 \\ 1 & 0 & & & & & \\ 0 & 2 & \cdot & & & & \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & & \cdot & \cdot \\ 0 & & & & 0 & r & 0 \end{pmatrix}$$

has eigenvalues $\lambda_r = 0$ and $\lambda_0, \ldots, \lambda_{r-1}$ satisfying the equation $\lambda^r = (r-1)!c$: here, $c = \rho r v(K)$. Thus $\lambda_0 = \{c(r-1)!\}^{\frac{1}{r}}$ is real and positive, and $\lambda_l = \lambda_0\omega_l$, where $\omega_l = e^{2\pi i l/r}$ are the complex r'th roots of unity.

The eigenvectors $\mathbf{e}^{(l)}$ of A are also easy to determine. The r'th eigenvector $\mathbf{e}^{(r)}$ is the coordinate vector $\varepsilon^{\mathbf{r}} = (0,\ldots,0,1)^T$, and all others have components

$(e_0, \ldots, e_r)$ satisfying the equations

$$\begin{aligned} ce_{r-1} &= \lambda e_0; \quad e_0 = \lambda e_1; \quad 2e_1 = \lambda e_2; \quad \ldots; \\ (r-1)e_{r-2} &= \lambda e_{r-1}; \quad re_{r-1} = \lambda e_r, \end{aligned}$$

for $\lambda = \lambda_0, \ldots, \lambda_{r-1}$. These in turn give

$$\begin{aligned} e_1 &= \lambda^{-1} e_0; \quad e_2 = 2\lambda^{-2} e_0; \quad \ldots; \\ e_{r-1} &= (r-1)!\lambda^{-(r-1)} e_0 = \lambda c^{-1} e_0; \quad e_r = r!\lambda^{-r} e_0. \end{aligned}$$

Thus it follows that, for $0 \leq l \leq r-1$,

$$\mathbf{e}^{(l)} = \left(1, \lambda_0^{-1}\omega_l, 2\lambda_0^{-2}\omega_l^2, \ldots, (r-1)!\lambda_0^{-(r-1)}\omega_l^{r-1}, c^{-1}r\right)^T. \tag{4.6}$$

The eigendecomposition can now be used to determine $\mathbf{M}(t)$ more explicitly. Writing $\varepsilon^{\mathbf{0}}$ in terms of the $\mathbf{e}^{(l)}$, $0 \leq l \leq r$, we find that

$$\varepsilon^{\mathbf{0}} = \sum_{l=0}^{r} \mu_l \mathbf{e}^{(l)} \quad \text{with} \quad \mu_0 = \cdots = \mu_{r-1} = r^{-1} \text{ and } \mu_r = -r/c.$$

Using this to evaluate (4.5), we obtain

$$\mathbf{M}(t) \quad = \quad \int_0^t \{(X_0(y) - 1) + 1\} \sum_{l=0}^{r-1} \mathbf{e}^{(l)} \mu_l \lambda_l e^{\lambda_l (t-y)} dy + X_0(t)\varepsilon^{\mathbf{0}}. \tag{4.7}$$

Turning to the moments of $\mathbf{M}$, we immediately have

$$\begin{aligned} \mathbf{EM}(t) &= \int_0^t \sum_{l=0}^{r-1} \mathbf{e}^{(l)} \mu_l \lambda_l e^{\lambda_l (t-y)} dy + \varepsilon^{\mathbf{0}} \\ &= \sum_{l=0}^{r-1} \frac{1}{r} \left(e^{\lambda_0 t \omega_l} - 1\right) \mathbf{e}^{(l)} + \varepsilon^{\mathbf{0}} \sim \frac{1}{r} e^{\lambda_0 t} \mathbf{e}^{(0)}. \end{aligned}$$

To estimate second moments, observe that, given $\mathcal{F}_t$, $dX_0(t)$ is a centred Poisson innovation, so that

$$\mathbf{E}\left(d^2 X_0(t)\right) = c\mathbf{E}(M_{r-1}(t))dt \sim cr^{-1} e^{\lambda_0 t} \mathbf{e}_{r-1}^{(0)}\, dt = r^{-1}\lambda_0 e^{\lambda_0 t}\, dt. \tag{4.8}$$

But now, in (4.7), we can write

$$M^{(il)}(t) := \int_0^t (X_0(y) - 1)\lambda_l e^{\lambda_l (t-y)}\, dy = \int_0^t \{e^{\lambda_l (t-v)} - 1\}\, dX_0(v),$$

since X_0 is a.s. of bounded variation on finite intervals. Hence, from (4.8),

$$\mathbf{E}|M^{(il)}(t)|^2 \leq 2r^{-1}\lambda_0 \int_0^t e^{\lambda_0 v}\{1 + e^{2\Re\lambda_l(t-v)}\}\, dv. \tag{4.9}$$

For $1 \leq l \leq r-1$, we can bound the integral in (4.9), uniformly in l and ρ, in terms of ω_* and $\lambda_0(\rho)t$. If $1 \leq r \leq 4$, so that $\omega_* \leq 0$, it is immediate that $\mathbf{E}|M^{(il)}(t)|^2 \leq 4r^{-1}e^{\lambda_0 t}$. If $r = 5$, so that $0 < \omega_* < 1/2$, the bound has an extra factor of $1/(1-2\omega_*)$; if $r = 6$, then $\omega_* = 1/2$, and $\mathbf{E}|M^{(il)}(t)|^2 \leq 4r^{-1}\lambda_0 t e^{\lambda_0 t}$; for $r \geq 7$, $\mathbf{E}|M^{(il)}(t)|^2 \leq 4r^{-1}(2\omega_* - 1)^{-1}e^{2\lambda_0\omega_* t}$. In all cases, for all $1 \leq l \leq r-1$, we have $\mathbf{E}|M^{(il)}(t)|^2 = o(e^{2\lambda_0 t})$, and since also

$$\mathbf{E}|M^{(i0)}(t)|^2 \leq r^{-1}\lambda_0 \int_0^t e^{2\lambda_0 t - \lambda_0 v}\, dv \leq r^{-1}e^{2\lambda_0 t},$$

it therefore follows easily that

$$\mathbf{E}M_0^2(t) = O(e^{2\lambda_0 t}) \quad \text{as } t \to \infty, \tag{4.10}$$

with the constant implied in the order estimate uniform for all ρ.

The convergence of $\mathbf{W}(t)$ can now be proved by second moment arguments. It follows directly from (4.8) that $\operatorname{Var} X_0(t) = r^{-1}(e^{\lambda_0 t} - 1)$, and thus, using Kolmogorov's inequality on the martingale X_0, we can easily show that, for any $\delta > 0$,

$$\sup_{t>0} |X_0(t) - 1| e^{-(\lambda_0+\delta)t/2} < \infty. \tag{4.11}$$

Now, from (4.7),

$$\begin{aligned} \mathbf{W}(t) &= e^{-\lambda_0 t}\mathbf{M}(t) \\ &= e^{-\lambda_0 t}\mathbf{E}\mathbf{M}(t) + e^{-\lambda_0 t}(X_0(t) - 1)\varepsilon^{\mathbf{0}} \\ &\quad + \int_0^t (X_0(y) - 1)\sum_{l=0}^{r-1} \mathbf{e}^{(l)}\mu_l\lambda_l e^{\lambda_l(t-y)} dy e^{-\lambda_0 t}, \end{aligned} \tag{4.12}$$

and we consider the various terms in (4.12), using (4.11). First, it is immediate that

$$\lim_{t\to\infty} e^{-\lambda_0 t}(X_0(t) - 1) = 0 \text{ a.s.}$$

Then, distinguishing the cases $\Re(\lambda_l) \leq 0$, $0 < \Re(\lambda_l) \leq \lambda_0/2$ and $\lambda_0/2 < \Re(\lambda_l) < \lambda_0$, it again follows from (4.11) that, for $1 \leq l \leq r-1$,

$$\lim_{t\to\infty} \int_0^t (X_0(y) - 1)\mu_l \lambda_l e^{\lambda_l(t-y)-\lambda_0 t}\, dy = 0 \text{ a.s.}$$

Finally, for $l = 0$, we have

$$\lim_{t\to\infty} \int_t^\infty (X_0(y) - 1)\lambda_0 e^{-\lambda_0 y}\, dy = 0 \text{ a.s.}$$

Hence, letting $t \to \infty$ in (4.12), it follows that the limit

$$\lim_{t\to\infty} \mathbf{W}(t) \quad = \quad \mathbf{e}^{(0)}\frac{1}{r}\left\{1 + \int_0^\infty \lambda_0(X_0(y) - 1)e^{-\lambda_0 y} dy\right\} = W\mathbf{e}^{(0)}$$

exists a.s. To show that $\mathbf{P}[W > 0] = 1$, note that

$$W = \sum_{l\geq 1} e^{-\lambda_0 U_l} W^{(l)},$$

where $(U_l,\, l \geq 1)$ are the times of the births of new neighbourhoods as children of the original neighbourhood around P, and $(W^{(l)},\, l \geq 1)$ are independent copies of W. This immediately implies that $\mathbf{P}[W = 0] \in \{0, 1\}$, and $\mathbf{P}[W = 0] = 1$ is excluded by $\mathbf{E}W = 1/r$.

Note also that, again because X_0 is locally of bounded variation,

$$\begin{aligned}
&\mathbf{W}(\infty) - \mathbf{W}(t) \\
&= \mu_0\mathbf{e}^{(0)}e^{-\lambda_0 t} - \sum_{l=1}^{r-1} \mathbf{e}^{(l)}\mu_l e^{-\lambda_0 t}\left(e^{\lambda_l t} - 1\right) \\
&\quad -\varepsilon^{\mathbf{0}} e^{-\lambda_0 t} + \mu_0\mathbf{e}^{(0)} \int_t^\infty e^{-\lambda_0 y} dX_0(y) \\
&\quad\quad - \sum_{l=1}^{r-1} \mathbf{e}^{(l)}\mu_l\lambda_l e^{-(\lambda_0-\lambda_l)t} \int_0^t \frac{e^{-\lambda_l y} - e^{-\lambda_l t}}{\lambda_l} dX_0(y) \\
&= \mu_0\mathbf{e}^{(0)}e^{-\lambda_0 t} - \sum_{l=1}^{r-1} \mathbf{e}^{(l)}\mu_l e^{-\lambda_0 t}\left(e^{\lambda_l t} - 1\right) \\
&\quad -\varepsilon^{\mathbf{0}} e^{-\lambda_0 t} + \mu_0\mathbf{e}^{(0)} \int_t^\infty e^{-\lambda_0 y} dX_0(y) \\
&\quad\quad - \sum_{l=1}^{r-1} \mathbf{e}^{(l)}\mu_l \int_0^t dX_0(y)\left\{e^{-\lambda_0 t + \lambda_l(t-y)} - e^{-\lambda_0 t}\right\}.
\end{aligned}$$

Thus

$$(j!)^{-1}\lambda_0^j|\mathbf{E}(W_j(\infty)-W_j(t))| \le c_{r1}e^{-\lambda_0(1-\omega_*)t} \tag{4.13}$$

and, from (4.8),

$$\mathrm{Var}\,\{(j!)^{-1}\lambda_0^j(W_j(\infty)-W_j(t))\} \le c_{r2}e^{-2\lambda_0(1-\omega_*)t}\{1+(\lambda_0 t)1_{\{r=6\}}\}, \tag{4.14}$$

for some constants c_{r1} and c_{r2} and for all $0 \le j \le r$, completing the proof. []

Now choose

$$\tau_x := \frac{1}{\lambda_0(\rho)}\left\{\frac{1}{2}\log(L\rho)+x\right\}, \tag{4.15}$$

for any x such that $|x| \le \frac{1}{4}\log(L\rho)$, agreeing with the previous choice in the case $r=1$. Let $\widehat{V}_x$ as before be the number of pairs of overlapping neighbourhoods, one from each of two independent pure growth processes $\mathbf{M}(\tau_x)$ and $\mathbf{N}(\tau_x)$. Then, from (4.3) and Theorem 3.1, it follows easily that

$$\begin{aligned}\left|\mathbf{P}[\widehat{V}_x=0]-\mathbf{E}\exp\left\{-L^{-1}\sum_{l=0}^{r}\alpha_l M_l(\tau_x)N_{r-l}(\tau_x)\right\}\right|&\\ =O((L\rho)^{-1/2}\log^r(L\rho)e^x).&\end{aligned} \tag{4.16}$$

since the probability of a given pair of neighbourhoods overlapping cannot exceed

$$L^{-1}\kappa v(K)\tau_x^r = O((L\rho)^{-1}\log^r(L\rho)),$$

for some constant κ depending on the shape of K, and because $\mathbf{E}M_0(\tau_x) \sim r^{-1}(L\rho)^{1/2}e^x$ from Theorem 4.1. Then, as for Corollary 3.5, it follows from Theorem 3.4 (still with $c=1$ in (3.1)) and Theorem 4.1 that

$$\mathbf{P}[\widehat{V}_x \ne V_x] = O((L\rho)^{-1/2}\log^{2r}(L\rho)e^{3x}), \tag{4.17}$$

where V_x is the number of overlaps between the two associated growth and merge processes. These observations lead to the following theorem.

Theorem 4.2 *Let D denote the distance between two randomly chosen points of C on the graph with a $\mathrm{Po}\,(L\rho/2)$–distributed random number of shortcuts. Then*

$$\begin{aligned}\mathbf{P}&\left[D > \frac{2}{\lambda_0(\rho)}\left\{\frac{1}{2}\log(L\rho)+x\right\}\right]\\ &= \mathbf{E}\left(\exp\left\{-\frac{e^{2x}WW'}{v(K)}\sum_{l=0}^{r}\frac{\alpha_l}{\binom{r}{l}}\right\}\right)\\ &\qquad +O\left\{(L\rho)^{-(1-\omega_*)/2}\log^{2r+1}(L\rho)(e^{3x}+e^x)\right\},\end{aligned}$$

uniformly in $|x| \le \frac{1}{4}\log(L\rho)$, where W and W' are independent copies of the limiting random variable of Theorem 4.1.

Proof: In view of (4.16) and (4.17), it is enough to show that

$$\begin{aligned}&\mathbf{E}\exp\left\{-L^{-1}\sum_{l=0}^{r}\alpha_l M_l(\tau_x)N_{r-l}(\tau_x)\right\}\\ &\quad = \mathbf{E}\left(\exp\left\{-\frac{e^{2x}WW'}{v(K)}\sum_{l=0}^{r}\frac{\alpha_l}{\binom{r}{l}}\right\}\right)+O(\tau_x\exp\{2x-\lambda_0(1-\omega_*)\tau_x\}).\end{aligned}$$

However, direct calculation shows that

$$L^{-1}\sum_{l=0}^{r}\alpha_l M_l(\tau_x)N_{r-l}(\tau_x) = \frac{e^{2x}}{r!v(K)}\sum_{l=0}^{r}\alpha_l\{\lambda_0^l W_l(\tau_x)\}\{\lambda_0^{r-l}W'_{r-l}(\tau_x)\},$$

where, by Theorem 4.1,

$$\mathbf{E}|(j!)^{-1}\lambda_0^j W_j(t) - W| \le c_r(1+(\lambda_0 t)^{1/2}1_{\{r=6\}})e^{-\lambda_0(1-\omega_*)t}, \quad 0\le j\le r.$$

Then the inequality $|e^{-a}-e^{-b}| \le |a-b|$ for all $a,b \ge 0$ completes the proof. []

Note that in the r–dimensional model, the distances D have $\lambda_0(\rho)$ in the denominator, of order $O(\rho^{1/r})$, in place of 2ρ in the case $r=1$. This scaling can be understood as follows. If there were no shortcuts, the average distance between pairs of points would be of order $L^{1/r}$. With about $L\rho/2$ shortcuts, this distance is reduced by a factor of order $(L\rho)^{-1/r}\log(L\rho)$. Thus, in a higher dimensional space, the reduction in distance as a result of introducing shortcuts is correspondingly smaller.

Acknowledgement. The authors would like to thank A.G. Pakes for sharing his expertise on the limiting variable W. Moreover the authors are grateful to the anonymous referees for valuable comments.

References

[1] Ball, F., Mollison, D. and Scalia-Tomba, G. (1997). Epidemics with two levels of mixing. *Ann. Appl. Probab.* **7**, 46 – 89.

[2] Barbour, A.D., Holst, L., and Janson, S. (1992). *Poisson Approximation.* Oxford Science Publications.

[3] Barrat and Weigt (2000). On the properties of small-world network models. *Europhysics Journal* **B**, **13**, 547 – 560.

[4] Bollobas, B. (1985). *Random Graphs.* Academic Press, London.

[5] Moore, C. and Newman, M.E.J. (1999). Epidemics and percolation in small-world networks. Santa Fe Institute working paper 00-01-002. Also `http://xxx.lanl.gov/archive/cond-mat/9911492`.

[6] Newman, M.E.J. and Watts, D.J. (1999). Scaling and percolation in the small-world network model. *Phys. Rev. E*, **60**, 7332-7344.

[7] Newman, M.E.J., Moore, C. and Watts, D.J. (1999). Mean-field solution of the small-world network model. Santa Fe Institute working paper 99-09-066. Also `http://xxx.lanl.gov/archive/cond-mat/9909165`.

[8] Ross, S.M. (1996). *Stochastic Processes.* Wiley, New York etc, 2^{nd} edition.

[9] Watts, D.J. (1999). *Small Worlds.* Princeton University Press.

[10] Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of "small-world" networks. *Nature* **393**, 440 – 442.

Figure 1.

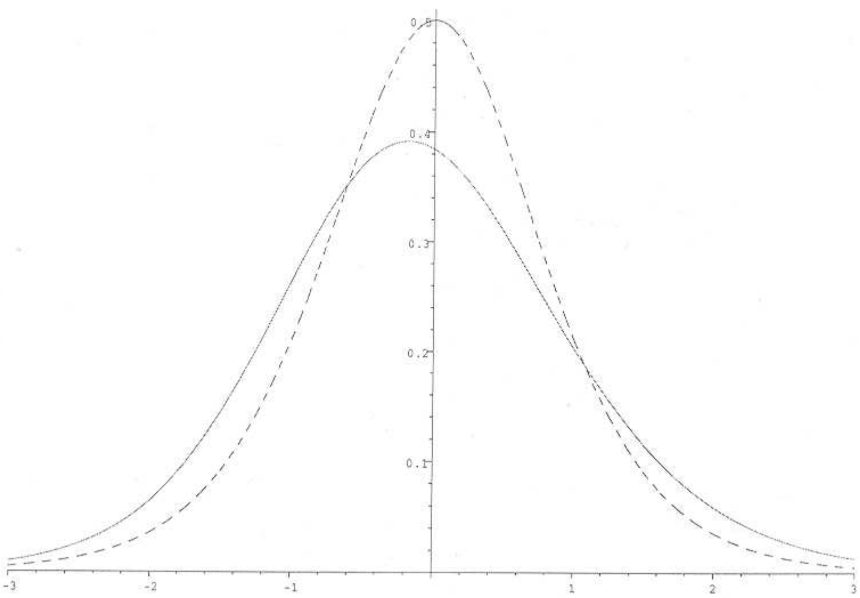

Figure 1: The asymptotic distribution function of $\rho D - \frac{1}{2}\log(L\rho)$ (solid line) and that predicted by the NMW heuristic (dashed line)