*J. Appl. Prob.* **50,** 861–870 (2013)
*Printed in England*


# JOINT DISTRIBUTION OF DISTANCES IN LARGE RANDOM REGULAR NETWORKS

JUSTIN SALEZ,* *University of California, Berkeley*

**Abstract**

We study the array of point-to-point distances in random regular graphs equipped with exponential edge lengths. We consider the regime where the degree is kept fixed while the number of vertices tends to $\infty$. The marginal distribution of an individual entry is now well understood, thanks to the work of Bhamidi, van der Hofstad and Hooghiemstra (2010). The purpose of this note is to show that the whole array, suitably recentered, converges in the weak sense to an explicit infinite random array. Our proof consists in analyzing the invasion of the network by several mutually exclusive flows emanating from different sources and propagating simultaneously along the edges.

*Keywords:* Random regular graph; distance matrix; first passage percolation; multitype Richardson process; configuration model; branching process approximation

2010 Mathematics Subject Classification: Primary 60C05
Secondary 05C80; 90B15

## 1. Introduction

Assigning independent exponential random lengths $(\ell_e)_{e\in E}$ to the edges of a graph $G = (V, E)$ induces a natural random metric on the vertex set $V$, namely,

$$\mathcal{D}_{i,j} = \min\left\{\sum_{e\in\pi} \ell_e : \pi \text{ is a path from } i \text{ to } j\right\}.$$

A now classical problem is that of understanding the asymptotic shape of a ball $B_i(t) = \{j \in V : \mathcal{D}_{i,j} \le t\}$ as $t \to \infty$, when $G$ is an infinite transitive graph (see, e.g. [12] and [25]). More recently, growing attention has been devoted to the way in which random edge lengths affect the inherent geometry of large finite graphs, and in particular their typical distance, flooding time, and diameter [22], [26], [27], [2], [16], [6], [8], [7], [9], [3]. The present note is concerned with one specific aspect of the above body of works, namely the remarkable second-order behaviour established by Bhamidi *et al.* [8] for the distance between two typical points in the configuration model. Although their result holds for an arbitrary prescribed degree distribution with finite variance, we will for simplicity restrict ourselves to the regular case.

Throughout the paper, $d \ge 3$ is a fixed integer. For $n \ge 1$, such that $dn$ is even, we consider the random $d$-regular multigraph on $\{1, \dots, n\}$ obtained by attaching $d$ half-edges to every vertex, and pairing these $dn$ half-edges uniformly at random to create edges. The edges are then assigned independent unit-rate exponential random lengths. As $n \to \infty$, it was shown in [8] that the distance between two fixed nodes—say 1 and 2—satisfies

$$\mathcal{D}_{1,2} - \frac{\log n}{d-2} \xrightarrow{\text{D}} W \quad \text{as } n \to \infty,$$

Received 17 April 2012; revision received 12 December 2012.
* Current address: Université Paris Diderot, LPMA, Site Chevaleret, case 7012, 75205 Paris Cedex 13, France. Email address: justin.salez@univ-paris-diderot.fr

where $W$ is a nondegenerate real-valued random variable whose distribution has been explicitly described. Specifically,

$$W \stackrel{\mathrm{D}}{=} X + X' + Y, \tag{1.1}$$

where $X$, $X'$, and $Y$ are independent with the following explicit densities over $\mathbb{R}$:

$$f_X(u) = f_{X'}(u) = \frac{d-2}{\Gamma(d/(d-2))} \mathrm{e}^{-du} \mathrm{e}^{-\mathrm{e}^{-(d-2)u}}, \tag{1.2}$$

$$f_Y(u) = 2d \mathrm{e}^{(d-2)u} \mathrm{e}^{-\mathrm{e}^{(d-2)u}/2d(d-2)}. \tag{1.3}$$

However, two-point marginals do not capture correlations between distances, and it is natural to look for a slightly more refined description of the relative geometry of the network. This is precisely the aim of the present note.

### 1.1. Results

In light of the results obtained on the complete graph [2], [6], a reasonable hope is that the whole array of re-centered distances converges in distribution to some infinite random array $\mathbb{W} = \{W_{i,j} \colon 1 \le i < j < \infty\}$. Weak convergence of arrays is understood with respect to the usual product topology, i.e. for each fixed $k \ge 1$,

$$\left\{ \mathcal{D}_{i,j} - \frac{\log n}{d-2} \colon 1 \le i < j \le k \right\} \xrightarrow{\mathrm{D}} \{W_{i,j} \colon 1 \le i < j \le k\} \quad \text{as } n \to \infty.$$

Note that such a limiting random array, if it exists, must be *partially exchangeable* in the sense described in Equation (2.2) of the survey paper [1], which discusses a much broader picture of representing complex random structures via induced substructures on randomly chosen points. In addition, the marginal distribution of each entry must coincide with (1.1). A simple candidate would be

$$W_{i,j} = X_i + X_j + Y_{i,j}, \tag{1.4}$$

where $\{X_i \colon i \ge 1\} \cup \{Y_{i,j} \colon 1 \le i < j < \infty\}$ is a collection of mutually independent random variables, with each $X_i$ admitting the density (1.2) on $\mathbb{R}$, and each $Y_{i,j}$ admitting the density (1.3) on $\mathbb{R}$. Decomposition (1.4) admits the following intuitive interpretation: the random variable $X_i$ may be thought of as a 'local cost' to escape (or enter) the neighborhood of node $i$, regardless of the destination (or source), and the random variable $Y_{i,j}$ as a 'transfer cost' to switch from the peripheral neighborhood of $i$ to that of $j$, as schematized in Figure 1.

Our result is that this remarkably simple metric structure is indeed the correct one.

**Theorem 1.1.** *For the $d$-regular multigraph generated by the configuration model on $\{1, \dots, n\}$ with independent rate-1 exponential edge lengths,*

$$\left\{ \mathcal{D}_{i,j} - \frac{\log n}{d-2} \colon 1 \le i < j \le n \right\} \xrightarrow{\mathrm{D}} \mathbb{W} = \{W_{i,j} \colon 1 \le i < j < \infty\} \quad \textit{as } n \to \infty,$$

*where $\mathbb{W}$ is the partially exchangeable random array described above.*

The configuration model was originally introduced by Bollobás [11] as a mechanism to simplify the analysis of the uniform $d$-regular simple graph on $\{1, \dots, n\}$: with probability bounded away from 0 as $n \to \infty$, the $d$-regular multigraph produced by the configuration model is simple, and conditioning on that event, it is uniformly distributed among all simple $d$-regular graphs on $\{1, \dots, n\}$. As a result, any event that occurs with high probability under

https://doi.org/10.1239/jap/1378401241 Published online by Cambridge University Press

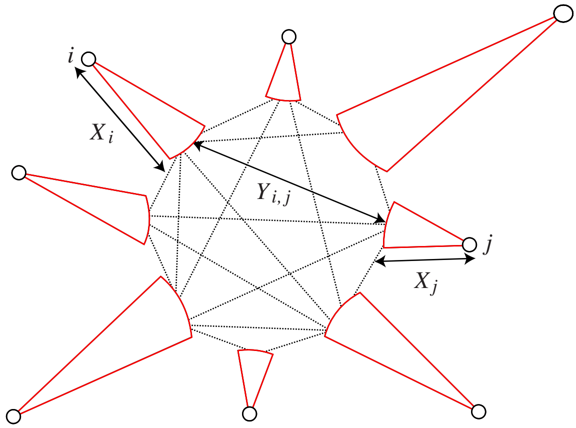


FIGURE 1: The typical relative geometry of a large random network.

the configuration model also does so under the uniform model (see [23] for a generalization to graphs with an arbitrary prescribed degree sequence). However, this does not directly carry over to weak convergence results such as Theorem 1.1, for which one needs to check that the random object under study is asymptotically independent from the presence of loops or multiple edges (see [24, Lemma 5] for a typical result along these lines). In Section 4 we provide a direct argument that exploits the independence inherent to the limiting array $\mathbb{W}$. Consequently, our result applies to the uniform $d$-regular simple graph as well.

**Theorem 1.2.** *For the uniform $d$-regular simple graph on $\{1, \ldots, n\}$ with independent exponential edge lengths,*

$$\left\{ \mathcal{D}_{i,j} - \frac{\log n}{d-2} : 1 \le i < j \le n \right\} \xrightarrow{\text{D}} \mathbb{W} = \{W_{i,j} : 1 \le i < j < \infty\} \quad \textit{as } n \to \infty,$$

*where $\mathbb{W}$ is the partially exchangeable random array described above.*

### 1.2. Discussion and possible extensions

As already mentioned, the two-point analysis conducted in [8] is by no means restricted to the regular case, and there is no doubt in the author's mind that the multipoint refinement established in the present paper should extend *mutatis mutandis* to arbitrary degree distributions with finite variance. However, several technical hurdles arise in the nonregular case, and it would be pleasant to find a clever way of dealing with them without substantially increasing the length of the proof.

Also, the assumption that the edge lengths are exponentially distributed may seem somewhat restrictive, especially from the point of view of modelling real-world networks. In a recent preprint, Bhamidi *et al.* [10] managed to generalize the results in [8] to arbitrarily distributed edge lengths. It would certainly be interesting to establish the joint convergence of the whole array of point-to-point distances in such a general setting.

### 1.3. Outline of the proof

Our proof consists in analyzing the invasion of the network by $k$ mutually exclusive flows emanating from different sources and propagating simultaneously at unit speed along the edges. Since the edge lengths are independent exponential random variables, this is exactly the multitype Richardson process introduced by Häggström and Pemantle [18] as a model for competing spatial growth. A considerable amount of work has been devoted to understanding the long-time behaviour of this process on lattices and other fixed infinite graphs [19], [20], [14], [13], [15], [17], [21]. More recently, a version of the Richardson process was analyzed on random regular graphs [4]. We note that our concern here is quite different, since we are interested in the second-order behaviour of the times at which the various species collide. The process is defined in Section 2, and analyzed on the configuration model in Section 3. The results are finally transferred to uniform simple regular graphs in Section 4.

## 2. Competing first passage percolation

We first recall the definition of the multitype Richardson process on an arbitrary multigraph $G$. We then specialize to the situation where $G$ is generated by the configuration model, in which case the competition can be coupled with an infinite branching process that will greatly simplify our asymptotic analysis.

### 2.1. Competition on a general multigraph

Let $G = (V, E)$ be a finite multigraph with $k$ distinguished vertices $o_1, \dots, o_k$, called the roots. Consider a continuous-time Markov process that colours vertices with integers $\{1, \dots, k\}$ as follows. Initially, all vertices are uncoloured except the roots, whose colours are $1, \dots, k$, respectively. Every edge is equipped with two independent unit-rate exponential clocks, one for each orientation. A clock begins to tick as soon as its start vertex $x$ gets coloured. When it rings, the action depends on the state of the end vertex $y$:

- if $y$ is uncoloured then $y$ gets the same colour as $x$ (invasion),
- if $y$ is already coloured then nothing changes (collision).

Alternatively, one may view this process as describing the invasion of the network by $k$ mutually exclusive flows emanating from the roots and propagating simultaneously at unit speed along the edges. For this interpretation to be valid, each edge must be considered as having a random length equal to the sum of the values shown by its two clocks when the first of them rings. This choice precisely results in the edge lengths being independent and exponentially distributed with mean 1, thanks to the memoryless property of the exponential distribution.

A consequence of this coupling is that the (random, possibly infinite) time $\tau_x$ at which a given vertex $x$ gets coloured is precisely its distance to the roots (that is, to the closest root), and that the colour it receives is that of the closest root. Similarly, if one lets $\tau_{i\to j}$ denote the time at which the first collision from a node with colour $i$ to a node with colour $j$ occurs, then $2(\tau_{i\to j} \wedge \tau_{j\to i})$ is the minimum length of a path from $o_i$ to $o_j$ passing only through nodes that are closer to the set $\{o_i, o_j\}$ than to the set $\{o_1, \dots, o_k\} \setminus \{o_i, o_j\}$. In particular, almost surely under this coupling,

$$\mathcal{D}_{o_i,o_j} \leq 2(\tau_{i\to j} \wedge \tau_{j\to i}) \tag{2.1}$$

for all $1 \leq i < j \leq k$, with equality when $k = 2$.

It should be here noted that the distance $\mathcal{D}_{o_i,o_j}$ does not depend upon the number and choice of the other roots, whereas the upper bound $2(\tau_{i\to j} \wedge \tau_{j\to i})$ implicitly does. In particular, the distribution of the random array $\{2(\tau_{i\to j} \wedge \tau_{j\to i})\colon 1 \leq i < j \leq k\}$ is not consistent as one increases the number $k$ of roots. However, we will see that it becomes asymptotically consistent when $G$ is a large multigraph generated by the configuration model. More precisely, the asymptotic joint distribution of all collision times will be determined in Section 3, implying in particular the following result.

**Proposition 2.1.** *For the competing first passage percolation from k fixed roots on a d-regular n-vertex multigraph generated by the configuration model,*

$$\left\{2(\tau_{i\to j}\wedge\tau_{j\to i})-\frac{\log n}{d-2}\colon 1\le i<j\le k\right\}\xrightarrow{\mathrm{D}}\{W_{i,j}\colon 1\le i<j\le k\}\quad \text{as } n\to\infty,$$

*where $\mathbb{W}$ is the random infinite array defined by (1.4).*

*From Proposition 2.1 to Theorem 1.1.* Combining Proposition 2.1 with inequality (2.1), we already see that $\mathbb{W}$ is a limiting upper bound (in the weak sense) for the random array $\mathcal{D}-\log n/(d-2)$. But, since the equality holds in (2.1) when $k=2$, then the single marginals of $\mathcal{D}-\log n/(d-2)$ converge to those of $\mathbb{W}$. These two facts together imply Theorem 1.1, and we can now focus on the proof of Proposition 2.1. The latter relies on a standard property of the configuration model, namely that the pairings can be revealed progressively, as we explore our way through the network.

### 2.2. Competition on the configuration model

An elementary and well-known (yet crucial) observation about the configuration model is that the uniform perfect matching on the $dn$ half-edges can be constructed sequentially: at each step, one selects an arbitrary unpaired half-edge $e$ and pairs it with a uniformly chosen other unpaired half-edge $e'$. The distribution of the resulting multigraph does not depend on the rule for selecting $e$, and we may exploit this degree of freedom to couple the construction of the graph with the competition process. Specifically, we equip each half-edge with an independent exponential clock programmed to start ticking as soon as the associated vertex gets coloured. Initially, all vertices are uncoloured except the $k$ roots, whose colours are $1,\dots,k$. Whenever the clock of an unpaired half-edge $e$ rings, $e$ is paired to a uniformly chosen remaining half-edge $e'$. If the vertex to which $e'$ was attached is uncoloured then it gets coloured with the same colour as the endpoint of $e$ (invasion); otherwise, it keeps its own colour (collision).

When $n$ is large, the local structure of the random multigraph around the $k$ roots resembles that of a forest consisting of $k$ rooted $d$-regular trees $(T_1,o_1),\dots,(T_k,o_k)$ equipped with independent exponential edge lengths. This infinite object is of course only an idealized approximation, since every collision actually produces a transverse edge that violates the forest structure. In fact, the whole competition process can be directly described as a mechanism that randomly builds 'bridges' between the trees of the idealized forest $(T_1,o_1),\dots,(T_k,o_k)$. We will now define this bridging process more formally. It turns out that its analysis is more conveniently performed in a quenched way, i.e. by treating the edge lengths as deterministic.

2.2.1. *The bridging process.* Consider a forest consisting of $k$ infinite $d$-regular rooted trees $(T_1,o_1),\dots,(T_k,o_k)$, equipped with deterministic edge lengths. For each node $x$, we write $\tau_x$ for its distance to the roots and $[o,x]$ for the set of vertices along the unique path from the roots to $x$. Also, we set $V_t=\{x\colon \tau_x<t\}$ and write $\partial V_t$ for the outer boundary of $V_t$, consisting of all nodes that do not belong to $V_t$ but whose parent does. We shall make the following assumptions (which hold almost surely for random independent exponential edge lengths):

(A1) (unique distances) the $\tau_x$, $x\notin\{o_1,\dots,o_k\}$ are pairwise distinct,

(A2) (exponential growth) $|T_i\cap\partial V_t|\sim\omega_i\mathrm{e}^{(d-2)t}$ as $t\to\infty$ for some $\omega_1,\dots,\omega_k>0$,

(A3) (well balancedness) $\max_{x\in V_t}|[o,x]|=o(|V_t|)$ as $t\to\infty$.

https://doi.org/10.1239/jap/1378401241 Published online by Cambridge University Press

Given a parameter $n \geq k$, we add directed edges—called *bridges*—to this deterministic forest by applying the following random procedure, which dynamically maintains a certain set of *active* nodes and a decreasing counter $N \in \mathbb{N}$. Initially, the active nodes are the $d \times k$ nodes in the first generation of the forest, and $N = n - k$. Then, as long as the set of active nodes is not empty, the following action is repeated.

- Deactivate the closest active node $x$, and pick $y$ uniformly at random among the remaining active nodes plus $d \times N$ tokens.
- If $y$ is a token, activate the children of $x$ and decrease $N$ by 1 (invasion). Otherwise, deactivate $y$ and build a bridge from $x$ to $y$ (collision).

We write $\{x \to y\}$ for the event that there is a bridge from $x$ to $y$, and we call $\tau_x$ the height of the bridge. We also define $\{x \to T_j\} = \bigcup_{y \in T_j} \{x \to y\}$. For each $1 \leq i, j \leq k$, we are interested in the quantity $\tau_{i \to j} = \min\{\tau_x : x \in T_i,\ x \to T_j\}$, which is the height of the closest bridge from $T_i$ to $T_j$. More generally, we introduce a random point process $\mathcal{N}^{(n)}_{i \to j}$ on $\mathbb{R}_+$ that records the (suitably rescaled) heights of all bridges from $T_i$ to $T_j$, i.e.

$$\mathcal{N}^{(n)}_{i \to j} = \sum_{x \in T_i} \mathbf{1}_{\{x \to T_j\}} \delta_{\psi_n(\tau_x)},$$

where the scaling function $\psi_n$ is given by $\psi_n(t) = \mathrm{e}^{2(d-2)t}/n$. We will establish the following weak convergence.

**Theorem 2.1.** (Quenched distribution of the collision times.) *If the edge lengths satisfy assumptions (A1), (A2), and (A3), then the* $k^2$ *point processes* $\{\mathcal{N}^{(n)}_{i \to j} : 1 \leq i, j \leq k\}$ *converge jointly in distribution as* $n \to \infty$ *to* $k^2$ *independent Poisson point processes with respective rates* $\{\omega_i \omega_j / 2d(d-2) : 1 \leq i, j \leq k\}$ *on* $\mathbb{R}_+$*. In particular, jointly for all* $1 \leq i < j \leq k$,

$$2(\tau_{i \to j} \wedge \tau_{j \to i}) - \frac{\log n}{d-2} \xrightarrow{\mathrm{D}} \frac{1}{d-2}(\log \xi_{i,j} - \log \omega_i - \log \omega_j) \quad \text{as } n \to \infty,$$

*where the* $\{\xi_{i,j} : 1 \leq i < j \leq k\}$ *are independent exponential random variables with mean* $2d(d-2)$.

*From Theorem 2.1 to Proposition 2.1.* When the edge lengths are independent exponential random variables, the bridging process is clearly equivalent to the competition process described above: $N$ is the number of not yet coloured vertices, active nodes represent those half-edges whose endpoint is coloured, and building a bridge from an active node $x$ to an active node $y$ corresponds to pairing the associated half-edges. In particular, the exact structure of the finite multigraph can be obtained by simply deleting the endpoints of every bridge (as well as all descendants) and placing a transverse edge between their parents. Now, each tree in the forest realizes an independent continuous-time branching process for which it is classical (see, e.g. [5]) that the assumptions (A1), (A2), and (A3) hold almost surely, with the random variables $w_1, \ldots, w_k$ being independent gamma random variables with mean and variance $d/(d-2)$. Thus, Theorem 2.1 implies Proposition 2.1, with $Y_{i,j} = \log \xi_{i,j}/(d-2)$ and $X_i = (d-2)^{-1} \log \omega_i{}^{-1}$. We may now focus on the proof of Theorem 2.1.

## 3. Analysis of the bridging process: proof of Theorem 2.1

At any given time $t \geq 0$ during the bridging process, the set of active nodes is included in $\partial V_t$. More precisely, a node $x \in \partial V_t$ is active if and only if its unique path to the root does not contain the head or tail of a bridge. Therefore, the bridging process can be equivalently described as follows. For each nonroot node $x$, we first independently generate a random target $Y_x$ uniformly in $\partial V_{\tau_x} \setminus \{x\}$ and a random mark $U_x$ uniformly in $[0, 1]$. Then we build bridges by applying the following deterministic rule, sequentially for every nonroot node $x$ in the order of increasing distance to the roots.

1. Determine the current set $\mathcal{F}$ of nodes whose unique path to the roots contains the head or tail of a bridge.

2. Build a bridge from $x$ to $Y_x$ if neither $x$ nor $Y_x$ belongs to $\mathcal{F}$, and
$$U_x \le \frac{|\partial V_{\tau_x}| - 1}{dn + 2k - 2|V_{\tau_x} \setminus \mathcal{F}| - 1}.$$

Strictly speaking, this new description is valid only until one reaches a vertex $x$ for which the right-hand side of the above inequality is larger than 1. However, any such vertex must satisfy $|V_{\tau_x}| > n$ and, hence, $\psi_n(\tau_x) \to \infty$ as $n \to \infty$, by our assumption (A2). Thus, bridges starting from such remote nodes will not affect the convergence stated in Theorem 2.1, and we can now safely analyze this new version.

Let us first observe that $\{x \to y\} \subseteq \{x \rightharpoonup y\}$, where, for any $x, y \in V$,
$$\{x \rightharpoonup y\} := \{Y_x = y\} \cap \left\{ U_x \le \frac{|\partial V_{\tau_x}| - 1}{dn + 2k - 2|V_{\tau_x}| - 1} \right\}.$$

The condition $x \rightharpoonup y$ can be viewed as obtained from $x \to y$ by setting $\mathcal{F} = \varnothing$ in step 1 above. The advantage of this approximation is that it suppresses the dependencies between bridges. This makes the corresponding point processes $\mathcal{N}^{(n)}_{i \rightharpoonup j}$ much easier to analyze than $\mathcal{N}^{(n)}_{i \to j}$.

**Lemma 3.1.** *Theorem 2.1 holds if we consider $\mathcal{N}^{(n)}_{i \rightharpoonup j}$ instead of $\mathcal{N}^{(n)}_{i \to j}$.*

*Proof.* Since the events $\{x \rightharpoonup y\}$ are independent as $x$ varies, it is enough, by the standard Poisson approximation theorem, to fix $1 \le i, j \le k$ and $0 \le s < t$ and to verify that the following two conditions hold:
$$\max_{\{x \in T_i \colon s \le \psi_n(\tau_x) < t\}} \mathbb{P}(x \rightharpoonup T_j) \to 0 \quad \text{as } n \to \infty, \tag{3.1}$$
$$\sum_{\{x \in T_i \colon s \le \psi_n(\tau_x) < t\}} \mathbb{P}(x \rightharpoonup T_j) \to \frac{\omega_i \omega_j (t - s)}{2d(d-2)} \quad \text{as } n \to \infty. \tag{3.2}$$

By construction we have
$$\mathbb{P}(x \rightharpoonup T_j) = f_n(\tau_x) \quad \text{with} \quad f_n(h) = \frac{|\partial V_h \cap T_j| - \mathbf{1}_{\{i=j\}}}{dn + 2k - 2|V_h| - 1}.$$

Note that $h \mapsto f_n(h)$ is nondecreasing and, therefore, $s \le \psi_n(\tau_x) < t$ implies that
$$(f_n \circ \psi_n^{-1})(s) \le \mathbb{P}(x \rightharpoonup T_j) \le (f_n \circ \psi_n^{-1})(t).$$

Since $T_j$ is $d$-regular, we have $|\partial V_h \cap T_j| = (d-2)|V_h \cap T_j| + 2$ for any $h > 0$. Using assumption

(A2), we obtain the following asymptotics:

$$\sqrt{n}(f_n \circ \psi_n^{-1})(s) \to \frac{\omega_j \sqrt{s}}{d} \quad \text{as } n \to \infty,$$
$$\sqrt{n}(f_n \circ \psi_n^{-1})(t) \to \frac{\omega_j \sqrt{t}}{d} \quad \text{as } n \to \infty,$$
$$\frac{1}{\sqrt{n}} |\{x \in T_i : s \leq \psi_n(\tau_x) < t\}| \to \frac{\omega_i(\sqrt{t} - \sqrt{s})}{d-2} \quad \text{as } n \to \infty.$$

Thus, condition (3.1) is guaranteed, as well as the following inequalities:

$$\liminf_{n\to\infty} \sum_{\{x \in T_i : s \leq \psi_n(\tau_x) < t\}} \mathbb{P}(x \rightharpoonup T_j) \geq \frac{\omega_i \omega_j}{d(d-2)} \sqrt{s}(\sqrt{t} - \sqrt{s}),$$
$$\limsup_{n\to\infty} \sum_{\{x \in T_i : s \leq \psi_n(\tau_x) < t\}} \mathbb{P}(x \rightharpoonup T_j) \leq \frac{\omega_i \omega_j}{d(d-2)} \sqrt{t}(\sqrt{t} - \sqrt{s}).$$

When $t$ is sufficiently close to $s$, $\sqrt{s}(\sqrt{t} - \sqrt{s}) \approx \sqrt{t}(\sqrt{t} - \sqrt{s}) \approx (t-s)/2$. Thus, condition (3.2) follows by subdividing $[s, t)$ into $M$ small intervals, adding up the inequalities obtained on each of them, and finally letting $M \to \infty$.

To prove Theorem 2.1, it now simply remains to show that our upper bound $\mathcal{N}_{i \rightharpoonup j}^{(n)}$ is asymptotically indistinguishable from the true process $\mathcal{N}_{i \to j}^{(n)}$.

**Lemma 3.2.** *For any* $1 \leq i, j \leq k$ *and* $t > 0$,

$$\mathcal{N}_{i \rightharpoonup j}^{(n)}[0, t) - \mathcal{N}_{i \to j}^{(n)}[0, t) \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty.$$

*Proof.* Set $\mathcal{X}_n = \{x \in V : \psi_n(\tau_x) < t\}$ and $\bar{\mathcal{X}}_n = \mathcal{X}_n \cup \partial \mathcal{X}_n$. We want to show that the set of pairs $(x, y) \in \mathcal{X}_n \times \bar{\mathcal{X}}_n$ satisfying $\{x \rightharpoonup y\}$ coincides with high probability (as $n \to \infty$) with the set of pairs $(x, y) \in \mathcal{X}_n \times \bar{\mathcal{X}}_n$ satisfying $\{x \to y\}$. By construction, there are only two possibilities for those sets to differ.

The first possibility is that, for some $(x, y) \in \mathcal{X}_n \times \bar{\mathcal{X}}_n$, $\{x \rightharpoonup y\}$ occurs but one of the two branches $[o, x]$, $[o, y]$ already contains the head or tail of a bridge, thereby violating the requirement $x \notin \mathcal{F}$, $y \notin \mathcal{F}$. In particular, $\{u \rightharpoonup v\}$ must occur for some $u \in [o, x) \cup [o, y)$, $v \in \bar{\mathcal{X}}_n$ or some $u \in \mathcal{X}_n$, $v \in [o, x] \cup [o, y]$. Assumptions (A2) and (A3) guarantee that this is asymptotically unlikely, by a simple union bound: the number of choices for $(x, y, u, v)$ is bounded above by

$$4 \times \Big(\max_{x \in \bar{\mathcal{X}}_n} |[o, x]|\Big) \times |\bar{\mathcal{X}}_n|^3 = o(n^2),$$

and, for each such choice, the probability of joint occurrence $\{x \rightharpoonup y\} \cap \{u \rightharpoonup v\}$ is (by independence) at most

$$\Big(\max_{(x,y) \in \mathcal{X}_n \times \bar{\mathcal{X}}_n} \mathbb{P}_n(x \rightharpoonup y)\Big)^2 = \bigg(\frac{1}{dn + 2k - 2|\mathcal{X}_n| - 1}\bigg)^2 = \Theta\bigg(\frac{1}{n^2}\bigg).$$

The second possibility is that, for some $x \in \mathcal{X}_n$, the uniformly distributed mark $U_x \in [0, 1]$ falls between the real threshold and its approximation, i.e.

$$\frac{|\partial V_{\tau_x}| - 1}{dn + 2k - 2|V_{\tau_x} \setminus \mathcal{F}| - 1} < U_x \leq \frac{|\partial V_{\tau_x}| - 1}{dn + 2k - 2|V_{\tau_x}| - 1}.$$

https://doi.org/10.1239/jap/1378401241 Published online by Cambridge University Press

Again, this is highly unlikely since the difference between these thresholds is at most

$$\frac{2|\bar{\mathcal{X}}_n|^2}{(dn + 2k - 2|\mathcal{X}_n| - 1)^2} = \Theta\left(\frac{1}{n}\right),$$

while there are only $\Theta(\sqrt{n})$ nodes in $\mathcal{X}_n$ by assumption (A2).

## 4. From the configuration model to the uniform simple model

In this last section we transfer the result obtained for the configuration model (Theorem 1.1) to the uniform simple graph model (Theorem 1.2). Write SIMPLE for the event that the random $d$-regular multigraph $G$ generated by the configuration model on $\{1, \ldots, n\}$ is simple. Recall from [11] (see also [23] for a generalization to graphs with an arbitrary prescribed degree sequence) that, conditionally on SIMPLE, $G$ is a uniform simple $d$-regular graph, and also that

$$\mathbb{P}(\text{SIMPLE}) \to \varrho := \mathrm{e}^{(1-d^2)/4} \quad \text{as } n \to \infty. \tag{4.1}$$

Fix $k \geq 2$ and a continuous bounded function $\psi\colon \mathbb{R}^{k(k-1)/2} \to \mathbb{R}$. Writing $\Xi_n = \{\mathcal{D}_{i,j} - \log n/(d-2)\colon 1 \leq i < j \leq k\}$ and $\Xi_* = \{W_{i,j} : 1 \leq i < j \leq k\}$, we want to show that

$$\mathbb{E}[\psi(\Xi_n) \mid \text{SIMPLE}] \to \mu := \mathbb{E}[\psi(\Xi_*)] \quad \text{as } n \to \infty. \tag{4.2}$$

In order to do so, let us fix an integer $m \geq 1$ and introduce, for each $1 \leq \ell \leq m$, the 'shifted replica'

$$\Xi_n^{\ell} := \left\{\mathcal{D}_{i,j} - \frac{\log n}{d-2}\colon (\ell - 1)k + 1 \leq i < j \leq \ell k\right\}.$$

Since $km$ is fixed, we may use our Theorem 1.1 and the independence of disjoint diagonal blocks in the limiting random array $\mathbb{W}$ to deduce that

$$\{\Xi_n^{\ell}\colon 1 \leq \ell \leq m\} \xrightarrow{\text{D}} \{\Xi_*^{\ell}\colon 1 \leq \ell \leq m\} \quad \text{as } n \to \infty, \tag{4.3}$$

where $\{\Xi_*^{\ell}\colon 1 \leq \ell \leq m\}$ are independent and identically distributed copies of $\Xi_*$. Now, for each $n \geq km$, the sequence $\{\Xi_n^{\ell}\colon 1 \leq \ell \leq m\}$ is exchangeable and, thus,

$$|\mathbb{E}[\psi(\Xi_n) \mid \text{SIMPLE}] - \mu| \leq \frac{1}{\mathbb{P}(\text{SIMPLE})}\mathbb{E}\left[\left|\frac{1}{m}\sum_{\ell=1}^{m}\psi(\Xi_n^{\ell}) - \mu\right|\right].$$

Letting $n \to \infty$ and using (4.1) and (4.3), we obtain

$$\limsup_{n\to\infty} |\mathbb{E}[\psi(\Xi_n) \mid \text{SIMPLE}] - \mu| \leq \frac{1}{\varrho}\mathbb{E}\left[\left|\frac{1}{m}\sum_{\ell=1}^{m}\psi(\Xi_*^{\ell}) - \mu\right|\right].$$

Since this holds for each $m \geq 1$ and since the $\{\Xi_*^{\ell}\colon 1 \leq \ell \leq m\}$ are independent copies of $\Xi_*$, we may finally let $m \to \infty$ to obtain (4.2).

## Acknowledgements

This research was supported by the NSF grants DMS-0704159 and CNS-0910702. The author warmly thanks David Aldous for suggesting the problem to him, and Louigi Addario-Berry, Shankar Bhamidi, and Gabor Lugosi for their helpful comments after a talk on the subject at the Banff Centre.

## References


[1] Aldous, D. J. (2010). More uses of exchangeability: representations of complex random structures. In *Probability and Mathematical Genetics* (London Math. Soc. Lecture Notes **378**), Cambridge University Press, pp. 35–63.

[2] Aldous, D. J. and Bhamidi, S. (2010). Edge flows in the complete random-lengths network. *Random Structures Algorithms* **37,** 271–311.

[3] Amini, A. and Lelarge, M. (2011). The diameter of weighted random graphs. Preprint. Available at http://arXiv.org/abs/1112.6330v1.

[4] Antunović, T., Dekel, Y., Mossel, E. and Peres, Y. (2011). Competing first passage percolation on random regular graphs. Preprint. Available at http://arXiv.org/abs/1109.2575v1.

[5] Athreya, K. B. and Ney, P. E. (2004). *Branching Processes*. Dover Publications, Mineola, NY.

[6] Bhamidi, S. (2008). First passage percolation on locally treelike networks. I. Dense random graphs. *J. Math. Phys.* **49,** 125218, 27pp.

[7] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2010). Extreme value theory, Poisson-Dirichlet distributions, and first passage percolation on random networks. *Adv. Appl. Prob.* **42,** 706–738.

[8] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2010). First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Prob.* **20,** 1907–1965.

[9] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2011). First passage percolation on the Erdős-Rényi random graph. *Combinatorics Prob. Comput.* **20,** 683–707.

[10] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2012). Universality for first passage percolation on sparse random graphs. Preprint. Available at http://arXiv.org/abs/1210.6839v1.

[11] Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combinatorics* **1,** 311–316.

[12] Cox, J. T. and Durrett, R. (1981). Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Prob.* **9,** 583–603.

[13] Deijfen, M. and Häggström, O. (2006). The initial configuration is irrelevant for the possibility of mutual unbounded growth in the two-type Richardson model. *Combinatorics Prob. Comput.* **15,** 345–353.

[14] Deijfen, M. and Häggström, O. (2006). Nonmonotonic coexistence regions for the two-type Richardson model on graphs. *Electron. J. Prob.* **11,** 331–344.

[15] Deijfen, M. and Häggström, O. (2007). The two-type Richardson model with unbounded initial configurations. *Ann. Appl. Prob.* **17,** 1639–1656.

[16] Ding, J., Kim, J. H., Lubetzky, E. and Peres, Y. (2010). Diameters in supercritical random graphs via first passage percolation. *Combinatorics Prob. Comput.* **19,** 729–751.

[17] Garet, O. and Marchand, R. (2008). First-passage competition with different speeds: positive density for both species is impossible. *Electron. J. Prob.* **13,** 2118–2159.

[18] Häggström, O. and Pemantle, R. (1998). First passage percolation and a model for competing spatial growth. *J. Appl. Prob.* **35,** 683–692.

[19] Häggström, O. and Pemantle, R. (2000). Absence of mutual unbounded growth for almost all parameter values in the two-type Richardson model. *Stochastic Process. Appl.* **90,** 207–222.

[20] Hoffman, C. (2005). Coexistence for Richardson type competing spatial growth models. *Ann. Appl. Prob.* **15,** 739–747.

[21] Hoffman, C. (2008). Geodesics in first passage percolation. *Ann. Appl. Prob.* **18,** 1944–1969.

[22] Janson, S. (1999). One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combinatorics Prob. Comput.* **8,** 347–361.

[23] Janson, S. (2009). The probability that a random multigraph is simple. *Combinatorics Prob. Comput.* **18,** 205–225.

[24] Joseph, A. (2011). The component sizes of a critical random graph with given degree sequence. Preprint. Available at http://arXiv.org/abs/1012.2352v2.

[25] Kesten, H. (1986). Aspects of first passage percolation. In *École d'été de Probabilités de Saint-Flour, XIV—1984* (Lecture Notes Math. **1180**), Springer, Berlin, pp. 125–264.

[26] Van der Hofstad, R., Hooghiemstra, G. and Van Mieghem, P. (2001). First-passage percolation on the random graph. *Prob. Eng. Inf. Sci.* **15,** 225–237.

[27] Van der Hofstad, R., Hooghiemstra, G. and Van Mieghem, P. (2005). Distances in random graphs with finite variance degrees. *Random Structures Algorithms* **27,** 76–123.