Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

One Half or Two Thirds? Epidemic Spread on Random 2-Regular Partnership Graphs

Zhivko Nedev

nedevz@biust.ac.bw

Botswana International University of Science and Technology

Zdravko Botev

UNSW Sydney

Research Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11219447/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

One Half or Two Thirds? Epidemic Spread on Random 2-Regular Partnership Graphs

Zhivko Nedev[1*] and Zdravko Botev[2]

[1]Department of Computing and Informatics, Botswana International University of Science and Technology, Palapye, Botswana.
[2*]School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.

*Corresponding author(s). E-mail(s): nedevz@biust.ac.bw; znedev@gmail.com;
Contributing authors: botev@unsw.edu.au;

Abstract

We model the spread of a sexually transmitted infection on a partnership network in which every individual has exactly two partners of the opposite sex and no pair shares more than one partnership. The partnership graph is a random 2-regular simple bipartite graph on $\mathbf{2}\boldsymbol{N}$ vertices, which breaks into separate loops of people. Under certain transmission the infection travels all the way around whichever loop contains the first case, and cannot leave it: each loop is a group in which all would be infected together if the disease reached any one of them. We show that the central risk quantity—the probability that a named individual is infected by a single introduced case—depends decisively on *how* the network is randomised. Under the *Derangement Model*, appropriate when partnerships form in two organised rounds, this probability converges to $\mathbf{1/2}$; under the *Uniform Model*, appropriate for organic partnership formation, it converges to $\mathbf{2/3}$. We give exact $\boldsymbol{O(N)}$-running-time simulation algorithms for both models and a Monte Carlo study that verifies empirically the validity both algorithms. A mixed-degree variant, in which one man and one woman have one partner each while everyone else has two, illustrates the robustness of the $\mathbf{2/3}$ limit under the Uniform law. It also yields an exact finite-population result: those two one-partner (degree-one) individuals are always linked through a single chain of partnerships (are endpoints of the same path component), so if one is infected the other is eventually infected too.

Keywords: random bipartite graph, epidemic cluster, random permutation, derangement, configuration model, sexually transmitted infection

MSC Classification: 05C80 , 60C05 , 92D30 , 65C05

1 Introduction and Background

Consider a population of N men and N women in which every individual has exactly two partners of the opposite sex, and no specific pair shares more than one partnership. A sexually transmitted infection (STI) enters through one individual and passes with certainty along every partnership. Given one initial infected case, three questions face a public-health analyst: Which individual will eventually be infected?; What is the probability that a named uninfected individual is infected?; How large is the resulting outbreak expected to be?

The partnership graph here is a 2-regular simple bipartite graph on $2N$ vertices. Every vertex has exactly two incident edges, so each connected component is a cycle, and bipartiteness forces every cycle to have even length. Under certain transmission each cycle is a closed outbreak: infection reaches the whole cycle and never leaves it. The first question is therefore answered by the cycle containing the initial case, and the second reduces to a precise combinatorial quantity—*the probability that two named vertices lie in the same cycle of a random 2-regular bipartite graph.*

Epidemic spread on configuration-model networks is a well-developed field of study [8–10] that has found applications in the spread of STIs [6]. The contribution of this paper is to show that this probability is not well defined until one decides *how* the graph is randomised, and that two natural randomisations give materially different answers. This matters because the same contact-tracing data, analyzed under two equally defensible modelling conventions, yields expected outbreak sizes that differ by a third: about N people infected under one convention against about $4N/3$ under the other.

The two randomizations of the graph are summarized in the following examples.

Example 1 (Two organised rounds) N men and N women attend a closed event and are paired twice—two rounds of introductions—under the rule that no pair from the first round repeats in the second. The first round is a uniform random perfect matching M_1; the second is a uniform random perfect matching M_2 sharing no edge with M_1; the partnership graph is $G = M_1 \cup M_2$. We call this the *Derangement Model*, because M_2 is specified relative to M_1 by a uniform random derangement. It is the natural model whenever partnerships are formed in two structured waves (think of formally organized "speed dating" events etc).

Example 2 (Organic formation) N men and N women form partnerships over time through independent, unstructured encounters, with no record constraining which pairings occur. The appropriate model gives equal probability to every valid partnership graph. We call this the *Uniform Model*. It is the default when the mechanism of partnership formation is unknown.

The two models place different distributions on the same graph class. The Derangement Model assigns a graph G probability proportional to $2^{c(G)}$, where $c(G)$ is its number of cycles, because a graph with $c(G)$ cycles arises from $2^{c(G)}$ ordered pairs of edge-disjoint perfect matchings (Lemma 1). It therefore over-represents fragmented networks relative to the uniform distribution. Our main results, Theorems 2 and 3, quantify the consequence: the infection probability converges to $1/2$ under the Derangement Model and to $2/3$ under the Uniform Model.

The rest of the article is organized as follows. Section 2 fixes the model and states the bijection with random permutations that drives everything else. Section 2.1 states the two infection-probability theorems and the exact small-N values. Section 4 gives two exact simulation algorithms corresponding to the two models and Section 2.2 records the cluster-size and largest-cluster consequences. Section 2.3 treats the mixed-degree variant, where the $2/3$ value proves robust and the two degree-one individuals are certainly co-infected. Section 4 reports the Monte Carlo study and finally Section 5 discusses modelling implications and open problems. All proofs, together with supporting lemmas and definitions, are in Appendix A.

2 The Model and The Permutation Correspondence

Let $M = \{m_1, \ldots, m_N\}$ and $W = \{w_1, \ldots, w_N\}$ denote the "men" and "women" nodes, respectively. We write $\mathcal{B}_2\mathcal{G}(N)$ for the set of graphs on $M \cup W$ that are bipartite between M and W, 2-regular, and simple. In addition, we write $c(G)$ for the number of cycles of G, and $C_{2k}(G)$ for the number of its cycles of length $2k$.

Every component of $G \in \mathcal{B}_2\mathcal{G}(N)$ is an even cycle: each vertex has two incident edges, so a walk can neither branch nor stop, and bipartiteness makes every cycle even. Each such cycle is one outbreak; see Figure 1.

Every $G \in \mathcal{B}_2\mathcal{G}(N)$ splits into two edge-disjoint perfect matchings. Encoding a perfect matching as a permutation π via the edge set $\{(m_i, w_{\pi(i)})\}$, the requirement that the second matching repeats no edge of the first becomes $\sigma(i) \neq i$ for all i, where $\sigma = \pi_1^{-1}\pi_2$; that is, σ is a *derangement*. Let

$$D_N = N! \sum_{k=0}^{N} (-1)^k / k! \approx N!/e$$

be the number of derangements. We can now formally state the two models.

Derangement Model

Draw π_1 uniformly from the $N!$ permutations and, independently, draw σ uniformly from the D_N derangements; set $M_1 = \{(m_i, w_{\pi_1(i)})\}$, $M_2 = \{(m_i, w_{\pi_1(\sigma(i))})\}$, and $G = M_1 \cup M_2$.

Uniform Model

Draw G uniformly from $\mathcal{B}_2\mathcal{G}(N)$. The uniform, independent choice of π_1 is essential: it is what makes the induced law on G proportional to $2^{c(G)}$ (Lemma 1(ii)); fixing π_1 would instead restrict the support to graphs containing that one matching.



Fig. 1 Two elements of $\mathcal{B}_2\mathcal{G}(N)$. Solid edges form M_1, dashed edges M_2. Left: a single outbreak of all six individuals. Right: two independent outbreaks of four; infection in one never reaches the other.

The next lemma is the engine of the paper: it reduces every question about outbreaks to the cycle structure of a random permutation, and makes the difference between the two models precise. Its proof is in Appendix A.

Lemma 1 (Correspondence and model weights) *Under the Derangement Model:*

(i) $C_{2k}(G) = X_k(\sigma)$ *for every* k*, where* $X_k(\sigma)$ *is the number of* k*-cycles of* σ*; in particular a* k*-cycle of* σ *is an outbreak of* k *men and* k *women.*

(ii) *Each* $G \in \mathcal{B}_2\mathcal{G}(N)$ *occurs with probability proportional to* $2^{c(G)}$.

One immediate consequence of Lemma 1(ii) is that in general $|\mathcal{B}_2\mathcal{G}(N)| \neq N!D_N/2$. For $N = 4$ there are $N!D_N = 216$ ordered pairs, comprising 72 Hamiltonian graphs ($c = 1$, two decompositions each) and 18 two-cycle graphs ($c = 2$, four decompositions each), so $|\mathcal{B}_2\mathcal{G}(4)| = 90$ while $N!D_N/2 = 108$. The gap is exactly the $2^{c(G)}$ weighting that separates the two models.

2.1 The infection probability

Fix two individuals $A \neq B$, with A the initial case. Under certain transmission B is infected exactly when A and B lie in the same cycle. For two men m_a, m_b this is the event that the indices $a \neq b$ lie in the same cycle of σ (Lemma 1(i)), and the same holds for two women after relabelling. A man and a woman need a separate treatment, because they may be joined directly by an edge of M_1; this is handled by equation (2) below and its Uniform analogue. Write $\mathbb{P}_D$ and $\mathbb{P}_U$ for probabilities under the two models.

Theorem 2 (Infection probability, Derangement Model) *Let* $q_N = D_N^{\text{same}}/D_N$ *be the probability that two indices* $a \neq b$ *share a cycle of a uniform random derangement, where* D_N^{same} *counts derangements placing* a, b *in one cycle. Then the successor-swap involution (Appendix A) yields the exact closed form*

$$q_N = \frac{D_N^{\text{same}}}{D_N} = \frac{1}{2} + \frac{1}{N-1} - \frac{D_{N-2}}{2D_N}, \qquad N \geq 2. \tag{1}$$

Under the Derangement Model the named-individual infection probability then depends on the pair type: for two men (or two women) it equals q_N, while for a man and a woman, with π_1 uniform,

$$\mathbb{P}_D(\textit{a man and a woman infected}) = \frac{1}{N} + \frac{N-1}{N}\, q_N. \tag{2}$$

Note that $D_{N-2}/D_N \sim 1/\{N(N-1)\} = \Theta(N^{-2})$, giving $q_N = \frac{1}{2} + \frac{1}{N} + O(N^{-2})$. Moreover $q_N > \frac{1}{2}$ for every $N \geq 2$, because $D_N = (N-1)(D_{N-1} + D_{N-2}) \geq (N-1)D_{N-2}$ makes the last term of (1) at most $\frac{1}{2(N-1)}$. As a result of this, in Theorem 2 both expressions converge to 1/2. There are no same-sex partnerships in the model; the distinction here is only between two named individuals on the same side of the bipartition and on opposite sides.

We next state the corresponding result for the Uniform Model.

Theorem 3 (Infection probability, Uniform Model) *Under the Uniform Model, let p_N be the probability that two men (or two women) lie in the same outbreak. Then $p_N > \frac{2}{3}$ for every $N \geq 2$,*

$$p_N = \frac{2}{3} + \frac{2}{3N} + O(N^{-2}),$$

and an exact finite-N expression is given in (A2) (Appendix A).

From Theorem 3 we can conclude that for a man and a woman the probability of infection is $\frac{1}{N} + \frac{N-1}{N}\, p_N$, which converges to 2/3.

Both theorems rest on one construction, the successor-swap involution ϕ (Appendix A): ϕ exchanges same-cycle and different-cycle permutations, and raises the cycle count by one on the same-cycle side. For the counting measure this gives equal class sizes, hence 1/2; for the Uniform Model's $2^{-c(\sigma)}$ weighting the extra cycle contributes a factor $\frac{1}{2}$, tilting the ratio to 2:1 and hence 2/3. This is the main finding of this paper.

Given one initial case A, let $|\mathcal{C}(A)|$ be the number of people in its outbreak. Under the Derangement Model, Theorem 2 then suggests that

$$\mathbb{E}_D[|\mathcal{C}(A)|] = 2\big\{1 + (N-1)q_N\big\} = N + 3 - \frac{(N-1)D_{N-2}}{D_N} = N + 3 + O(N^{-1}),$$

with q_N as in (1); for instance $\mathbb{E}_D[|\mathcal{C}(A)|] = 20/3$ at $N = 4$.

In contrast, under the Uniform Model Theorem 3 gives

$$\mathbb{E}_U[|\mathcal{C}(A)|] = 2\{1 + (N-1)p_N\} = \tfrac{4N}{3} + 2 + O(N^{-1}),$$

with p_N as in Theorem 3. Thus for large N the expected outbreak is about N under organised formation and about $4N/3$ under organic formation—the modelling convention alone changes it by roughly a factor of 1/3.

Table 1 records the numerically exact value of the probability q_N of (1) that two named individuals lie in the same outbreak (for a pair of two men or women). For a man and a woman, we use equation (2) or its Uniform analogue in Theorem 3. For

$N \leq 10$ they are obtained by complete enumeration of all D_N derangements (with the 2^{-c} reweighting of Lemma 1(ii) for the final column), which agrees with the closed forms (1) and (A2); for $N \geq 20$ they are computed from those closed forms. The integer counts D_N and D_N^{same} are shown only for $N \leq 10$.

Table 1 Probability that two named individuals lie in the same outbreak, for a pair of two men (or two women); this is q_N of (1).

N	D_N	D_N^{same}	$\mathbb{P}_D(\text{same})$	$\mathbb{P}_U(\text{same})$
2	1	1	1.000000	1.000000
3	2	2	1.000000	1.000000
4	9	7	0.777778	0.866667
5	44	32	0.727273	0.823529
6	265	181	0.683019	0.793377
7	1 854	1 214	0.654800	0.773188
8	14 833	9 403	0.633924	0.758523
9	133 496	82 508	0.618056	0.747407
10	1 334 961	808 393	0.605556	0.738690
20	—	—	0.551316	0.701293
50	—	—	0.520204	0.680203
100	—	—	0.510051	0.673384
200	—	—	0.505013	0.670013
500	—	—	0.502002	0.668002
∞	—	—	1/2	2/3

If the analyst knows partnerships formed in organised rounds, the named-contact infection probability is near 1/2; if the formation mechanism is unknown, the conservative figure is 2/3. The gap between the two is already within one percent of its limit (a factor of 1/3) when N is in the few hundreds.

2.2 Outbreak-size structure

Beyond the pairwise infection probability, an analyst needs to know the size of the outbreak an initial case sits in, and the size of the largest outbreak in the population. Both follow from Lemma 1 and classical random-permutation asymptotics.

Definition 1 (Poisson–Dirichlet distribution) For a parameter $\theta > 0$, let $V_1, V_2, \ldots$ be independent with $V_i \sim \text{Beta}(1, \theta)$, and set the stick-breaking fractions

$$P_1 = V_1, \quad P_2 = (1 - V_1)V_2, \quad P_3 = (1 - V_1)(1 - V_2)V_3, \ \ldots$$

The sequence $(P_1, P_2, \ldots)$ follows the Griffiths–Engen–McCloskey distribution $\text{GEM}(0, \theta)$, and its decreasing rearrangement follows the Poisson–Dirichlet distribution $\text{PD}(0, \theta)$; here θ is the concentration parameter, with smaller θ concentrating more mass on the largest atoms [11]. (These are the $\alpha = 0$ cases of the two-parameter family $\text{PD}(\alpha, \theta)$, obtained by taking $V_i \sim \text{Beta}(1 - \alpha, \theta + i\alpha)$; only $\alpha = 0$ occurs in this paper.)

Proposition 4 (Cluster counts) *For each fixed $k \geq 2$, as $N \to \infty$,*

$$C_{2k} \xrightarrow{d} \text{Poisson}(1/k) \text{ under } \mathbb{P}_D, \qquad C_{2k} \xrightarrow{d} \text{Poisson}(1/(2k)) \text{ under } \mathbb{P}_U.$$

The expected number of outbreaks satisfies $\mathbb{E}_D[c(G)] = H_N - 1 + O(1)$ with $\mathbb{E}_D[c(G)] \sim \log N$, and $\mathbb{E}_U[c(G)] \sim \frac{1}{2}\log N$, where H_N is the Nth harmonic number.

Note that in the result above, the conditioning on no fixed points shifts the finite-N value away from $H_N - 1$ exactly; e.g. $\mathbb{E}_D[c(G)] = 4/3$ at $N = 4$ against $H_4 - 1 = 13/12$.

In addition, the fixed-k limits of Proposition 4 must not be summed over k to obtain $\mathbb{E}_U[c(G)]$: the approximation fails once k is of order N, and $\sum_k 1/(2k)$ understates the truth. At $N = 300$, simulation gives $\mathbb{E}_U[c(G)] \approx 3.3$ against $(H_N - 1)/2 \approx 2.6$, and the model ratio is about 1.6, not 2.

Proposition 5 (Largest outbreak) *The size-ordered outbreak fractions converge to* $\text{PD}(0, 1)$ *under $\mathbb{P}_D$ and to* $\text{PD}(0, \frac{1}{2})$ *under $\mathbb{P}_U$.*

The expected largest-outbreak fraction is ≈ 0.6243 (the Golomb–Dickman constant) under $\mathbb{P}_D$ and ≈ 0.758 under $\mathbb{P}_U$; the latter is the expected largest atom of $\text{PD}(0, \frac{1}{2})$, obtained here by $\text{GEM}(0, \frac{1}{2})$ stick-breaking simulation (2×10^5 draws) and cross-checked against direct graph simulation (Section 4). The probability that the whole population is one outbreak is $(N-1)!/D_N \sim e/N$ under $\mathbb{P}_D$ and of order $N^{-1/2}$ under $\mathbb{P}_U$.

Thus a single introduced case can, with the stated probability, reach the entire population; short of that, it typically reaches a dominant cluster covering roughly 62% (organised formation) to 76% (organic formation) of everyone. The reweighting that raises the infection probability from 1/2 to 2/3 is the same one that enlarges the dominant outbreak: fewer, larger cycles. Again, this observation is the main contribution of this article.

2.3 A mixed-degree variant

Contact data routinely contain individuals who report a single partner. We model the simplest such perturbation: one woman a and one man z have degree 1, all others degree 2, and G is uniform among simple bipartite graphs with this degree sequence. Two questions arise: whether lowering these degrees changes the infection probability, and what happens to the two degree-one individuals themselves. The answers are that the limit is unchanged—it remains the 2/3 of the Uniform law (Theorem 3)—while the two degree-one individuals acquire an exact and absolute link.

Proposition 6 (Path–cycle decomposition) *Every such G consists of exactly one path with endpoints a and z, of odd length $2T - 1$ for some $T \in \{1, \ldots, N\}$, together with even cycles covering the remaining vertices.*

Because the path is a single connected chain, a and z are always co-infected: $\mathbb{P}(z \text{ infected} \mid a \text{ initial}) = 1$ exactly, for every N. The substantive question concerns a generic degree-2 pair.

Theorem 7 (Mixed-degree infection probability) *Let A, B be distinct degree-2 individuals. Then*

$$\mathbb{P}(A, B \text{ in the same component}) \longrightarrow \frac{2}{3} \qquad (N \to \infty).$$

The proof (Appendix A) conditions on the path half-length T, whose rescaled law T/N converges to the density $f(x) = \frac{1}{2}(1-x)^{-1/2}$ on $(0,1)$; the residual cycles form a Uniform-Model graph, so their cycle-sharing probability is the 2/3 of Theorem 3, and the two combine through

$$\int_0^1 \left[x^2 + \tfrac{2}{3}(1-x)^2\right] \frac{dx}{2\sqrt{1-x}} = \frac{8}{15} + \frac{2}{15} = \frac{2}{3}.$$

Two warnings are in order about what the mixed-degree result does and does not say!

First, the comparison must hold the randomisation law fixed. Under the Uniform law the 2-regular infection probability is already 2/3 (Theorem 3), and the mixed-degree value is also 2/3; so lowering the degree of one couple does *not* change the limit, and the mixed-degree calculation should be read as a robustness check, not as an effect of reduced degree. A comparison against the 2-regular Derangement value 1/2 would change both the degree sequence and the law at once and is not a like-for-like statement. Numerically the two Uniform-law models agree at every finite N we tested, to within Monte Carlo error.

Second, the one genuinely finite-population fact here is exact and does not depend on the law: the two degree-one individuals a and z are always the endpoints of the single path component (Proposition 6) and are therefore certain to lie in the same outbreak, $\mathbb{P}(z \text{ infected} \mid a \text{ initial}) = 1$ for every N. For an analyst the operational point is thus not that low degree raises system-wide risk—it does not, under a fixed law—but that a single-partner individual is not a dead end: it sits at one end of the graph's unique open chain, which covers on average two-thirds of the population, and the two such individuals are bound to the same outbreak with certainty.

3 Recommendations for contact tracing strategies

The structural results above are not only descriptive; they dictate an efficient control procedure. Because every outbreak is a single cycle (Section 2), the set of infected individuals reachable from a diagnosed case has no branching: each infected person has exactly two partners, one "upstream" and one "downstream" along the chain. This is the operational difference from tracing on a general contact network, where each interview (by the health authorities) can spawn several new leads and effort can grow combinatorially. Here it cannot.

Algorithm 1 below states our suggested protocol for contact tracing.

Algorithm 1 Bidirectional outbreak tracing

Require: diagnosed index case v_0

interview v_0; add the partners it names to found, together with v_0
make each named partner an active frontier ▷ one frontier only, if v_0 names one partner
repeat
    for each active frontier x **do**
        interview x
        if x names a partner $y \notin$ found **then**
            add y to found; move this frontier to y
        else
            mark this frontier stopped ▷ it met the other frontier, or reached a path endpoint
until no frontier is active
return found ▷ the complete outbreak

Starting from a diagnosed index case, the tracer advances two frontiers outward along the chain, one round at a time. Each frontier runs independently and stops on its own when it can go no further—because it has reached the other frontier (a cycle) or because it has reached a person with no further partner, a degree-one endpoint (a path). Three facts make this reliable in the field.

(a) *Completeness.* Every member of the index case's outbreak is found, because the outbreak is a single connected chain—a cycle or a path—and the two frontiers together sweep it entirely (see Proposition 8 below).
(b) *A definite stopping rule.* Each frontier carries its own stop signal, available in the field: it halts when the next contact named is one already interviewed (the frontiers have met, on a cycle) or when the person interviewed names no second partner (an endpoint, on a path). Tracing ends when both frontiers have stopped, so the tracer never has to guess whether more cases remain.
(c) *Bounded, predictable effort.* Each member is interviewed once. On a cycle of L people every member is identified within $\lfloor L/2 \rfloor$ rounds, against $L-1$ for one-directional tracing. On a path the figure is $\max(d_1, d_2)$, where d_1, d_2 are the index case's distances to the two endpoints; this is smallest when the index case is central and as large as $L-1$ when it is itself an endpoint (see Proposition 8 below).

All of these properties are formally summarized in the following proposition.

Proposition 8 (Tracing is complete, with explicit depth) *Algorithm 1 identifies every member of the index case's outbreak and interviews each member exactly once. Count the interview of the index case as round 1, and in each later round interview every active frontier. If the outbreak is a cycle of L people, every member is identified after $\lfloor L/2 \rfloor$ rounds, against $L-1$ for one-directional tracing. If it is a path of L people and the index case is at distances d_1, d_2*

from its endpoints, every member is identified after $\max(d_1, d_2)$ *rounds, which ranges from* $\lceil (L-1)/2 \rceil$ *(central index case) to* $L-1$ *(index case at an endpoint). In both cases the procedure halts one round later, when the last frontier interviews name no new person.*

What do these results imply for resource planning?

Proposition 5 supplies planning averages, not guarantees. The share of the population in the largest outbreak is itself random. Its mean is about 0.62 when partnerships formed in organised rounds and about 0.76 when they formed organically, but in the large-N limit the middle 80% of outcomes runs from about 0.38 to 0.91 and from about 0.48 to 0.99 respectively (the 10th and 90th percentiles of $\mathrm{PD}(0,1)$ and $\mathrm{PD}(0,\frac{1}{2})$, from 2×10^5 stick-breaking draws). A uniformly chosen index case lies in the largest outbreak with probability equal to that share, so tracing need not begin there. For a population of $N = 1000$ men and women, a largest outbreak of average size needs about 620 rounds (Derangement) or 760 rounds (Uniform) under Algorithm 1, while the percentile ranges above correspond to roughly 380–910 and 480–990 rounds. A unit planning interviewer-days should therefore work with a range. What the depth bound does give firmly is conditional: once tracing is inside an outbreak of L people, every member is identified within $\lfloor L/2 \rfloor$ rounds.

Scope of the guarantees.

The completeness, stopping rule, and depth bound above are exact for the model as stated: certain transmission, exactly two partners per person, and a static network. In practice all three idealisations loosen. Sub-certain transmission breaks an outbreak into sub-clusters, so a single sweep may recover only the part of the cycle actually infected; higher-degree individuals reintroduce branching and with it the lead proliferation the cycle structure removes; and partnership turnover during the tracing window can connect what was, at diagnosis, a closed chain. The protocol therefore gives an exact procedure for the idealised model and a starting approximation otherwise, to be refined as the assumptions are relaxed (Section 5). We make no one-directional claim about how the reachable set or the tracing depth change when the assumptions loosen: higher degrees or temporal connections can enlarge the reachable set, while imperfect transmission can shrink both it and the depth.

The degree-1 warning, operationally.

Section 2.3 showed that a single-partner individual is not a dead end: it is an endpoint of the graph's one path component, and the two such individuals are certain to share an outbreak. For a tracer this reframes a common instinct—a single-partner contact looks like a low priority, but it sits at one end of the graph's unique open chain, and its counterpart degree-one individual is guaranteed to be in the same outbreak. Algorithm 1 traces such a chain correctly because each frontier stops on its own: the frontier that reaches a degree-one endpoint halts while the other continues to the far endpoint (Proposition 8). No special handling is needed, but such contacts should not be deprioritised.

4 Exact Simulation via Rejection Sampling

Generation of random regular graphs is a common computational tool [2, 13]. Both of our models admit exact sampling in $O(N)$ expected time by rejection sampling, with no burn-in and no approximation; see Algorithms 2 and 3 below. Numerical simulation via Algorithms 2 and 3 is what an analyst would use to estimate any functional of the outbreak not covered in closed form in this article.

Algorithm 2 Exact sampler, Derangement Model

Require: population size N

draw a uniform permutation π_1 by Fisher–Yates
repeat
    draw a uniform permutation σ by Fisher–Yates
until $\sigma(i) \neq i$ for all i
return $G = \{(m_i, w_{\pi_1(i)})\} \cup \{(m_i, w_{\pi_1(\sigma(i))})\}$

Algorithm 3 Exact sampler, Uniform Model

Require: population size N

repeat
    form the stub list $L = (w_1, w_1, w_2, w_2, \ldots, w_N, w_N)$
    shuffle L uniformly by Fisher–Yates
    assign to m_i the entries of L in positions $2i - 1$ and $2i$
until no man receives the same woman twice
return the resulting graph G

The correctness of the simulation in Algorithms 2 and 3 is formally established by the following, whose proof is in Appendix A.

Proposition 9 (Correctness and cost) *Algorithm 2 returns a Derangement-Model sample; its acceptance probability is $D_N/N! \to e^{-1}$, so it uses e trials on average. Algorithm 3 returns a Uniform-Model sample; its acceptance probability tends to $e^{-1/2}$, so it uses $e^{1/2}$ trials on average. For both algorithms each trial costs $O(N)$.*

The two samplers/algorithms look alike—each shuffles a list and rejects on a coincidence—but they shuffle different objects (indices versus a stub multiset), and that difference is exactly what separates the two laws on $\mathcal{B}_2\mathcal{G}(N)$.

We now proceed to numerically validated the two algorithms. Because Table 1 is exact, simulation serves here to check the accuracy of the two samplers and any quantities that have no simple closed form. All runs use a fixed seed. The exact values in Table 1 lie above their limits, as (1) and (A2) require, and decrease in N over the range tabulated. Monte Carlo estimates from 8×10^4 draws of each sampler at

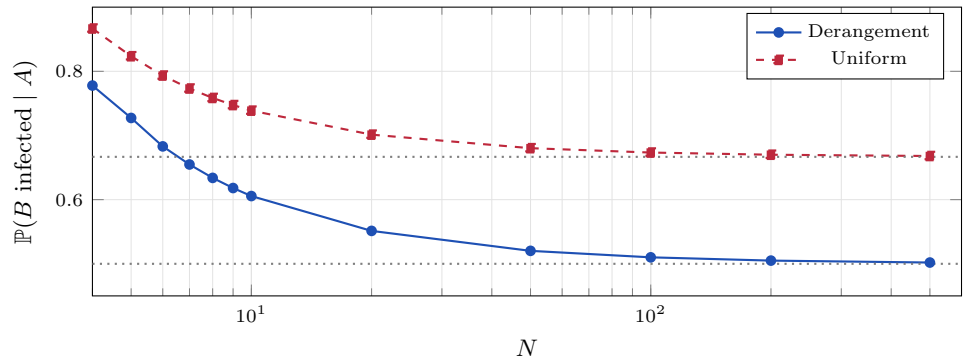


Fig. 2 Exact infection probability for two men (or two women) against N under both models, approaching the limits 1/2 and 2/3 (dotted) from above, with $q_N - \frac{1}{2} \sim \frac{1}{N}$ and $p_N - \frac{2}{3} \sim \frac{2}{3N}$.

$N = 20, 50, 100, 200, 500$ agree with the exact entries to within 0.0030, inside the 95% half-width of 0.0035.

For Propositions 4 and 5, 4000 draws of each model at $N = 300$ give $\mathbb{E}_D[C_4] \approx 0.499$ and $\mathbb{E}_D[C_6] \approx 0.343$ against 1/2 and 1/3, $\mathbb{E}_U[C_4] \approx 0.252$ and $\mathbb{E}_U[C_6] \approx 0.158$ against 1/4 and 1/6, and mean largest-outbreak fractions 0.631 and 0.758 against 0.6243 and 0.758. Under the Uniform Model the probability that the whole population forms one outbreak, multiplied by $\sqrt{N}$, is 1.48, 1.47 and 1.47 at $N = 50, 100, 200$, consistent with order $N^{-1/2}$.

Table 2 confirms the mixed-degree results: the endpoint co-infection probability is 1 to three decimals on every row, and both the generic-pair infection probability and the path's population share converge to 2/3. Sampling from the exact simple-graph law also confirms that the path-fraction T/N follows the density $f(x) = \frac{1}{2}(1-x)^{-1/2}$ predicted for the conditioned model in the proof of Theorem 7: for example its mean converges to 2/3 and its CDF at $x = \frac{1}{2}$ to $1 - \frac{1}{\sqrt{2}} \approx 0.293$, with the boundary value $t = N - 1$ never occurring, as simplicity requires.

Table 2 Mixed-degree variant, stub-matching sampler with rejection (1.5×10^4 samples for $N \le 100$, 2×10^4 for $N = 200$, 8×10^3 for $N = 500$). Confidence half-widths for the generic-pair column are below 0.011. Column 2 is exactly 1 by Proposition 6.

N	$\mathbb{P}(z \mid a)$	$\mathbb{P}(B \mid A)$, deg-2	$\mathbb{E}$[path vertices]$/2N$
20	1.000	0.7033	0.6976
50	1.000	0.6815	0.6781
100	1.000	0.6745	0.6725
200	1.000	0.6743	0.6706
500	1.000	0.6690	0.6687
∞	1.000	2/3	2/3

5 Final Discussion and Conclusions

The practical message of this work has two parts. First, a control rule: because every outbreak is a single non-branching chain, bidirectional tracing (Algorithm 1) finds the whole outbreak and carries a definite stopping signal, identifying every member of a cyclic outbreak of L people within $\lfloor L/2 \rfloor$ rounds against $L-1$ for one-directional tracing, and single-partner contacts—path endpoints—should not be deprioritised.

Second, a risk figure: on a two-partner contact network, the probability that a named individual is reached by one introduced case is $1/2$ if partnerships were formed in organised rounds and $2/3$ if they formed organically, and this gap propagates to a factor of one-third difference in expected outbreak size.

Because the two models differ only in the $2^{c(G)}$ weighting, an analyst who cannot justify the organised-rounds mechanism should default to the Uniform Model, which is the more conservative choice and the natural uniform baseline given only the degree constraints (equal probability to all admissible labelled graphs is itself a modelling assumption, not the absence of one). This model choice matters only when the graph is unobserved: given the full partnership graph, the outbreak of any index case is already determined and neither prior applies. The distinction is therefore about inference from partial information—a known degree sequence or formation mechanism but an unknown graph—which is the usual situation in practice.

The exact samplers of Section 4 make either model available at negligible cost for any functional not in closed form here, and the mixed-degree result warns against reading low reported degree as low risk.

Three limitations frame the natural extensions. First, transmission is certain here, which is not always realistic; with per-contact probability $p < 1$ over T contacts the effective edge probability is $1-(1-p)^T$. Conditional on the cycle length, the component of a seed under independent bond percolation has an elementary truncated-geometric description; obtaining the unconditional outbreak-size law by mixing this over the random cycle structure is the concrete open problem.

Second, the network is static; temporal reordering of partnerships is known to change reach [4] and would be the most consequential generalisation.

Finally, relaxing the degree of exactly two; the same successor-swap argument should extend to mixed two-and-three-partner populations, where we expect the interpolation between the constants computed here and the giant-component regime of higher-degree configuration models. Both infection probabilities now have exact finite-N forms, (1) and (A2), with $q_N - \frac{1}{2} \sim \frac{1}{N}$ and $p_N - \frac{2}{3} \sim \frac{2}{3N}$; whether the tabulated monotone decrease in N holds for every N is left open.

Appendix A Proofs

A.1 The correspondence and the samplers

Proof of Lemma 1 (i) Follow the alternating walk from m_i: along M_1 to $w_{\pi_1(i)}$, then along M_2 to the man m_j with $\pi_1(\sigma(j)) = \pi_1(i)$, that is $j = \sigma^{-1}(i)$, and so on. Each two-step move applies σ^{-1} once to the man-index, so the walk first returns to m_i after k moves, where k is the length of the cycle of σ^{-1} through i, equal to that of the cycle of σ through i. It visits k

men and k women, a $2k$-cycle in G; distinct cycles of σ give disjoint cycles of G, and every cycle of G arises so. Hence $C_{2k}(G) = X_k(\sigma)$.

(ii) The ordered pairs (M_1, M_2) of edge-disjoint perfect matchings with union in $\mathcal{B}_2\mathcal{G}(N)$ number $N!\,D_N$: $N!$ choices for M_1, then D_N for M_2 since the relative permutation must be a derangement. A given G arises from exactly $2^{c(G)}$ such pairs, because on each cycle of G the two alternating edge classes may be assigned to M_1 and M_2 in two ways, independently across cycles. As the Derangement model draws (π_1, σ) uniformly, the induced law on G is proportional to $2^{c(G)}$. □

Proof of Proposition 9 Algorithm 2 is the definition of the Derangement model verbatim; the rejection accepts a Fisher–Yates permutation exactly when it is a derangement, with probability $D_N/N! \to e^{-1}$, so the trial count is geometric with mean e, each trial $O(N)$ [7]. For Algorithm 3, every simple G arises from exactly 4^N of the $(2N)!$ orderings of the stub list L: each of the N men may receive his two partners in either order, and each of the N women may have her two stubs assigned to her two partners in either order. Since 4^N is independent of G, the accepted sample is uniform on $\mathcal{B}_2\mathcal{G}(N)$. A double bond at man m_i occurs with probability $1/(2N-1)$; these N events are asymptotically independent, so the acceptance probability tends to $\exp(-N/(2N)) = e^{-1/2}$. □

A.2 The successor-swap involution and the two constants

For a permutation σ define $\phi(\sigma)$ by exchanging the images of a and b:

$$\phi(\sigma)(a) = \sigma(b), \quad \phi(\sigma)(b) = \sigma(a), \quad \phi(\sigma)(i) = \sigma(i) \ (i \neq a, b).$$

Lemma 10 *ϕ is an involution on S_N. It exchanges the set of permutations placing a, b in one cycle with the set placing them in two, and $c(\phi(\sigma)) = c(\sigma) + 1$ when a, b share a cycle of σ, while $c(\phi(\sigma)) = c(\sigma) - 1$ otherwise.*

Proof That $\phi^2 = \mathrm{id}$ is immediate. If a, b lie in one cycle $\cdots \to a \to s \to \cdots \to b \to t \to \cdots \to a$ with $s = \sigma(a)$, $t = \sigma(b)$, then in $\phi(\sigma)$ the element a maps to t; following from a traverses $a \to t \to \cdots \to a$, closing without meeting b, and the rest forms $b \to s \to \cdots \to b$. One cycle splits into two, c rises by one, and a, b separate. The reverse case is the same computation read backwards: two cycles merge, c falls by one, and a, b unite. □

Proof of Theorem 2, equation (1) By Lemma 10, ϕ is a bijection from same-cycle onto different-cycle permutations, so on all of S_N each class has size $N!/2$. We restrict ϕ to derangements and account exactly for where it leaves that class. Since ϕ alters σ only at a and b, the image $\phi(\sigma)$ fails to be a derangement precisely when it creates a fixed point at a—which needs $\sigma(b) = a$—or at b—which needs $\sigma(a) = b$.

The number of derangements with $\sigma(b) = a$ is not D_{N-1}. Deleting the value a from position b leaves a bijection on the remaining $N-1$ positions with a single forbidden fixed point removed; a direct count (or the identity $\sum_{a'} \#\{\sigma : \sigma(b) = a'\} = D_N$ over the $N-1$ admissible images $a' \neq b$, all equinumerous by relabelling) gives

$$\#\{\sigma \in \mathcal{D}_N : \sigma(b) = a\} = \frac{D_N}{N-1},$$

and likewise for $\sigma(a) = b$. These two exceptional sets are the same-cycle derangements that ϕ maps out of $\mathcal{D}_N$; they intersect in the derangements with both $\sigma(a) = b$ and $\sigma(b) = a$, i.e. containing the transposition $(a\,b)$ as a 2-cycle, of which there are D_{N-2}. On the complement ϕ is an involutive bijection between same-cycle and different-cycle derangements, so those cardinalities are equal there, and the imbalance is carried entirely by the exceptional sets, all of which are same-cycle:

$$D_N^{\text{same}} - D_N^{\text{diff}} = \frac{2D_N}{N-1} - D_{N-2}.$$

With $D_N^{\text{same}} + D_N^{\text{diff}} = D_N$ this gives

$$\frac{D_N^{\text{same}}}{D_N} = \frac{1}{2} + \frac{1}{N-1} - \frac{D_{N-2}}{2D_N},$$

which is (1). Complete enumeration confirms the identity for all $N \le 10$ (Table 1). Equation (2) follows because, with π_1 uniform, a man–woman pair (m_a, w_b) shares an M_1-edge with probability $1/N$ (then trivially co-infected) and otherwise reduces to a two-men (or two-women) pair, contributing q_N. □

Proof of Theorem 3 Reduction. Each $G \in \mathcal{B}_2\mathcal{G}(N)$ arises from $2^{c(G)}$ ordered pairs (π_1, σ), so the uniform law on $\mathcal{B}_2\mathcal{G}(N)$ is the image of the law on pairs that makes π_1 uniform and, independently, gives σ weight proportional to $2^{-c(\sigma)}$. Write $W(\mathcal{A}) = \sum_{\sigma\in\mathcal{A}} 2^{-c(\sigma)}$ and $Z_N = W(\mathcal{D}_N)$, so that for two men $p_N = W(\text{same})/Z_N$.

Exact identity. Let $\mathcal{S}_1$ be the same-cycle derangements with $\sigma(a) = b$ or $\sigma(b) = a$, and $\mathcal{S}_0$ the other same-cycle derangements. A different-cycle derangement τ has $\tau(b) \neq a$ and $\tau(a) \neq b$, so $\phi(\tau)$ is again a derangement; hence ϕ is a bijection between the different-cycle derangements and $\mathcal{S}_0$, and by Lemma 10 it raises the cycle count by one on $\mathcal{S}_0$. Therefore $W(\text{diff}) = \frac{1}{2}\{W(\text{same}) - E_W\}$ with $E_W = W(\mathcal{S}_1)$, and substituting into $W(\text{same}) + W(\text{diff}) = Z_N$ gives

$$p_N = \frac{2}{3} + \frac{E_W}{3Z_N}. \tag{A1}$$

Since $\mathcal{S}_1 \neq \emptyset$ for $N \ge 2$, $p_N > \frac{2}{3}$.

Computing E_W. Let A_N be the weight of the derangements with $\sigma(a) = b$. If the cycle through a has length $\ell \ge 2$, its other $\ell-2$ elements can be chosen and ordered in $(N-2)!/(N-\ell)!$ ways, the cycle contributes weight $\frac{1}{2}$, and the remaining elements form a derangement of weight $Z_{N-\ell}$. With $m = N - \ell$,

$$A_N = \tfrac{1}{2}\,(N-2)! \sum_{m=0}^{N-2} \frac{Z_m}{m!}.$$

The same holds for $\sigma(b) = a$, and the two sets meet in the derangements containing the 2-cycle $(a\,b)$, of weight $\frac{1}{2}Z_{N-2}$. Hence $E_W = 2A_N - \frac{1}{2}Z_{N-2}$, and

$$p_N = \frac{2}{3} + \frac{2A_N - \frac{1}{2}Z_{N-2}}{3Z_N}, \qquad Z_0 = 1,\ Z_1 = 0,\ Z_{n+1} = n\big(Z_n + \tfrac{1}{2}Z_{n-1}\big). \tag{A2}$$

This reproduces the enumerated values in Table 1 exactly for $N \le 10$ and gives the entries for $N \ge 20$.

Asymptotics. By the exponential formula, with each cycle of length at least two weighted by $\frac{1}{2}$,

$$F(z) = \sum_{n\ge0} \frac{Z_n}{n!} z^n = \exp\Big\{\tfrac{1}{2}\Big(\log\frac{1}{1-z} - z\Big)\Big\} = e^{-z/2}(1-z)^{-1/2},$$

and $(1-z)F'(z) = \frac{z}{2}F(z)$ yields the recurrence in (A2). Also $A_N = \frac{1}{2}(N-2)!\,[z^{N-2}]F(z)/(1-z)$. The function F is analytic in $\mathbb{C}\setminus[1,\infty)$, with $F(z) = e^{-1/2}(1-z)^{-1/2} + O\big((1-z)^{1/2}\big)$ as $z \to 1$, so the transfer theorems of singularity analysis [3, Thms VI.1 and VI.3] give

$$[z^n]F(z) = \frac{e^{-1/2}}{\sqrt{\pi n}}\big(1+O(n^{-1})\big), \qquad [z^n]\frac{F(z)}{1-z} = 2e^{-1/2}\sqrt{\frac{n}{\pi}}\big(1+O(n^{-1})\big).$$

Hence

$$\begin{aligned}\frac{A_N}{Z_N} &= \frac{(N-2)!}{2\,N!}\cdot\frac{2e^{-1/2}\sqrt{(N-2)/\pi}}{e^{-1/2}(\pi N)^{-1/2}}\big(1+O(N^{-1})\big)\\ &= \frac{\sqrt{N(N-2)}}{N(N-1)}\big(1+O(N^{-1})\big) = \frac{1}{N}+O(N^{-2}),\end{aligned}$$

while $Z_{N-2}/Z_N = O(N^{-2})$. Thus $E_W/Z_N = 2/N + O(N^{-2})$, and by (A1), $p_N = \frac{2}{3} + \frac{2}{3N} + O(N^{-2})$.

A man and a woman. Since π_1 is uniform and independent of σ under the law above, the argument for (2) applies unchanged and gives $\frac{1}{N} + \frac{N-1}{N}p_N$. □

A.3 Outbreak-size asymptotics, tracing, and the mixed-degree proof

Proof of Propositions 4 and 5 By Lemma 1(i), $C_{2k} = X_k(\sigma)$. For a uniform random permutation $X_k \xrightarrow{d}$ Poisson$(1/k)$ by the method of moments [1]; conditioning on a derangement removes only the $k=1$ term and perturbs fixed $k \geq 2$ by $O(1/N)$, giving the $\mathbb{P}_D$ limit and $\mathbb{E}_D[c(G)] = H_N - 1 + O(1) \sim \log N$ (the conditioning perturbs the finite-N value but not the leading order). The 2^{-c} tilt of the Uniform law is the Ewens sampling formula with parameter $\theta = \frac{1}{2}$, restricted—like the derangement law—to permutations with no one-cycles, since a fixed point of σ would be a self-partnership, excluded from $\mathcal{B}_2\mathcal{G}(N)$. As in the $\theta = 1$ case, this restriction removes only the $k=1$ term and perturbs each fixed-k count by $O(1/N)$, so it leaves the Poisson and Poisson–Dirichlet limits unchanged: the tilt halves each Poisson mean and yields PD$(0, \frac{1}{2})$, while the derangement law ($\theta = 1$) gives PD$(0,1)$ [5, 11]. The largest-fraction constants are those of PD$(0,\theta)$, namely 0.6243 (the Golomb–Dickman constant [12]) at $\theta = 1$ and ≈ 0.758 at $\theta = \frac{1}{2}$ (computed as in Section 4). Under Ewens(θ) the single-cycle probability is of order $N^{-\theta}$, i.e. e/N for $\theta = 1$ (exactly $(N-1)!/D_N$) and $N^{-1/2}$ for $\theta = \frac{1}{2}$. □

Proof of Proposition 6 A connected graph with all degrees in $\{1,2\}$ is a path or a cycle. A path has two degree-1 endpoints, and a, z are the only degree-1 vertices, so there is exactly one path component, with endpoints a and z. It starts in W and ends in M with alternating edges, hence odd length. The remaining components have all degrees 2 and are even cycles. □

Proof of Theorem 7 Take A, B degree-2 women (two men are handled identically; a man and a woman are treated at the end of the proof) and condition on the path half-length T, which uses $T-1$ of the $N-1$ degree-2 women. By exchangeability,

$$\mathbb{P}(A,B \text{ same}) = \sum_{t=1}^{N}\mathbb{P}(T=t)\left[\frac{(t-1)(t-2)}{(N-1)(N-2)} + \frac{(N-t)(N-t-1)}{(N-1)(N-2)}\,p_{\mathrm{res}}(N-t)\right],$$

where $p_{\mathrm{res}}(m)$ is the same-cycle probability in the residual graph on m degree-2 pairs. Conditionally on the path, the residual stub matching is uniform, so the residual is a Uniform-Model

graph and $p_{\mathrm{res}}(m) \to 2/3$ by Theorem 3. It remains to find the limiting law of T/N under the model, which is uniform over *simple* graphs. We obtain it in two steps: first the unconditioned stub-pairing law, then the effect of conditioning on simplicity.

Unconditioned law. Building the path from a's single stub, at step j there are $2N-2j+1$ available opposite stubs, one of which ends the path at z, so the unconditioned stub pairing gives

$$\mathbb{P}_{\mathrm{unc}}(T=t) = \Big(\prod_{j=1}^{t-1} \tfrac{2N-2j}{2N-2j+1}\Big)\tfrac{1}{2N-2t+1}.$$

With $x = t/N$ and $\log(1-\varepsilon) = -\varepsilon + O(\varepsilon^2)$, the product's logarithm is $-\sum_{j=1}^{t-1}(2N-2j+1)^{-1} + O(N^{-1}) \to \frac{1}{2}\log(1-x)$ uniformly on $x \le 1-\delta$, so under the unconditioned law T/N has limiting density $f(x) = \frac{1}{2}(1-x)^{-1/2}$ (one checks $\int_0^1 f = 1$, $\int_0^1 xf = 2/3$).

Conditioning on simplicity. The stated model is the stub pairing conditioned to be simple. Given $T = t$, the path itself is automatically simple, so the conditioning acts only on the residual 2-regular bipartite graph on $m = N-t$ degree-2 pairs; writing $s(m)$ for the probability that this residual is simple,

$$\mathbb{P}(T=t) = \frac{\mathbb{P}_{\mathrm{unc}}(T=t)\, s(N-t)}{\sum_{t'} \mathbb{P}_{\mathrm{unc}}(T=t')\, s(N-t')}.$$

The residual is a configuration-model 2-regular bipartite graph, for which $s(m) \to e^{-1/2}$ as $m \to \infty$ (the same computation as in Proposition 9, giving the acceptance probability of the Uniform-Model sampler), while $s(0) = 1$ and $s(1) = 0$—the case $t = N-1$ is forbidden, since a single leftover degree-2 pair would need a double edge. Write $\alpha = e^{-1/2}$ and fix $\delta \in (0,1)$ and $\varepsilon \in (0, \alpha/2)$. Because $0 \le s \le 1$ and $s(m) \to \alpha$, for all large N we have $|s(N-t) - \alpha| < \varepsilon$ for every t in the bulk $t \le (1-\delta)N$, where $N - t \ge \delta N$. Let g be bounded and continuous on $[0,1]$. Splitting both sums in

$$\mathbb{E}[g(T/N)] = \frac{\sum_t g(t/N)\, \mathbb{P}_{\mathrm{unc}}(T=t)\, s(N-t)}{\sum_t \mathbb{P}_{\mathrm{unc}}(T=t)\, s(N-t)}$$

into the bulk and the boundary $t > (1-\delta)N$, each bulk sum equals $\alpha + O(\varepsilon)$ times the corresponding unconditioned sum, and each boundary sum is at most $\|g\|_\infty\, \mathbb{P}_{\mathrm{unc}}(T > (1-\delta)N)$ in absolute value. By the unconditioned limit, $\mathbb{P}_{\mathrm{unc}}(T > (1-\delta)N) \to \int_{1-\delta}^1 f = \sqrt{\delta}$ and $\sum_{t \le (1-\delta)N} g(t/N)\mathbb{P}_{\mathrm{unc}}(T=t) \to \int_0^{1-\delta} gf$. Hence

$$\mathbb{E}[g(T/N)] = \frac{\alpha \int_0^{1-\delta} gf + O(\varepsilon) + O(\sqrt{\delta})}{\alpha \int_0^{1-\delta} f + O(\varepsilon) + O(\sqrt{\delta})} + o(1),$$

and letting $N \to \infty$, then $\varepsilon \to 0$, then $\delta \to 0$ gives $\mathbb{E}[g(T/N)] \to \int_0^1 gf$. So under the simple-graph law T/N still converges in distribution to the density $f(x) = \frac{1}{2}(1-x)^{-1/2}$. The conditioning reweights only the boundary region, whose mass vanishes as $\delta \to 0$; in particular the forbidden value $t = N-1$ carries no limiting mass.

Conclusion. As $N \to \infty$ with $t = xN$, the two combinatorial factors tend to x^2 and $(1-x)^2$, $p_{\mathrm{res}} \to 2/3$, and T/N has density f, so by bounded convergence

$$\mathbb{P}(A, B \text{ same}) \to \int_0^1 \Big[x^2 + \tfrac{2}{3}(1-x)^2\Big] \frac{dx}{2\sqrt{1-x}} = \frac{8}{15} + \frac{2}{15} = \frac{2}{3},$$

the integrals evaluated by $u = 1-x$; the tail $x > 1-\delta$ contributes at most $\int_{1-\delta}^1 f = \sqrt{\delta} \to 0$, so the limit holds on all of $(0,1)$.

For a man and a woman the same decomposition applies. The probability that both lie on the path given $T = t$ again tends to x^2, the probability that both lie in the residual tends to $(1-x)^2$, and in the residual their same-cycle probability is the man-and-woman probability of Theorem 3, which also tends to 2/3. The finite-N factors differ from the two-women case by $O(1/N)$ terms, including the $O(1/N)$ chance that the two are partners, and the limit 2/3 is unchanged. □

Proof of Proposition 8 The index case lies in one outbreak, which by Section 2 and Proposition 6 is a single cycle or a single path. Round 1 interviews the index case and identifies its partners, the people at distance one. In round $r \geq 2$ each active frontier sits at distance $r-1$, and its interview identifies the next person at distance r. A frontier only ever moves to a person not yet identified, so no one is a frontier twice, and each person is interviewed exactly once. After r rounds the identified set is exactly the set of people within distance r of the index case.

Cycle case. Every person on a cycle of L people is within distance $\lfloor L/2 \rfloor$ of the index case, and some person attains this, so every member is identified after $\lfloor L/2 \rfloor$ rounds. On an even cycle both frontiers reach the opposite person at the same distance; one frontier identifies that person and the other, finding them already identified, stops. On an odd cycle the two frontiers end on adjacent people, each of whom names only identified people. In both cases the procedure halts at round $\lfloor L/2 \rfloor + 1$. One-directional tracing identifies one new person per round and needs $L-1$ rounds.

Path case. Let the index case be at distances d_1, d_2 from the two endpoints, $d_1 + d_2 = L-1$; if it is itself an endpoint, it names one partner and there is a single frontier. The frontier towards the endpoint at distance d_i identifies that endpoint in round d_i and stops in the next round, when the endpoint names no further partner. Every member is therefore identified after $\max(d_1, d_2)$ rounds, and the procedure halts one round later. The value $\max(d_1, d_2)$ equals $\lceil (L-1)/2 \rceil$ when the index case is central and $L-1$ when it is an endpoint. Because each frontier stops on its own, a frontier that reaches an endpoint does not stop the other. □

References

[1] Arratia, R., Barbour, A.D., Tavaré, S.: Logarithmic Combinatorial Structures: A Probabilistic Approach. EMS Monographs in Mathematics. European Mathematical Society (2003)

[2] Bollobás, B.: Random Graphs, 2nd edn. Cambridge University Press (2001)

[3] Flajolet, P., Sedgewick, R.: Analytic Combinatorics. Cambridge University Press (2009)

[4] Holme, P., Saramäki, J.: Temporal networks. Phys. Rep. **519**, 97–125 (2012)

[5] Kingman, J.F.C.: Random partitions in population genetics. Proc. R. Soc. Lond. A **361**, 1–20 (1978)

[6] Kretzschmar, M., Morris, M.: Measures of concurrency in networks and the spread of infectious disease. Math. Biosci. **133**, 165–195 (1996)

[7] Martínez, C., Panholzer, A., Prodinger, H.: Generating random derangements. In: Proc. ANALCO 2008, pp. 234–240. SIAM (2008)

[8] Miller, J.C., Slim, A.C., Volz, E.J.: Edge-based compartmental modelling for infectious disease spread. J. R. Soc. Interface **9**, 890–906 (2012)

[9] Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. Random Struct. Algorithms **6**, 161–180 (1995)

[10] Newman, M.E.J.: Spread of epidemic disease on networks. Phys. Rev. E **66**, 016128 (2002)

[11] Pitman, J., Yor, M.: The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. Ann. Probab. **25**, 855–900 (1997)

[12] Shepp, L.A., Lloyd, S.P.: Ordered cycle lengths in a random permutation. Trans. Am. Math. Soc. **121**, 340–357 (1966)

[13] Wormald, N.C.: Models of random regular graphs. In: Surveys in Combinatorics 1999. LMS Lecture Note Series, vol. 267, pp. 239–298. Cambridge University Press (1999)