Short cycle distribution in random regular graphs recursively generated by pegging

Pu Gao and Nicholas Wormald*
Department of Combinatorics and Optimization
University of Waterloo
p3gao@math.uwaterloo.ca, nwormald@math.uwaterloo.ca

Abstract

We introduce a new method of generating random d-regular graphs by repeatedly applying an operation called pegging. The pegging operation is abstracted from a type of basic operation applied in a type of peer-to-peer network called the SWAN network. We prove that for the resulting graphs, the limiting joint distribution of the numbers of short cycles is independent Poisson. We also use coupling to bound the rate at which the distribution approaches its limit. The coupling involves two different, though quite similar, Markov chains that are not time-homogeneous.

1 Introduction

Random regular graphs have recently arisen in a peer-to-peer ad-hoc network, called the SWAN network, introduced by Bourassa and Holt [2]. In the SWAN network, clients arrive and leave randomly. The network maintains the underlying topology as a d-regular graph by using an operation called "clothespinning" (for arriving clients), and its reverse (for clients leaving). Clothespinning consists of deleting two links of the network and joining all four incident nodes to the newly arrived node, which can also be thought of pictorially as pinning the two links together at their midpoints using the new node. Occasionally some other adjustments are also required to repair the network when these operations cannot cope. For instance, when a client leaves and its neighbours already have many links between them, the reverse of clothespinning might cause multiple links between two nodes.

After a sequence of such operations applied randomly, the result is a random graph whose distribution is not fully understood. Bourassa and Holt found experimentally that it has good connectivity and diameter properties. More recently, Cooper, Dyer and Greenhill [3] defined a Markov chain on d-regular graphs with randomised size to model (a simplified version of) the SWAN network. The moves of the Markov chain are by clothespinning or the reverse. They showed that,

*Research supported by the Canadian Research Chairs Program and NSERC

restricted to the times when the network has a certain size, the stationary distribution is uniform, and they bounded the mixing time of the chain.

In this paper, we introduce the *pegging algorithm* to generate random d-regular graphs for constant d. The pegging algorithm simply repeats clothespinning (which we call *pegging*) operations, without performing the reverse. We will focus mainly on even d, in which case a pegging operation can be visualised as binding the middles of $d/2$ nonadjacent edges together using a new vertex. Thus the size of the graph increases linearly with the number of operations. This gives an extreme version of the SWAN network, in which no client ever leaves the network. Since the analysis of [3] does not apply if the network undergoes net long-term growth, by studying this extreme case we hope to gain knowledge of which properties of the random SWAN network are not sensitive to long-term growth.

We will study the joint distribution of short cycle counts in the random d-regular graph generated by pegging, for any $d \geq 3$. This is the most interesting small subgraph feature of most models of random regular graphs, since they look locally like trees. We will find that the distribution is quantitatively different from, but qualitatively very similar to, the model of random d-regular graphs with uniform distribution, for which the short cycle joint distribution is independent Poisson, first derived independently by Bollobás [1] and the second author [10]. Besides using the method of moments to obtain the limiting distribution, we use an application of coupling to bound the rate of convergence to the limiting distribution. Coupling has often been used in combinatorics in a Markov chain setting, where two copies of the same time-homogeneous Markov chain are coupled. In our application, we couple two different Markov chains, which has some similarity to the introductory application of coupling in [5] concerning the Poisson distribution.

The pegging algorithm provides a natural and very efficient way of generating random regular graphs. How to generate random d-regular graphs with the uniform probability distribution is a topic that has attracted considerable attention. It is only known how to do this efficiently when d is fairly small [6]. On the other hand, there are natural algorithmic approaches to generating random regular graphs, such as the d-process [9], and d-star-process [7, 8]. These do not generate d-regular graphs uniformly, but asymptotically almost surely terminate with a graph that is d-regular.

The pegging algorithm is defined in general in the next section, and our main results are stated there. The rest of the paper is mainly devoted to proofs. Some basic results about the moments of the short cycle counts for the case $d = 4$ are proved in Section 3, rate of convergence results are proved in Section 4, and the arguments are generalised to arbitrary $d \geq 3$ in Section 5. The final section has further discussion and a conjecture on the similarity between the graphs obtained by this model and the uniformly random d-regular graphs.

2 Main Results

In this section we define the pegging algorithm and give our main results. We define the *pegging operation* on a d-regular graph, where d is even, as follows.

Pegging Operation for Even d

Input: a d-regular graph G, where d is even.

Choose a set E_1 of $d/2$ pairwise non-adjacent edges in $E(G)$ u.a.r.

Let $\{u_1, u_2, \ldots, u_d\}$ denote the vertices incident with edges in E_1.
$V(H) := V(G) \cup \{u\}$, where u is a new vertex.
$E(H) := (E(G) \setminus E_1) \cup \{uu_1, uu_2, uu_3, \ldots, uu_d\}$.
Output: H.

The newly introduced vertex u is called the *peg vertex*, and we say that the edges deleted are *pegged*. Figure 1 illustrates the pegging operation with $d = 4$.

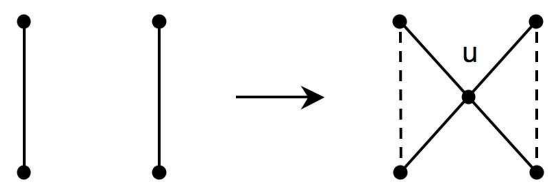

Figure 1: *Pegging operation when* $d = 4$

It is immediate that the graph outputted by the pegging operation is also d-regular. (There is some question as to whether the operation is well defined: does every nonempty d-regular graph have a set of $d/2$ independent edges when d is even? It is easy to see that the answer is yes, by a greedy algorithm.)

The *pegging algorithm* starts from a nonempty d-regular graph G_0 ($d \geq 4$ and even), for example, K_{d+1}, and repeatedly applies pegging operations. For $t > 0$, G_t is defined inductively to be the graph resulting when the pegging operation is applied to G_{t-1}. Clearly, G_t contains $n_t := n_0 + t$ vertices. We denote the resulting random graph process $(G_0, G_1, \ldots)$ by $\mathcal{P}(G_0)$, or $\mathcal{P}(G_0, d)$ if we wish to specify the degree d of the vertices of G_0.

The SWAN network used in practice has $d = 4$. For this reason and simplicity of notation, we treat this case separately. Here, the algorithm starts from a 4-regular graph G_0 with n_0 vertices. At each step, it randomly chooses two non-adjacent edges, deletes them, and joins a newly created vertex to the four end vertices of the deleted edges. Thus G_t contains $2n_t$ edges.

Let $Y_{t,k}$ denote the number of k-cycles in G_t. Note that initially, the number of triangles might be as big as $2n_0$. However, as we will show later, in such an extreme case, $\mathbf{E}Y_{t,3}$ will decrease quickly in the early stage of the algorithm.

For $k \geq 3$ we define

$$\mu_k = \frac{3^k - 9}{2k}. \tag{2.1}$$

Theorem 2.1 *Let G_0 and $k \geq 3$ be fixed. Then in $\mathcal{P}(G_0, 4)$,*

$$\mathbf{E}Y_{t,k} = \mu_k + O\left(n_t^{-1}\right).$$

Moreover, in $\mathcal{P}(G_0, 4)$ the joint distribution of $Y_{t,3}, \ldots, Y_{t,k}$ is asymptotically that of independent Poisson variables with means $\mu_3, \ldots, \mu_k$.

Let σ and π be probability distributions on the same countable set (state space) $\mathcal{S}$. The *total variation distance* between σ and π is defined as

$$d_{TV}(\sigma, \pi) = \max_{A \subset \mathcal{S}} \{\sigma(A) - \pi(A)\}. \tag{2.2}$$

Equivalently,

$$d_{TV}(\sigma, \pi) = \frac{1}{2} \sum_{x \in \mathcal{S}} |\ \sigma(x) - \pi(x)\ |.$$

Let $0 < \epsilon < 1$. The standard definition of the mixing time $\tau(\epsilon)$ of a Markov chain with state space $\mathcal{S}$ is the minimum time t, such that after t steps, the total variation distance between P_x^t, the distribution at time t starting from state x, and the stationary distribution π, is at most ϵ. Formally,

$$\tau(\epsilon) = \max_{x \in \mathcal{S}} \min\{T : d_{TV}\left(P_x^t, \pi\right) \leq \epsilon \text{ for all } t \geq T\}.$$

In practice, for mixing time results one chooses $\epsilon = 1/4$ say, and obtains results such as "the mixing time is $O(n \log n)$," referring to structures of size n and for the fixed value of ϵ. This makes sense because of the fact that for a fixed time-homogeneous Markov chain, $\tau(\epsilon^k)$ is well estimated by $k\tau(\epsilon)$. In other words, this sort of Markov chain usually has mixing time as that is logarithmic in $1/\epsilon$. So in such cases it is more interesting to study the mixing time as a function of n rather than as a function of the variable ϵ. Slow mixing, such as exponential mixing time, usually refers to the mixing time as an exponential function of n, rather than $1/\epsilon$. However, there are several differences in our case. Firstly, this random process is not time-homogeneous, since the transition probability from $Y_{t,k}$ to $Y_{t+1,k}$ depends on the time t. So we might not (and would not expect to) get logarithmic "mixing time," as is usual for a finite state Markov chain when the mixing time is viewed as a function of ϵ. The fact that the size of our structures is growing is a further complicating factor. Indeed, the random process $(Y_{t,k})_{t\geq 0}$, for some constant k, is not a Markov chain, since the distribution of $Y_{t,k}$ depends not only on $Y_{t-1,k}$, but also on the underlying graph G_{t-1}, and as a result is not independent of $Y_{t-2,k}$ given $Y_{t-1,k}$. Instead, we wish to consider the total variation distance between the distribution of the random variable $Y_{t,k}$ and the limiting distribution of $Y_{t,k}$ (if it exists). With these considerations in mind, we define ϵ-mixing time in a general way as follows. Let $(\sigma_t)_{t\geq 0}$ be a sequence of distributions which converge to the distribution π_k^*. The *ϵ-mixing time* of $(\sigma_t)_{t\geq 0}$ is

$$\tau_\epsilon^*\left((\sigma_t)_{t\geq 0}\right) = \min\{T \geq 0 :\ d_{TV}\left(\sigma_t, \pi_k^*\right) \leq \epsilon \text{ for all } t \geq T\ \}. \tag{2.3}$$

We now focus on a particular sequence of distributions. For any fixed k, let $\sigma_{t,k}$ denote the joint distribution of $Y_{t,3}, \ldots, Y_{t,k}$. Our second main result is to show that the ϵ-mixing time of $(\sigma_{t,k})_{t\geq 0}$ is $O(1/\epsilon)$.

Theorem 2.2 *For fixed G_0 and $k \geq 3$, the ϵ-mixing time of the sequence of short cycle joint distributions in $\mathcal{P}(G_0, 4)$ satisfies*

$$\tau_\epsilon^*\left((\sigma_{t,k})_{t\geq 0}\right) = O\left(\epsilon^{-1}\right).$$

We also extend these results to arbitrary integers $d \geq 3$. First, the pegging operation for d-regular graphs was only defined for d even, and we need to extend it to the case of d odd. This is as follows, illustrated in Figure 2 for $d = 3$.

Pegging Operation for Odd d
Input: a d-regular graph G, where d is odd.

1. Let $c := \lfloor d/2 \rfloor$ and choose a set $E_1 = \{u_1u_2, u_3u_4, \ldots, u_{2c-1}u_{2c}\}$ of c pariwise non-adjacent edges in $E(G)$ u.a.r., and another set $E_2 = \{u_{2c+1}u_{2c+2}, \ldots, u_{4c-1}u_{4c}\}$ of c pairwise non-adjacent edges in $E(G) \setminus E_1$ u.a.r.
2. $G := (G \setminus (E_1 \cup E_2)) \cup \{u, v\} \cup E_3 \cup \{uv\}$, where u and v are new vertices added to $V(G)$, and $E_3 = \{uu_1, \ldots, uu_{2c}, vu_{2c+1}, \ldots, vu_{4c}\}$.
3. Output: G.

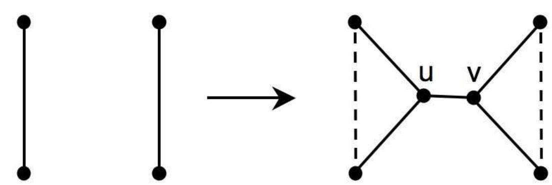

Figure 2: *Pegging operation when* $d = 3$

The definition of $\mathcal{P}(G_0, d)$ is now extended to odd integers d in the obvious way. Let $Y_{t,d,k}$ be the number of k-cycles in $G_t \in \mathcal{P}(G_0, d)$ and $\sigma_{t,d,k}$ be the joint distribution of $Y_{t,d,3}, Y_{t,d,4}, \ldots, Y_{t,d,k}$. We derive the following generalised result.

Theorem 2.3 *For fixed $k \geq 3$, any integer $d \geq 3$, and fixed initial d-regular graph G_0,*

$$\mathbf{E}Y_{t,d,k} = \frac{(d-1)^k - (d-1)^2}{2k} + O\left(n_t^{-1}\right).$$

Moreover

(i) $Y_{t,d,3}, Y_{t,d,4}, \ldots, Y_{t,d,k}$ have a limiting joint distribution equal to that of independent Poisson variables with means $\mu_{d,3}$, $\mu_{d,4}, \ldots, \mu_{d,k}$, where $\mu_{d,i} = ((d-1)^i - (d-1)^2)/(2i)$ for any fixed integer $i \geq 3$,

(ii) the ϵ-mixing time satisfies $\tau^*_\epsilon\left((\sigma_{t,d,k})_{t\geq 0}\right) = O\left(\epsilon^{-1}\right)$.

Couplings are often used to estimate the mixing time of a Markov chain, by bounding the distance between the distribution of the Markov chain with a given initial state, and a copy of the same chain beginning with the stationary distribution. However, we will need to estimate the total variation distance of the distributions derived from two different random processes, not just two copies of the same Markov chain.

A coupling of two random variables X_1 and X_2 (not necessarily defined on the same probability space) is a pair (Y_1, Y_2) of variables defined on a probability space such that the marginal distribution of Y_i is the distribution of X_i (for $i = 1$ and 2). With only a slight abuse of notation, this random pair is sometimes written as (X_1, X_2), and the coupling can be described as a construction of copies of X_1 and X_2 in a common probability space. We require that X_1 and X_2 have the same state space (range). Lindvall [5, pp. 9–12] gave a more elaborate definition of coupling that is equivalent for our purposes, and gave a corresponding general coupling lemma which we may state as follows.

Lemma 2.1 *Let (X_1, X_2) be a coupling and let σ_i denote the distribution of X_i. Then*

$$d_{TV}(\sigma_1, \sigma_2) \le \mathbf{P}(X_1 \ne X_2).$$

If $(X_t)_{t\ge0}$ and $(Y_t)_{t\ge0}$ are two random processes in the same state space, a random process $\big((X_t, Y_t)\big)_{t\ge0}$ is a coupling of the two processes if (X_t, Y_t) is a coupling of X_t and Y_t for all $t \ge 0$.

3 Factorial moments of short cycle counts

We begin with a simple technical lemma that will be used several times.

Lemma 3.1 *Let $(a_n)_{n\ge1}$ be a sequence of nonnegative real numbers, and let $c > 0$, and $p \ne c + 1$, be constants. If*

$$a_{n+1} = \left(1 - \frac{c}{n} + O(n^{-2})\right) a_n + O(n^{-p}),$$

then $a_n = O(n^{\delta})$ for all $n \ge 1$, where $\delta = \max\{-c, -p + 1\}$.

Proof. We have

$$a_{n+1} = \exp\left(-\frac{c}{n} + O(n^{-2})\right) a_n + O(n^{-p}).$$

Iterating this gives

$$\begin{aligned} a_n &= a_1 \exp\left(-\sum_{i=1}^{n-1} \frac{c}{i} + O(i^{-2})\right) + \sum_{i=1}^{n-1} \exp\left(-\sum_{j=i+1}^{n-1} \frac{c}{j} + O(j^{-2})\right) O(i^{-p}) \\ &= a_1 \exp\left(-c\log n + O(1)\right) + \sum_{i=1}^{n-1} \exp\left(-c\log(n/i) + O(i^{-1})\right) O(i^{-p}) \\ &= O(n^{-c}) + \sum_{i=1}^{n-1} \left(\frac{i^c}{n^c}\right) O(i^{-p}) \\ &= O(n^{\delta}). \quad \blacksquare \end{aligned}$$

To show that $Y_{t,3}, Y_{t,4}, \ldots, Y_{t,k}$ are asymptotically independent Poisson random variables, it is enough, by the method of moments, to check that their moments are asymptotic to those of independent Poisson random variables with fixed means. We will first prove several lemmas, which give the first and higher moments of $Y_{t,k}$ for any fixed k.

The notations $O()$ occurring in the following lemma and subsequently are defined as follows: for each occurrence of the notation $O(f)$, where f is a function of t and $G_0, \ldots, G_t$, there exists a constant C, depending only on n_0 and k, such that the term denoted $O(f)$ is at most $C|f|$. In particular, this is for all t in the following.

Lemma 3.2 *For $k \geq 3$,*

$$\mathbf{E}Y_{t,k} = \mu_k + O\left(n_t^{-1}\right).$$

Proof. Our analysis is based on the underlying graph G_t produced in step t. For step $t+1$, two non-adjacent edges e_1 and e_2 are chosen in the pegging operation. There are $2n_t$ choices for e_1, and then $2n_t - 7$ choices for e_2 to be non-adjacent to e_1. So the number of ways to choose an ordered pair (e_1, e_2) is $2n_t\,(2n_t - 7)$, and hence the total number of ways to do a pegging operation in step $t+1$ is

$$N_t = \frac{2n_t\,(2n_t - 7)}{2} = n_t\,(2n_t - 7). \tag{3.1}$$

We prove by induction that, for k fixed, $\mathbf{E}Y_{t,k} = \mu_k + O(n_t^{-1})$ for all $t \geq 0$. Note that in the inductive hypothesis, the notation $O()$ implicitly contains a constant that depends on k. For the base case, we consider $k = 3$.

For this and many similar calculations, to estimate the expected change in a variable counting copies of some subgraph, we consider the number of copies of the subgraph created in one step, and separately subtract the number destroyed. In particular, if a subgraph contains either of the pegged edges, it is destroyed.

We need to consider the creation of a new triangle. Given an edge e of G_t not in a triangle, a new triangle is created containing e if and only if the two pegged edges e_1 and e_2 are both adjacent to e. Of course, in view of the definition of pegging, they must be incident with different end-vertices of e. Since G_t is 4-regular, the number of ways to choose such e_1 and e_2 is precisely 9. Note also that only one edge of a given new triangle was already present in G_t. It follows that the expected number of new triangles created is at least $9\,(2n_t - 3Y_{t,3})\,/N_t$, with N_t given above. An obvious upper bound is $9 \cdot 2n_t/N_t$.

To destroy a triangle, either e_1 or e_2 must lie in the triangle, and there are of course $2n_t - 7$ choices for another edge to be pegged. So for each triangle in G_t, the probability that it is destroyed is $3\,(2n_t - 7)\,/N_t$. Thus, the expected number of existing triangles destroyed is $3Y_{t,3}\,(2n_t - 7)\,/N_t = 3Y_{t,3}/n_t$.

It follows that the expected value of $Y_{t+1,3} - Y_{t,3}$, given G_t, satisfies

$$\frac{18}{2n_t - 7} - \frac{3Y_{t,3}}{n_t}\left(1 + \frac{9}{2n_t - 7}\right) \leq \mathbf{E}\left(Y_{t+1,3} - Y_{t,3} \mid G_t\right) \leq \frac{18}{2n_t - 7} - \frac{3Y_{t,3}}{n_t}.$$

Thus

$$\mathbf{E}\left(Y_{t+1,3} \mid G_t\right) = \left(1 - \frac{3 + O(n_t^{-1})}{n_t}\right) Y_{t,3} + \frac{9}{n_t} + O\left(n_t^{-2}\right). \tag{3.2}$$

Taking expectation of both sides and applying the Tower Property of conditional expectations, we obtain

$$\mathbf{E}\, Y_{t+1,3} = \left(1 - \frac{3 + O(n_t^{-1})}{n_t}\right) \mathbf{E}\, Y_{t,3} + \frac{9}{n_t} + O\left(n_t^{-2}\right)$$

where the $O()$ terms are to be read as stated prior to the lemma statement, and in particular they are independent of $G_0, \ldots, G_t$. Putting $\lambda_{t,3} = \mathbf{E}(Y_{t,3} - 3)$ gives

$$\lambda_{t+1,3} = \left(1 - \frac{3}{n_t}\right) \lambda_{t,3} + O\left(\frac{1 + \lambda_{t,3}}{n_t^2}\right).$$

Applying Lemma 3.1, we have $\lambda_{t,3} = O(n_t^{-1})$ and hence $\mathbf{E}Y_{t,3} = 3 + O(n_t^{-1})$. This establishes the base case of the induction, i.e. for $k = 3$.

Now assume the inductive hypothesis is true of all integers smaller than k. There are two ways that one pegging operation can create a k-cycle. The first way occurs when two non-adjacent edges are pegged such that some $(k-1)$-cycle contains exactly one of them. The expected number of k-cycles created in this way is

$$\frac{(k-1)Y_{t,k-1}(2n_t - k - 3)}{N_t}.$$

The second way occurs when the two end edges of a k-path are chosen for pegging. The number of paths of length k in G_t starting from a fixed vertex v is at most $4 \cdot 3^{k-1}$, so the number of k-paths in G_t is at most $2 \cdot 3^{k-1} n_t$. This counts all walks of length k that do not immediately retrace a step, so is an over-count due to repeated vertices in the cases that the walk contains at least one cycle. There are $\sum_{i=1}^{k} Y_{t,i}$ cycles of size at most k in G_t. If we pick an edge in each of those cycles and exclude all walks containing the selected edges, we have an upper bound on the number of walks counted that are not paths. The number of selected edges is at most $\sum_{i=1}^{k} Y_{t,i}$, and each edge is contained in at most $k3^{k-1}$ walks. So G_t contains at least $2 \cdot 3^{k-1} n_t - k3^{k-1} \sum_{i=1}^{k} Y_{t,i}$ different k-paths. Thus the expected number of k-cycles created in this way, given G_t, is

$$\frac{2 \cdot 3^{k-1} n_t + O\left(\sum_{i=1}^{k} Y_{t,i}\right)}{N_t}.$$

Note that $N_t = 2n_t^2(1 + O(n_t^{-1}))$ and, by induction, $\mathbf{E}\, Y_{t,i} = O(1)$ for $i < k$. It thus follows from the two cases above that the expected number of new k-cycles created in going from G_t to G_{t+1} is

$$\frac{3^{k-1} + (k-1)\mathbf{E}Y_{t,k-1}}{n_t} + O\left(\frac{1 + \mathbf{E}Y_{t,k}}{n_t^2}\right).$$

Similar to the case of $k = 3$, given G_t, a given k-cycle is destroyed if and only if some edge contained in the k-cycle is pegged. The probability for that to occur is $k(2n_t - 7)/N_t - k(k-3)/(2N_t)$, where $k(k-3)/(2N_t)$ accounts for the over-counting in the first term when both pegged edges are in the k-cycle. Hence the expected number of k-cycles destroyed is $kY_{t,k}/n_t + O(Y_{t,k}/n_t^2)$. Combining the creation and destruction cases, we find that

$$\mathbf{E}Y_{t+1,k} - \mathbf{E}Y_{t,k} = \frac{3^{k-1} + (k-1)\mathbf{E}Y_{t,k-1} - k\mathbf{E}Y_{t,k}}{n_t} + O\left(\frac{1 + \mathbf{E}Y_{t,k}}{n_t^2}\right) \tag{3.3}$$

By induction, $\mathbf{E}Y_{t,k-1} = \mu_{k-1} + O(n_t^{-1})$, so

$$\begin{aligned}
\mathbf{E}Y_{t+1,k} &= \left(1 - \frac{k}{n_t} + O(n_t^{-2})\right) \mathbf{E}Y_{t,k} + \frac{3^{k-1}}{n_t} + \frac{(k-1)\,\mathbf{E}Y_{t,k-1}}{n_t} + O(n_t^{-2}) \\
&= \left(1 - \frac{k}{n_t} + O(n_t^{-2})\right) \mathbf{E}Y_{t,k} + \frac{3^{k-1}}{n_t} + \frac{k-1}{n_t}\left(\frac{3^{k-1}-9}{2} + O(n_t^{-1})\right) + O(n_t^{-2}) \\
&= \left(1 - \frac{k}{n_t} + O(n_t^{-2})\right) \mathbf{E}Y_{t,k} + \frac{k\mu_k}{n_t} + O(n_t^{-2}).
\end{aligned}$$

Letting $\lambda_{t,k} = \mathbf{E}Y_{t,k} - \mu_k$ gives

$$\lambda_{t+1,k} = \left(1 - \frac{k}{n_t}\right) \lambda_{t,k} + O\left(\frac{1+\lambda_{t,k}}{n_t^2}\right),$$

and so by Lemma 3.1, we have $\lambda_{t,k} = O(n_t^{-1})$, and hence $\mathbf{E}Y_{t,k} = \mu_k + O(n_t^{-1})$ for any constant $k \geq 3$. Lemma 3.2 follows. ∎

Define $\Psi(i,r)$ to be the set of graphs with i vertices, minimum degree at least 2, and excess r, where the excess of a graph is the number of edges minus the number of vertices. Define $W_{t,i,r}$ to be the number of subgraphs of G_t in $\Psi(i,r)$.

For the following lemma the constants implicit in $O()$ depend on i.

Lemma 3.3 *For fixed $i > 0$ and $r \geq 0$,*

$$\mathbf{E}W_{t,i,r} = O(n_t^{-r}).$$

Proof. We prove by induction on r and i. Any graph in $\Psi(i,r)$ contains at least one cycle since it has minimum degree at least 2. Thus $\Psi(i,r) = \emptyset$ for $i = 1,2$. The base case is $r = 0$ and $i = 3$. So $H \in \Psi(3,0)$ is a triangle. Hence The base case holds by Lemma 3.2.

Assume $W_{t,i-1,0} = O(1)$ for any $i \geq 4$. Let H be any graph in $\Psi(i,0)$. Since the excess of H is a 0, every component of H is a cycle.

We bound the expected number of subgraphs in $\Psi(i,0)$ created when going from G_t to G_{t+1}. We omit some simple details that are virtually the same as those in the proof of Lemma 3.2. We also note that for any fixed i, $|\Psi(i,0)| < \infty$, namely, there are only finitely many graphs in $\Psi(i,r)$.

As in the proof of Lemma 3.2, by linearity of expectation we can deal separately with the expected numbers of subgraphs created and destroyed in a single step. Any new subgraph, which is a union of cycles, in $\Psi(i,0)$ can be created either by pegging an edge of a short cycle with any other edge (to make a cycle with length increased by 1), or by pegging together the end edges of a short path.

Case 1: One edge in a graph H' in $\Psi(i-1,0)$ is pegged. (Hence one cycle in H' get longer.) Since each $H' \in \Psi(i-1,0)$ contains $i-1$ vertices, thus $i-1$ edges, there are $O(W_{t,i-1,0})$ ways to choose an edge contained in $\Psi(i-1,0)$ and at most $2n_t$ choices for the other edge to be pegged. The expected number of $\Psi(i,0)$ arising this way is $O(W_{t,i-1,0}/n_t)$. By the inductive hypothesis that $\mathbf{E}W_{t,i-1,0} = O(1)$, the total expected number of graphs created in $\Psi(i,0)$ due to this case is $O(1/n_t)$.

Case 2: A new cycle of size at most i is created by pegging two edges within distance i, which, together with a graph in $\Psi(i',0)$ for some $i' < i$ will form a new graph in $\Psi(i,0)$. There are $O(n_t)$ paths of length at most i. So the expected number of $\Psi(i,0)$ created this way is at most $O(W_{t,i',0}/n_t)$. The number of choices of $i' < i$ is bounded. So summing over all possible value of i', and again by induction, the total contribution from this case is $O(1/n_t)$.

Since subgraphs are destroyed if they contain a pegged edge, the expected number of graphs in $\Psi(i,0)$ destroyed is at least $W_{t,i,0}/n_t$.

Putting it all together, we have

$$\mathbf{E}W_{t+1,i,0} \le \left(1 - \frac{1}{n_t}\right)\mathbf{E}W_{t,i,0} + O(n_t^{-1}).$$

and hence $\mathbf{E}W_{t,i,0} = O(1)$ for by Lemma 3.1. By induction, we obtain $\mathbf{E}W_{t,i,0} = O(1)$ for any $i \ge 3$.

Next we fix any $r \ge 1$, and $i \ge 3$, and assume that $\mathbf{E}W_{t,j,r-1} = O(n_t^{-(r-1)})$ for any $j \ge 3$ and $\mathbf{E}W_{t,j,r} = O(n_t^{-r})$ for any $j \le i$.

We use the same procedure to prove $\mathbf{E}W_{t,i,r} = O(n_t^{-r})$. Consider the expected number of subgraphs in $\Psi(i,r)$ created in going from G_t to G_{t+1}, treating separate cases for creation as above.

Case 1: Similar to the first case above, a subgraph in $\Psi(i,r)$ arises from a subgraph in $\Psi(i-1,r)$, so by induction, we have the total contribution as $O(\mathbf{E}W_{t,i-1,r}/n_t) = O(n_t^{-(r+1)})$.

Case 2: One subcase is that the end edges of a path of length at most i are pegged, which will convert some graph in $\Psi(i',r)$ to one in $\Psi(i,r)$, where $i' < i$. The only other case is that the edges pegged are both within distance i of some graph in $\Psi(j',r-1)$, where $j' < i$. For any fixed subgraph of G_t in $\Psi(j',r-1)$, there are only finite many choices for two such edges to be pegged. So the expected increase in this case will be a sum of a finite number of terms of the form $O(W_{t,i',r}/n_t) + O(W_{t,j',r-1}/n_t^2)$. By induction, $\mathbf{E}W_{t,i',r} = O(n_t^{-r})$ and $\mathbf{E}W_{t,j',r-1} = O(n_t^{-(r-1)})$, and again the total contribution from this case is $O(n_t^{-(r+1)})$.

Analogous to the case of $\Psi(i,0)$, for the expected number of subgraphs in $\Psi(i,r)$ destroyed in a single step is at least $W_{t,i,r}/n_t$. Thus

$$\mathbf{E}W_{t+1,i,r} \le \left(1 - \frac{1}{n_t}\right)\mathbf{E}W_{t,i,r} + O(n_t^{-(r+1)}).$$

Hence $\mathbf{E}W_{t,i,r} = O(n_t^{-r})$ by Lemma 3.1. ∎

In later arguments, we especially need to bound the number of subgraphs consisting of two distinguished cycles and sharing at least one edge. Of course the number of such subgraphs is bounded above by $\sum_{i=1}^{2k} W_{t,i,1}$, where k is length of the longer cycle given. Define $W^*_{t,k} = \sum_{i=1}^{2k} W_{t,i,1}$. Since $\mathbf{E}W_{t,i,1} = O(n_t^{-1})$, and the summation is taken over finitely many values of i, the following comes immediately from Lemma 3.3.

Corollary 3.1 $\mathbf{E}W^*_{t,k} = O(n_t^{-1})$.

Gearing up for the proof of Theorem 2.1, we next give some simple lemmas bounding some rare events. Let $\mathbf{Y}_t^{(l)} := (Y_{t,3}, Y_{t,4}, \ldots, Y_{t,l})$. In the following lemmas, the choice of the norm $\|\mathbf{Y}_t^{(l)}\|$ does not change the strength of the statement, and one may for instance settle on the L^∞ norm.

Lemma 3.4 *Fix the graph G_t. For any fixed $k \geq 3$, the probability that more than one cycle of length at most $k+1$ in G_{t+1} contains the peg vertex is $O(\|\mathbf{Y}_t^{(2k)}\|^2/n_t^2 + W^*_{t,k}/n_t)$.*

Proof. There are several cases to consider. The first case is that one edge pegged is contained in more than one cycle of length at most k, so that at least two cycles of length at most k will pass through the peg vertex. Since the subgraph consisting of two cycles of length at most k sharing common edges is involved, the probability this happens is at most $O(W^*_{t,k}/n_t)$. The second case is that one edge pegged is contained in a cycle of length at most k, and the other edge pegged is of distance at most k from the first edge. In this case, a new cycle is created using the path joining the pegged edges. There are at most $O(\|\mathbf{Y}_t^{(k)}\|)$ ways to choose the first edge, and for each such choice, there are at most $2d^k = O(1)$ ways to choose the second edge. So the probability that this case happens is $O(\|\mathbf{Y}_t^{(k)}\|/n_t^2)$. The third case is that the two edges pegged are both contained in some cycle of length at most $2k$. The probability this happens is $O(\|\mathbf{Y}_t^{(2k)}\|/n_t^2)$, since there are at most $O(\|\mathbf{Y}_t^{(2k)}\|)$ ways to choose such two edges. The fourth case is that each of the two edges pegged is contained in a cycle of length at most k. The probability for this to happen is $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2)$. Then Lemma 3.4 follows. ∎

We will use Lemma 3.4 to show that the only significant things that can happen with respect to short cycles are (a) an edge of a short cycle is pegged and no other cycles are created or destroyed, or (b) a short cycle is created by pegging the ends of a short path and no other short cycles are created or destroyed.

Note that a cycle is destroyed only if at least one of its edges is pegged. So to create or destroy more than one k-cycle in one step, there must be at least two cycles of length at most $k+1$ containing the peg vertex. Hence the following result comes immediately from Lemma 3.4.

Corollary 3.2 $\mathbf{P}(|Y_{t+1,k} - Y_{t,k}| > 1 \mid \mathbf{Y}_t^{(2k)}, W^*_{t,k}) = O\big(\|\mathbf{Y}_t^{(2k)}\|^2/n_t^2 + W^*_{t,k}/n_t\big)$.

By taking expectation of both sides of the equation in the statement of Lemma 3.2, and by Lemma 3.2 and Corollary 3.1, we have the following corollary.

Corollary 3.3 $\mathbf{P}(|Y_{t+1,k} - Y_{t,k}| > 1) = O\left(1/n_t^2\right)$.

Similarly, we may bound the simultaneous creation and destruction of cycles, except for a special case. (The following bounds are sufficient for our purposes and can easily be improved by examining the cases in the proof of Lemma 3.4.)

Corollary 3.4 *For any fixed integers l_1, $l_2 \geq 3$, such that $l_1 \neq l_2 + 1$, the probability of creating a new l_1-cycle and simultaneously destroying an existing l_2-cycle in the same step is $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2) + O(W^*_{t,k}/n_t)$, where $k = \max\{l_1, l_2\}$.*

Proof. The peg vertex is contained in the l_1-cycle that is created, and also in at least one of the edges in the l_2-cycle that is destroyed. If only one edge in this l_2-cycle is pegged, then a new (l_2+1)-cycle is created which contains the peg vertex. Since $l_1 \neq l_2 + 1$, the peg vertex is contained in both the l_1-cycle and the (l_2+1)-cycle. By Lemma 3.4, this happens with probability $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2) + O(W^*_{t,k}/n_t)$. If two edges in this l_2-cycle are pegged, then two short cycles containing the peg vertex and the rest of the edges of this l_2-cycle are created. By Lemma 3.4, this happens with probability $O(\|\mathbf{Y}_t^{(k)}\|^2/n_t^2) + O(W^*_{t,k}/n_t)$. Thus Corollary 3.4 follows. ∎

Next we check the moments $\mathbf{E}[Y_{t,3}]_j$ of $Y_{t,3}$, for any fixed $j \geq 0$. We set $Y_{t,2} = 0$ for any t, since the random graph generated is simple.

Lemma 3.5 *For any fixed nonnegative integer j,*

$$\lim_{t\to\infty} \mathbf{E}([Y_{t,3}]_j) = 3^j.$$

Proof. The proof is by induction on j. Lemma 3.2 shows that $\lim_{t\to\infty} \mathbf{E}([Y_{t,3}]_1) = 3$. So we may assume that $j \geq 2$ and $\mathbf{E}([Y_{t,3}]_{j-1}) \to 3^{j-1}$.

Instead of calculating $[Y_{t,3}]_j$ directly, we will calculate $[Y_{t,3}]_j/j!$, which is the number of j-sets of distinct i-cycles. We first consider the creation of a new j-set of triangles in moving from G_t to G_{t+1}, beginning with the j-sets that use an existing $(j-1)$-set of triangles, together with one newly created triangle.

We know that the expected number of triangles created at step t is $9/n_t + O(Y_{t,3})/n_t^2$. Each such new triangle creates a new j-set with each $(j-1)$-set of existing triangles except for those that simultaneously have one of their triangles destroyed. So the expected number of new j-sets created this way is

$$\left(\frac{9 + O(Y_{t,3}/n_t)}{n_t}\right)\frac{[Y_{t,3}]_{j-1}}{(j-1)!} + O\left(\frac{Y_{t,3}^2}{n_t^2} + \frac{W_{t,3}^*}{n_t}\right)\frac{[Y_{t,3}]_{j-1}}{(j-1)!}.$$

Here, the first term arises from the assumption that no existing triangles in the $(j-1)$-set are destroyed when the new triangle is created. The second, purely error term bounds the expected number of j-sets counted in the main term that should be discounted because, simultaneously with the new triangle being created, one of the triangles in the existing $(j-1)$-set is destroyed. The factor $O(Y_{t,3}^2/n_t^2 + W_{t,3}^*/n_t)$ comes from Corollary 3.4 for the probability of simultaneously creating and destroying triangles, and is multiplied by a bound on how many $(j-1)$-sets of existing triangles can contain one of the (bounded number of) triangles destroyed.

There are also j-sets that include more than one newly created triangle. It is straightforward to check that in one step it is possible to create at most four triangles, and destroy at most six. By Corollary 3.2, the probability of creating more than one triangle in one step, given $Y_{t,3}$ and $Y_{t,4}$, is $O(\|\mathbf{Y}_t^{(4)}\|^2/n_t^2 + W_{t,3}^*/n_t)$. Hence, the expected number of new j-sets created this way is at most

$$O\left(\frac{\|\mathbf{Y}_t^{(4)}\|^2}{n_t^2} + \frac{W_{t,3}^*}{n_t}\right)\sum_{i=2}^{6}\frac{[Y_{t,3}]_{j-i}}{(j-i)!}.$$

Note that $W_{t,3}^*[Y_{t,3}]_{j-i}$ is bounded above by $W_{t,3(j-i)+4,1}$ (representing the structures that come from the union of a structure in a $\Psi(3(j-i)+4,1)$). By Lemma 3.3 the expected number of such complex structures of bounded size is $O(n_t^{-1})$. Thus, using also the first part of Lemma 3.3,

$$\begin{aligned}
\mathbf{E}\left(\left(\frac{Y_{t,3}^2}{n_t^2} + \frac{W_{t,3}^*}{n_t}\right)[Y_{t,3}]_{j-2}\right) &= O(n_t^{-2}),\\
\mathbf{E}\left(O\left(\frac{\|\mathbf{Y}_t^{(4)}\|^2}{n_t^2} + \frac{W_{t,3}^*}{n_t}\right)\sum_{i=2}^{6}[Y_{t,3}]_{j-i}/(j-i)!\right) &= O(n_t^{-2}).
\end{aligned}$$

Now consider destroying an existing j-set. Firstly, assuming the j triangles are disjoint, then pegging any edge contained in those edges with any other non-adjacent edge will destroy the j-set. It follows that the expected number of j-sets being destroyed, given $Y_{t,3}$, is

$$\frac{3j[Y_{t,3}]_j/j!}{n_t} + \sum_{i' \leq 3j} \frac{O(W_{t,i',1})}{n_t}.$$

The error term $O(W_{t,i',1}/n_t)$ accounts for the case that the j-set of triangles share some common edges. So by Lemma 3.3

$$\mathbf{E}\left(\sum_{i' \leq 3j} \frac{O(W_{t,i',1})}{n_t} \right) = O(n_t^{-2}).$$

Thus

$$\begin{aligned} \mathbf{E}([Y_{t+1,3}]_j/j!) - \mathbf{E}([Y_{t,3}]_j/j!) &= \left(\frac{9 + O(n_t^{-1})}{n_t} \right) \mathbf{E}([Y_{t,3}]_{j-1}/(j-1)!) \\ &\quad - \frac{3j\mathbf{E}([Y_{t,3}]_j/j!) + O(n_t^{-1})}{n_t} + O(n_t^{-2}). \end{aligned}$$

By the inductive assumption, we have $\mathbf{E}([Y_{t,3}]_i) \to 3^i$ for any $i \leq j-1$. We use once again the argument as in the proof of Lemma 3.2. This calls for setting

$$\mathbf{E}([Y_{t+1,3}]_j/j!) - \mathbf{E}([Y_{t,3}]_j/j!) = 0$$

and then produces

$$\mathbf{E}([Y_{t,3}]_j/j!) \to \frac{9 \cdot 3^{j-1}/(j-1)!}{3j}.$$

Simplifying this gives

$$\mathbf{E}([Y_{t,3}]_j/j!) \to \frac{3^j}{j!}.$$

So $\mathbf{E}([Y_{t,3}]_j) \to 3^j$, as required. ∎

Proof of Theorem 2.1. It is enough to show that, for any fixed constant $k \geq 3$, and a given sequence of nonnegative integers $(j_3, j_4, \ldots, j_k)$,

$$\lim_{t \to \infty} \mathbf{E}([Y_{t,3}]_{j_3}[Y_{t,4}]_{j_4} \cdots [Y_{t,k}]_{j_k}) = \prod_{i=3}^{k} u_i^{j_i}.$$

We prove this by induction on the sequence of $(j_3, j_4, \ldots, j_k)$. The base case is $(j_3, 0, \ldots, 0)$, for any nonnegative integer j_3. Lemma 3.2 shows that $\mathbf{E}(Y_{t,j_3}) \to \mu_3^{j_3}$.

Let $\mathcal{S}(j_3, j_4, \ldots, j_k)$ denote the family of all collections (in whatever graph is under consideration) consisting a j_3-set of distinct 3-cycles, a j_4-set of distinct 4-cycles,..., and a j_k-set of distinct k-cycles. Let $\#(t, j_3, j_4, \ldots, j_k)$ be the number of elements of $\mathcal{S}(j_3, j_4, \ldots, j_k)$ in G_t. Note that

$$\#(t, j_3, j_4, \ldots, j_k) = \prod_{i=3}^{k} \frac{[Y_{t,i}]_{j_i}}{j_i!}.$$

We will estimate

$$\Delta := \mathbf{E}\left(\prod_{i=3}^{k} \frac{[Y_{t+1,i}]_{j_i}}{j_i!}\right) - \mathbf{E}\left(\prod_{i=3}^{k} \frac{[Y_{t,i}]_{j_i}}{j_i!}\right) = \mathbf{E}\#(t+1, j_3, j_4, \ldots, j_k) - \mathbf{E}\#(t, j_3, j_4, \ldots, j_k).$$

Case 1: Analogous to the creation of new triangles considered in the proof of Lemma 3.5, if a new i-cycle is created by pegging together the end edges of an i-path, then a new element of $\mathcal{S}(j_3, j_4, \ldots, j_k)$ can be created from an existing element of $\mathcal{S}(j_3, \ldots, j_i - 1, \ldots, j_k)$ together with the new i-cycle. The argument is similar to the proof of Lemma 3.5, and we omit the precise error terms since they have a similar nature and can be bounded in the same way. Instead we find that the contribution to Δ is

$$\sum_{i=3}^{k} \frac{3^{i-1}}{n_t} \mathbf{E}\left(\#(t, j_3, \ldots, j_i - 1, \ldots, j_k)\right) + O(n_t^{-2}).$$

Here we use the convention that $[x]_{-1} = 0$ for all x.

Case 2: If an edge of an $(i-1)$-cycle is pegged, for $i \le k$, then a new element of $\mathcal{S}(j_3, j_4, \ldots, j_k)$ can be created in several ways. The typical way is from an element of $\mathcal{S}(j_3, \ldots, j_{i-1} + 1, j_i - 1, \ldots, j_k)$, for some $4 \le i \le k$, that contains the $(i-1)$-cycle pegged. The expected number of elements of $\mathcal{S}(j_3, j_4, \ldots, j_k)$ created in this way is

$$\sum_{i=4}^{k} \frac{(i-1)(j_{i-1}+1) + O(n_t^{-1})}{n_t} \#(t, j_3, \ldots, j_{i-1} + 1, j_i - 1, \ldots, j_k)$$
$$+ O\Big(\sum W_{t,i',0}/n_t^2 + \sum W_{t,i',1}/n_t\Big).$$

The error term $O(n^{-2})$ accounts for the approximation for the number of possible peggings as before. The event that the two edges pegged are both in short cycles is accounted for by $O\big(\sum W_{t,i',0}/n_t^2\big)$. This also accounts for the case that a new short cycle is created as in Case 1 at the same time that an edge of a short cycle is pegged. The final error term accounts for the case that cycles in an element of $\mathcal{S}(j_3, \ldots, j_{i-1} + 1, j_i - 1, \ldots, j_k)$ share common edges, one of which is pegged; these cases should be discounted. It also accounts for other atypical ways to produce an element of $\mathcal{S}(j_3, \ldots, j_k)$, where an edge in two or more short cycles is pegged. These sums are taken over finitely many possible i'. By Lemma 3.3, the expected value of the error terms is $O(n_t^{-2})$.

An existing $\mathcal{S}(j_3, j_4, \ldots, j_k)$ can be destroyed by pegging any of its edges. Arguing as in the proof of Lemma 3.5, the contribution to Δ from destroying these configurations is

$$-\left(\sum_{i=3}^{k} i j_i\right) \frac{1}{n_t} \mathbf{E}\left(\#(t, j_3, j_4, \ldots, j_k)\right) + O(n_t^{-2}).$$

So we get

$$\begin{aligned}\Delta &= \sum_{i=3}^{k} \frac{3^{i-1}}{n_t} \mathbf{E}\left(\#(t, j_3, \ldots, j_i - 1, \ldots, j_k)\right) \\ &\quad + \sum_{i=4}^{k} \frac{(i-1)(j_{i-1}+1)}{n_t} \mathbf{E}\left(\#(t, j_3, \ldots, j_{i-1}+1, j_i - 1, \ldots, j_k)\right) \\ &\quad - \left(\sum_{i=3}^{k} i j_i\right) \frac{1}{n_t} \mathbf{E}\left(\#(t, j_3, j_4, \ldots, j_k)\right) + O(n_t^{-2}).\end{aligned}$$

By induction,

$$\mathbf{E}\left(\#(t, j_3, \ldots, j_i - 1, \ldots, j_k)\right) \to \frac{u_3^{j_3}}{j_3!} \cdots \frac{u_i^{j_i-1}}{(j_i-1)!} \cdots \frac{u_k^{j_k}}{j_k!} \qquad \text{for all } 3 \le i \le k.$$

$$\mathbf{E}\left(\#(t, j_3, \ldots, j_{i-1}+1, j_i - 1, \ldots, j_k)\right) \to \frac{u_3^{j_3}}{j_3!} \cdots \frac{u_{i-1}^{j_{i-1}+1}}{(j_{i-1}+1)!} \frac{u_i^{j_i-1}}{(j_i-1)!} \cdots \frac{u_k^{j_k}}{j_k!}$$

for all $4 \le i \le k$. So arguing as in the proof of Lemma 3.2, we set $\Delta = 0$ and obtain

$$\begin{aligned}\mathbf{E}\left(\#(t, j_3, j_4, \ldots, j_k)\right) &\to \left(\frac{1}{\sum_{i=3}^{k} i j_i}\right) \left(\sum_{i=3}^{k} 3^{i-1} \frac{u_3^{j_3}}{j_3!} \cdots \frac{u_i^{j_i-1}}{(j_i-1)!} \cdots \frac{u_k^{j_k}}{j_k!}\right. \\ &\quad \left. + \sum_{i=4}^{k} (i-1)(j_{i-1}+1) \frac{u_3^{j_3}}{j_3!} \cdots \frac{u_{i-1}^{j_{i-1}+1}}{(j_{i-1}+1)!} \frac{u_i^{j_i-1}}{(j_i-1)!} \cdots \frac{u_k^{j_k}}{j_k!}\right) \\ &= \prod_{i=3}^{k} \frac{u_i^{j_i}}{j_i!} \left(\frac{1}{\sum_{i=3}^{k} i j_i}\right) \left(\sum_{i=3}^{k} 3^{i-1} \frac{j_i}{\mu_i} + \sum_{i=4}^{k} (i-1)(j_{i-1}+1) \frac{\mu_{i-1}}{j_{i-1}+1} \frac{j_i}{\mu_i}\right).\end{aligned}$$

We only need to prove that

$$\sum_{i=3}^{k} 3^{i-1} \frac{j_i}{\mu_i} + \sum_{i=4}^{k} (i-1)(j_{i-1}+1) \frac{\mu_{i-1}}{j_{i-1}+1} \frac{j_i}{\mu_i} = \sum_{i=3}^{k} i j_i.$$

By calculating the left hand side, we get

$$\begin{aligned}&\sum_{i=3}^{k} 3^{i-1} \frac{j_i}{\mu_i} + \sum_{i=4}^{k} (i-1)(j_{i-1}+1) \frac{\mu_{i-1}}{j_{i-1}+1} \frac{j_i}{\mu_i} \\ &\qquad = \sum_{i=3}^{k} \frac{2 i j_i}{3^i - 9} 3^{i-1} + \sum_{i=4}^{k} (i-1) j_i \frac{3^{i-1} - 9}{2(i-1)} \frac{2i}{3^i - 9} \\ &\qquad = \sum_{i=3}^{k} \frac{2 i j_i}{3^i - 9} 3^{i-1} + \sum_{i=4}^{k} \frac{2 i j_i}{3^i - 9} \frac{3^{i-1} - 9}{2} \\ &\qquad = \sum_{i=3}^{k} i j_i.\end{aligned}$$

So we have shown that

$$\mathbf{E}\left(\prod_{i=3}^{k}\frac{[Y_{t,i}]_{j_i}}{j_i!}\right)\to\prod_{i=3}^{k}\frac{\mu_i^{j_i}}{j_i!},$$

and hence

$$\mathbf{E}\left(\prod_{i=3}^{k}[Y_{t,i}]_{j_i}\right)\to\prod_{i=3}^{k}\mu_i^{j_i}.$$

Theorem 2.1 then follows. ∎

4 Rate of convergence

In this section we prove Theorem 2.2. Because of the complexities of the proof, we treat the case $k=3$ first in detail. For simplicity, we simply use the notation Y_t in this proof to denote $Y_{t,3}$ in the case of $k=3$. With the aim of approximating Y_t, we define a Markov chain $(X_t)_{t\geq 0}$, a random walk on the nonnegative integers. To define this walk, we observe from the proof of Lemma 3.2 that the expected numbers of triangles created or destroyed in one step are approximately $9/n_t$ and $3Y_t/n_t$ respecively. Corollaries 3.2 and 3.4 show that typically creation and destruction of triangles occur disjointly, and no other events of significance occur. Hence, these two quantities give the significant transition probabilities for Y_t.

With this in mind, we define the transition probabilities for $(X_t)_{t\geq 0}$ as fallows. First, we define $\mathbf{B}_t := \{i\in\mathbf{Z}_+ : (9+3i)/n_t \leq 1\}$, and the boundary of $\mathbf{B}_t$ to be $\partial\mathbf{B}_t := \{i\in\mathbf{B}_t : i+1\notin\mathbf{B}_t\}$. Also w.p. denotes "with probability."

For $X_t\in\mathbf{B}_t\setminus\partial\mathbf{B}_t$,

$$X_{t+1}=\begin{cases} X_t-1 & \text{w.p. } 3X_t/n_t\\ X_t & \text{w.p. } 1-3X_t/n_t-9/n_t\\ X_t+1 & \text{w.p. } 9/n_t.\end{cases}\tag{4.1}$$

For $X_t\in\partial\mathbf{B}_t$,

$$X_{t+1}=\begin{cases} X_t-1 & \text{w.p. } 3X_t/n_t\\ X_t & \text{w.p. } 1-3X_t/n_t.\end{cases}\tag{4.2}$$

For $X_t\notin\mathbf{B}_t$,

$$X_{t+1}=X_t \quad \text{w.p. } 1.\tag{4.3}$$

We first show that $\mathbf{Po}(3)$, the Poisson distribution with mean 3, is a stationary distribution of the Markov chain $(X_t)_{t\geq 0}$. Assuming X_t has distribution $\mathbf{Po}(3)$, we have

$$\mathbf{P}(X_t=i)=e^{-3}\frac{3^i}{i!}\qquad\text{for all } i\in\mathbb{Z}_+$$

where $\mathbb{Z}_+$ denotes the set of nonnegative integers. Let $\mathbf{P}_{ij}=\mathbf{P}(X_{t+1}=j\mid X_t=i)$. For $j\in\mathbf{B}_t\setminus\partial\mathbf{B}_t$, we have

$$\begin{aligned}\mathbf{P}(X_{t+1}=j) &= \sum_{i\in\mathbb{Z}_+}\mathbf{P}(X_t=i)\mathbf{P}_{ij}\\ &= e^{-3}\frac{3^{j-1}}{(j-1)!}\frac{9}{n_t}+e^{-3}\frac{3^j}{j!}\left(1-\frac{9}{n_t}-\frac{3j}{n_t}\right)+e^{-3}\frac{3^{j+1}}{(j+1)!}\frac{3\,(j+1)}{n_t}\\ &= e^{-3}\frac{3^j}{j!}.\end{aligned}$$

For $j\in\mathbb{Z}_+$, such that $j\in\partial\mathbf{B}_t$, we have

$$\begin{aligned}\mathbf{P}(X_{t+1}=j) &= \sum_{i\in\mathbb{Z}_+}\mathbf{P}(X_t=i)\mathbf{P}_{ij}\\ &= e^{-3}\frac{3^{j-1}}{(j-1)!}\frac{9}{n_t}+e^{-3}\frac{3^j}{j!}\left(1-\frac{3j}{n_t}\right)\\ &= e^{-3}\frac{3^j}{j!}.\end{aligned}$$

For $j\in\mathbb{Z}_+$, such that $j\notin\mathbf{B}_t$, we have

$$\mathbf{P}(X_{t+1}=j) = \sum_{i\in\mathbb{Z}_+}\mathbf{P}(X_t=i)\mathbf{P}_{ij}=e^{-3}\frac{3^j}{j!}.$$

Thus $\mathbf{Po}(3)$ is invariant, so by definition it is a stationary distribution.

Let $(X_t)_{t\geq 0}$ have its stationary distribution $\mathbf{Po}(3)$ at $t=0$. Then its distribution remains $\mathbf{Po}(3)$ for all $t>0$. We aim to couple $(X_t)_{t\geq 0}$ with a process related to $(Y_t)_{t\geq 0}$. The specification of the transition probabilities in the coupled process is rather complicated because they are given in terms of some probabilities that are not known explicitly.

Define the following events for positive integers i:

$$\begin{aligned}L_{i,t} &: Y_{t+1}=Y_t-i,\\ R_{i,t} &: Y_{t+1}=Y_t+i,\\ S_t &: Y_{t+1}=Y_t,\\ \widetilde{L}_t &: X_{t+1}=X_t-1,\\ \widetilde{R}_t &: X_{t+1}=X_t+1,\\ \widetilde{S}_t &: X_{t+1}=X_t.\end{aligned}$$

(L, R and S stand for left, right and stay). The probabilities of the events $L_{i,t}$, $R_{i,t}$ and so on are determined by the values of t, Y_t and X_t and the dynamics of the Markov chain $(X_t)_{t\geq 0}$ and the process $(Y_t)_{t\geq 0}$. As we will see later, the probability of $L_{i,t}$ and $R_{i,t}$ for $i\geq 2$ is very small and finally determines only the size of the error terms. The following functions of X_t, Y_t and t are well defined:

$$\begin{aligned} p_1(t,j,m) &= \mathbf{P}(\widetilde{S}_t \mid X_t = m) + \mathbf{P}(S_t \mid Y_t = j) - 1, \\ \bar{a}_t(j) &= \max\{0, 9/n_t - \mathbf{P}(R_{1,t} \mid Y_t = j)\}, \\ \bar{b}_t(j) &= \max\{0, 3j/n_t - \mathbf{P}(L_{1,t} \mid Y_t = j)\}, \\ p_2(t,j) &= \mathbf{P}(S_t \mid Y_t = j) - \bar{a}_t(j) - \bar{b}_t(j). \end{aligned}$$

We will show later that, though $\bar{a}_t(Y_t)$ and $\bar{b}_t(Y_t)$ are involved in the analysis, their final contribution is only to the final error term.

We next define a coupling of $(X_t)_{t\geq 0}$ with a process $(Z_t)_{t\geq 0}$ such that Z_t has exactly the same distribution as Y_t for all $t \geq 0$. However the process $(Z_t)_{t\geq 0}$ is different from $(Y_t)_{t\geq 0}$, for $(Z_t)_{t\geq 0}$ is in fact a Markov chain (though not a time-homogeneous one), whereas $(Y_t)_{t\geq 0}$ is not. To make the definition, we define the initial distribution of (Z_0, X_0) and the transition probabilities for obtaining (Z_{t+1}, X_{t+1}) from (Z_t, X_t) as a Markov chain, and then check that the (marginal) distributions of Z_t and X_t agree with those of Y_t and the above definition of X_t.

For the distribution of (Z_0, X_0), we merely specify that X_0 and Z_0 are independent. We use the functions above in the definition of the transition probability of (Z_t, X_t) to (Z_{t+1}, X_{t+1}). To ensure that $p_1 \geq 0$, we make a special case of the following event:

$$\mathcal{C}_t := \{(Z_t, X_t) : 9 + 3X_t > \frac{1}{2} n_t \text{ or } Z_t \in A_t\} \tag{4.4}$$

where

$$A_t = \{m : \mathbf{P}(S_t \mid Y_t = m) < 1/2\}.$$

We show below that this event is unlikely. If it holds, we define Z_{t+1} and X_{t+1} by letting the two processes each take one step independently of each other; that is for all i and j, conditional upon (Z_t, X_t),

$$\mathbf{P}\big((Z_{t+1}, X_{t+1}) = (Z_t + i, X_t + j)\big) = \mathbf{P}(Y_{t+1} = Z_t + i \mid Y_t = Z_t)\mathbf{P}(X_{t+1} = X_t + j).$$

Now consider the other cases when $\mathcal{C}_t$ in (4.4) does not hold. Then $9 + 3X_t \leq n_t/2$, which implies that $X_t \in \mathbf{B}_t \setminus \partial\mathbf{B}_t$. Hence $\mathbf{P}(\widetilde{S}_t \mid X_t) = 1 - (9 + 3X_t)/n_t$. As mentioned in the proof of Lemma 3.4, each step creates at most four and destroys at most six triangles, and thus

$$\mathbf{P}(S_t \mid Y_t = j) = 1 - \sum_{i=1}^{6} \mathbf{P}(L_{i,t} \mid Y_t = j) - \sum_{i=1}^{4} \mathbf{P}(R_{i,t} \mid Y_t = j), \tag{4.5}$$

and hence

$$\begin{aligned} p_1(t, Y_t, X_t) &= 1 - 9/n_t - 3X_t/n_t - \sum_{i=1}^{6} \mathbf{P}(L_{i,t} \mid Y_t) - \sum_{i=1}^{4} \mathbf{P}(R_{i,t} \mid Y_t), \\ p_2(t, Y_t) &= 1 - \bar{a}_t(Y_t) - \bar{b}_t(Y_t) - \sum_{i=1}^{6} \mathbf{P}(L_{i,t} \mid Y_t) - \sum_{i=1}^{4} \mathbf{P}(R_{i,t} \mid Y_t). \end{aligned}$$

The assumption $(Z_t, X_t) \notin \mathcal{C}_t$ guarantees $p_1 \geq 0$ and $p_2 \geq 0$. To define the transitions in this case, we use two tables. Table 1 gives the transition probability when $X_t \neq Z_t$ and $\mathcal{C}_t$ is false. For example, the entry in the row labelled $j-i$ and column labelled m shows that

$$\mathbf{P}\big((Z_{t+1}, X_{t+1}) = (j-i, m) \mid (Z_t, X_t) = (j, m)\big) = \mathbf{P}(L_{i,t} \mid Y_t = j)$$

for $2 \leq i \leq 6$, where $\bar{a}$ and $\bar{b}$ are defined above. We emphasise that this probability is a number whose value we do not know explicitly, but is nevertheless a well defined function of t and j. Table 2 applies when $X_t = Z_t$ and $\mathcal{C}_t$ is false, and uses $\bar{a}$ and $\bar{b}$ as defined above.

	$m-1$	m	$m+1$
$j-i$ $(2 \leq i \leq 6)$	0	$\mathbf{P}(L_{i,t} \mid Y_t = j)$	0
$j-1$	0	$\mathbf{P}(L_{1,t} \mid Y_t = j)$	0
j	$3m/n_t$	$p_1(t, j, m)$	$9/n_t$
$j+1$	0	$\mathbf{P}(R_{1,t} \mid Y_t = j)$	0
$j+i$ $(2 \leq i \leq 4)$	0	$\mathbf{P}(R_{i,t} \mid Y_t = j)$	0

Table 1: Transition probabilities for (Z_t, X_t) when $\mathcal{C}_t$ is false — the case $j = Z_t \neq X_t = m$.

	$j-1$	j	$j+1$
$j-i$ $(2 \leq i \leq 6)$	0	$\mathbf{P}(L_{i,t} \mid Y_t = j)$	0
$j-1$	$3j/n_t - \bar{b}_t(j)$	$\mathbf{P}(L_{1,t} \mid Y_t = j) - 3j/n_t + \bar{b}_t(j)$	0
j	$\bar{b}_t(j)$	$p_2(t, j)$	$\bar{a}_t(j)$
$j+1$	0	$\mathbf{P}(R_{1,t} \mid Y_t = j) - 9/n_t + \bar{a}_t(j)$	$9/n_t - \bar{a}_t(j)$
$j+i$ $(2 \leq i \leq 4)$	0	$\mathbf{P}(R_{i,t} \mid Y_t = j)$	0

Table 2: The case $j = Z_t = X_t$

It is trivial to check that, for each table, the transition probabilities from X_t to X_{t+1} satisfy (4.1). Note for instance that in Table 2 the first column sums to $3j/n_t$, and in this table $X_t = j$. The same is clearly true when $\mathcal{C}_t$ holds. Hence, by induction, the marginal distribution of X_t is precisely that required to correctly call it X_t as defined in (4.1) etc.

Similarly, by definition of p_1 and p_2 and (4.5), the transition probabilities from Z_t to Z_{t+1} are exactly the same as those for going from Y_t to Y_{t+1}. This implies inductively that for all $t \geq 0$ the marginal distribution $\hat{\sigma}_{t,3}$ of Z_t satisfies

$$\hat{\sigma}_{t,3} = \sigma_{t,3} \tag{4.6}$$

(recalling that $\sigma_{t,3}$ is the distribution of Y_t). Hence, to prove the theorem, it is enough to bound the total variation distance between the marginal distributions of X_t and Z_t. This is what occupies the remainder of the proof.

Before proceeding we need to bound $\mathbf{P}(\mathcal{C}_t)$. In the derivation of (3.2), we saw that the expected number of triangles created in step t is $9/n_t+O((1+Y_t)/n_t^2)$, and the expected number destroyed is $3Y_t/n_t+O((1+Y_t)/n_t^2)$. By Corollary 3.4, the probability of simultaneously creating and destroying triangles in step t is $O(Y_t^2/n_t^2+W^*_{t,3}/n_t)$. Define $\xi(Y_t)=\mathbf{E}(W^*_{t,3} \mid Y_t)$, and by Lemma 3.3, $\mathbf{E}\xi(Y_t)=O(n_t^{-1})$. Therefore

$$\begin{aligned}\sum_{i=1}^{6} i\mathbf{P}(L_{i,t} \mid Y_t) &= \frac{3Y_t}{n_t}+O\left(\frac{1+Y_t^2}{n_t^2}+\frac{\xi(Y_t)}{n_t}\right)\\ \sum_{i=1}^{4} i\mathbf{P}(R_{i,t} \mid Y_t) &= \frac{9}{n_t}+O\left(\frac{1+Y_t^2}{n_t^2}+\frac{\xi(Y_t)}{n_t}\right).\end{aligned} \tag{4.7}$$

Note that we will use (4.7) to estimate $\mathbf{E}(\bar{a}_t(Z_t) \mid \cdot)$ and $\mathbf{E}(\bar{b}_t(Z_t) \mid \cdot)$ and show that their final contribution is negligible. We have $\mathbf{E}\xi(Y_t)=O(n_t^{-1})$, so $\mathbf{P}(\xi(Y_t)/n_t \geq a)=O(n_t^{-2})$ for any constant a, by Markov's inequality. We also have $\mathbf{E}Y_t=O(1)$ and $\mathbf{E}Y_t^2=O(1)$ from Lemma 3.5. Hence the variance of Y_t is $O(1)$, and so $\mathbf{P}(3Y_t/n_t \geq b) \leq O(n_t^{-2})$ for any constant b, by Chebyshev's inequality. Thus we have

$$\mathbf{P}\left(\sum_{i=1}^{6}\big(i\mathbf{P}(L_{i,t} \mid Y_t)+i\mathbf{P}(R_{i,t} \mid Y_t)\big) \geq \frac{1}{2}\right)=O(n_t^{-2}).$$

Hence

$$\begin{aligned}\mathbf{P}(Y_t \in A_t) &= \mathbf{P}\big(\mathbf{P}(S_t \mid Y_t)<1/2\big)\\ &= \mathbf{P}\left(\sum_{i=1}^{6}\big(\mathbf{P}(L_{i,t} \mid Y_t)+\mathbf{P}(R_{i,t} \mid Y_t)\big) \geq \frac{1}{2}\right)\\ &= O(n_t^{-2}).\end{aligned}$$

Similarly, since $\mathbf{E}X_t=3$, and $\sigma^2(X_t)=3$, we have $\mathbf{P}(9+3X_t \geq n_t/2)=O(n_t^{-2})$. Thus, referring to (4.4),

$$\mathbf{P}(\mathcal{C}_t)=O(n_t^{-2}). \tag{4.8}$$

We use $\overline{\mathcal{C}}_t$ for the complement of $\mathcal{C}_t$. We need to estimate the expected value $\bar{a}$ and $\bar{b}$ in Table 2, given that the Table is applicable, i.e. given that $Z_t=X_t$ and $\mathcal{C}_t$ is false. Many expectations and probabilities concerning Z_t and X_t will be conditional upon the event $\overline{\mathcal{C}}_t$. To simplify notation we denote $\mathbf{E}(\cdot \mid \overline{\mathcal{C}}_t)$ by $\widehat{\mathbf{E}}(\cdot)$ and $\mathbf{P}(\cdot \mid \overline{\mathcal{C}}_t)$ by $\widehat{\mathbf{P}}(\cdot)$.

At this point, we extend the definition of $L_{i,t}$ etc. to Z_t. Since the marginal distribution of $(Z_t)_{t\geq 0}$ has the same transition probabilities as $(Y_t)_{t\geq 0}$ we have for all j

$$\mathbf{P}(L_{i,t} \mid Y_t=j)=\mathbf{P}(L_{i,t} \mid Z_t=j) \tag{4.9}$$

and similarly for $R_{i,t}$. From (4.7), we thus have

$$\frac{9}{n_t} - \mathbf{P}(R_{1,t} \mid Z_t) = \sum_{i=2}^{4} i\mathbf{P}(R_{i,t} \mid Z_t) + O((1+Z_t^2)/n_t^2 + \xi(Z_t)/n_t).$$

By definition, $\bar{a}_t(Z_t)$ is the maximum of this quantity and 0, and $\mathbf{P}(R_{i,t} \mid Z_t) = O((1+Z_t^2)/n_t^2 + \xi(Z_t)/n_t)$ if $9/n_t - \mathbf{P}(R_{1,t} \mid Z_t) < 0$. Taking expectations of the equation above in the restricted space of appropriate (Z_t, X_t), we obtain the following bound to be used for $\bar{a}$ (using a similar argument for $\bar{b}$):

$$\begin{aligned} \widehat{\mathbf{E}}\left(\bar{a}_t(Z_t) \mid Z_t = X_t\right) &= \widehat{\mathbf{E}}\Bigg(\sum_{i=2}^{4} i\mathbf{P}(R_{i,t} \mid Z_t = X_t)\Bigg) + O\left(E\right) \\ &= \sum_{i=2}^{4} i\widehat{\mathbf{P}}(R_{i,t} \mid Z_t = X_t) + O\left(E\right), \\ \widehat{\mathbf{E}}(\bar{b}_t(Z_t) \mid Z_t = X_t) &= \sum_{i=2}^{6} i\widehat{\mathbf{P}}(L_{i,t} \mid Z_t = X_t) + O\left(E\right), \end{aligned} \tag{4.10}$$

where

$$E = \frac{1 + \widehat{\mathbf{E}}(Z_t^2 \mid Z_t = X_t)}{n_t^2} + \frac{\widehat{\mathbf{E}}(\xi(Z_t) \mid Z_t = X_t)}{n_t}.$$

Now we are ready to bound the difference between the distributions of Z_t and X_t (conditional upon $\overline{\mathcal{C}}_t$). We define

$$D_t = |Z_t - X_t|. \tag{4.11}$$

Restricting to the event $Z_t > X_t$, we see from Table 1 that $D_{t+1} - D_t$ increases by 1 if $X_{t+1} = X_t - 1$, decreases by 1 if $X_{t+1} = X_t + 1$ or $Z_{t+1} = Z_t - 1$, and increases by at most $|Z_{t+1} - Z_t|$ otherwise. Thus

$$\begin{aligned} \widehat{\mathbf{E}}(D_{t+1} - D_t \mid Z_t, X_t, Z_t > X_t) &\le \frac{3X_t}{n_t} - \mathbf{P}(L_{1,t} \mid Z_t) + \sum_{i=2}^{6} i\mathbf{P}(L_{i,t} \mid Z_t) + \sum_{i=1}^{4} i\mathbf{P}(R_{i,t} \mid Z_t) - \frac{9}{n_t} \\ &= \frac{3X_t}{n_t} - \frac{3Z_t}{n_t} + 2\sum_{i=2}^{6} i\mathbf{P}(L_{i,t} \mid Z_t) + O\left(\frac{1+Z_t^2}{n_t^2} + \frac{\xi(Z_t)}{n_t}\right) \end{aligned} \tag{4.12}$$

by (4.7) and (4.9). (In particular, note that these equations are applicable when conditioning upon $\overline{\mathcal{C}}_t$.)

Noting that $D_t = Z_t - X_t$ when $Z_t > X_t$, we have by the Tower Property

$$\begin{aligned} &\widehat{\mathbf{E}}(D_{t+1} - D_t \mid Z_t > X_t) \\ &\quad = \widehat{\mathbf{E}}(\widehat{\mathbf{E}}(D_{t+1} - D_t \mid X_t, Z_t, Z_t > X_t) \mid Z_t > X_t) \\ &\quad \le -\frac{3}{n_t}\widehat{\mathbf{E}}(D_t \mid Z_t > X_t) + 2\sum_{i=2}^{6} i\widehat{\mathbf{P}}(L_{i,t} \mid Z_t > X_t) \\ &\quad +O\left(\frac{1 + \widehat{\mathbf{E}}(Z_t^2 \mid Z_t > X_t)}{n_t^2} + \frac{\widehat{\mathbf{E}}(\xi(Z_t) \mid Z_t > X_t)}{n_t}\right). \end{aligned} \tag{4.13}$$

A similar calculation, for the case $Z_t < X_t$, gives

$$\widehat{\mathbf{E}}(D_{t+1} - D_t \mid Z_t < X_t) \;\le\; -\frac{3}{n_t}\widehat{\mathbf{E}}(D_t \mid Z_t < X_t) + 2\sum_{i=2}^{4} i\widehat{\mathbf{P}}(R_{i,t} \mid Z_t < X_t) + O\left(\frac{1+\widehat{\mathbf{E}}(Z_t \mid Z_t < X_t)}{n_t^2} + \frac{\widehat{\mathbf{E}}(\xi(Z_t) \mid Z_t < X_t)}{n_t}\right). \tag{4.14}$$

If $Z_t = X_t$ then $D_t = 0$, and from Table 2 we get

$$\begin{aligned}\widehat{\mathbf{E}}(D_{t+1} - D_t \mid Z_t, Z_t = X_t) &= \mathbf{P}(L_{1,t} \mid Z_t) - \frac{3Z_t}{n_t} + 2\bar{b}_t(Z_t) + \mathbf{P}(R_{1,t} \mid Z_t) - \frac{9}{n_t} + 2\bar{a}_t(Z_t) \\ &\quad + \sum_{i=2}^{6} i\mathbf{P}(L_{i,t} \mid Z_t) + \sum_{i=2}^{4} i\mathbf{P}(R_{i,t} \mid Z_t) \\ &= 2\bar{b}_t(Z_t) + 2\bar{a}_t(Z_t) + O\left(\frac{1+Z_t^2}{n_t^2} + \frac{\xi(Z_t)}{n_t}\right) \qquad (4.15)\end{aligned}$$

by (4.7) and (4.9) and the analogous equations concerning $L_{i,t}$. Applying the Tower Property as in the previous cases, and using (4.10), we obtain

$$\widehat{\mathbf{E}}(D_{t+1} - D_t \mid Z_t = X_t) = O\left(\frac{1+\widehat{\mathbf{E}}(Z_t^2 \mid Z_t = X_t)}{n_t^2} + \frac{\widehat{\mathbf{E}}(\xi(Z_t) \mid Z_t = X_t)}{n_t}\right) + \sum_{i=2}^{6} i\widehat{\mathbf{P}}(L_{i,t} \mid Z_t = X_t) + \sum_{i=2}^{4} i\widehat{\mathbf{P}}(R_{i,t} \mid Z_t = X_t). \tag{4.16}$$

Combining the three cases, and adding extra nonnegative terms $i\widehat{\mathbf{P}}(L_{i,t} \mid Z_t = X_t)$ and $i\widehat{\mathbf{P}}(R_{i,t} \mid Z_t = X_t)$ to the upper bounds where convenient,

$$\begin{aligned}\widehat{\mathbf{E}}\left(D_{t+1} - D_t\right) &= \widehat{\mathbf{E}}\left(D_{t+1} - D_t \mid Z_t > X_t\right)\widehat{\mathbf{P}}(Z_t > X_t) + \widehat{\mathbf{E}}\left(D_{t+1} - D_t \mid Z_t = X_t\right)\widehat{\mathbf{P}}(Z_t = X_t) \\ &\quad + \widehat{\mathbf{E}}\left(D_{t+1} - D_t \mid Z_t < X_t\right)\widehat{\mathbf{P}}(Z_t < X_t) \\ &\le -\frac{3}{n_t}\widehat{\mathbf{E}}\left(D_t \mid Z_t > X_t\right)\widehat{\mathbf{P}}(Z_t > X_t) - \frac{3}{n_t}\widehat{\mathbf{E}}\left(D_t \mid Z_t = X_t\right)\widehat{\mathbf{P}}(Z_t = X_t) \\ &\quad -\frac{3}{n_t}\widehat{\mathbf{E}}\left(D_t \mid Z_t < X_t\right)\widehat{\mathbf{P}}(Z_t < X_t) + O\left(\frac{1+\widehat{\mathbf{E}}Z_t^2}{n_t^2} + \frac{\widehat{\mathbf{E}}\xi(Z_t)}{n_t}\right) \\ &\quad + \sum_{2}^{6} i\widehat{\mathbf{P}}(L_{i,t}) + \sum_{2}^{4} i\widehat{\mathbf{P}}(R_{i,t}) \\ &= -\frac{3}{n_t}\widehat{\mathbf{E}}D_t + \sum_{2}^{6} i\widehat{\mathbf{P}}(L_{i,t}) + \sum_{2}^{4} i\widehat{\mathbf{P}}(R_{i,t}) + O\left(\frac{1+\widehat{\mathbf{E}}Z_t^2}{n_t^2} + \frac{\widehat{\mathbf{E}}\xi(Z_t)}{n_t}\right). \qquad (4.17)\end{aligned}$$

By (4.8), $\mathbf{P}(\overline{\mathcal{C}}_t) = 1 + O(n_t^{-2})$. Hence using Lemma 3.2 and (4.6), we have $\widehat{\mathbf{E}}Z_t^2 = \mathbf{E}(Z_t^2 \mid \overline{\mathcal{C}}_t) \le \mathbf{E}(Z_t^2)/\mathbf{P}(\overline{\mathcal{C}}_t) = O(1)$, $\widehat{\mathbf{E}}\xi(Z_t) \le \mathbf{E}\xi(Z_t)/\mathbf{P}(\overline{\mathcal{C}}_t) = O(n_t^{-1})$. Similarly by Corollary 3.2 and recalling (4.9), we know that for all i with $2 \le i \le 6$,

$$\widehat{\mathbf{P}}(L_{i,t}) = \mathbf{P}(L_{i,t} \mid \overline{\mathcal{C}}_t) = O(n_t^{-2})$$

and the analogue for $R_{i,t}$. Thus we now have

$$\widehat{\mathbf{E}}(D_{t+1} - D_t) = -\frac{3}{n_t}\widehat{\mathbf{E}}D_t + O(n_t^{-2}). \tag{4.18}$$

Since $|D_{t+1} - D_t| = O(1)$ always, (4.8) tells us that $\mathbf{E}(D_{t+1} - D_t \mid \mathcal{C}_t)\mathbf{P}(\mathcal{C}_t) = O(n_t^{-2})$. Hence

$$\begin{aligned}
\mathbf{E}(D_{t+1} - D_t) &= \widehat{\mathbf{E}}(D_{t+1} - D_t)\mathbf{P}(\overline{\mathcal{C}}_t) + \mathbf{E}(D_{t+1} - D_t \mid \mathcal{C}_t)\mathbf{P}(\mathcal{C}_t) \\
&= \left(-\frac{3}{n_t}\widehat{\mathbf{E}}D_t + O(n_t^{-2})\right)(1 + O(n_t^{-2})) + O(n_t^{-2}) \\
&= -\frac{3}{n_t}\widehat{\mathbf{E}}D_t + O(n_t^{-2}). \qquad (4.19)
\end{aligned}$$

Since

$$\mathbf{E}D_t = \widehat{\mathbf{E}}D_t\mathbf{P}(\overline{\mathcal{C}}_t) + \mathbf{E}(D_t \mid \mathcal{C}_t)\mathbf{P}(\mathcal{C}_t),$$

and $\mathbf{E}(D_t \mid \mathcal{C}_t) = O(n_t)$ as $D_t = O(n_t)$, we have

$$\widehat{\mathbf{E}}D_t = \mathbf{E}D_t(1 + O(n_t^{-2})) + O(n_t^{-1}),$$

and this, together with (4.19), gives

$$\mathbf{E}D_{t+1} = \left(1 - \frac{3}{n_t} + O(n_t^{-3})\right)\mathbf{E}D_t + O(n_t^{-2}). \tag{4.20}$$

By Lemma 3.1, this gives $\mathbf{E}D_t = O(1/n_t)$. Since $\mathbf{P}\left(Z_t \neq X_t\right) \leq \mathbf{E}D_t$, Lemma 2.1 now implies

$$d_{TV}\left(\hat{\sigma}_{t,3}, \mathbf{Po}(3)\right) = O(1/n_t)$$

and (4.6) now implies that $\tau_\epsilon^*\big((\sigma_{t,3})_{t\geq 0}\big) = O(1/\epsilon)$. This proves the upper bound on the ϵ-mixing time for $k = 3$.

Note that the error terms above can be simplified. The probability of simultaneously creating and destroying triangles in one step is actually $O(Y_t/n_t^2)$, rather than $O(Y_t^2/n_t^2 + W_{t,3}^*/n_t)$. Since triangles are only created by pegging the end edges of paths of length 3, we only need to consider the second case in Lemma 3.4. Thus, the error term in (4.7) will be $O((1 + Y_t)/n_t^2)$. However, we retain the extra error terms, to make it clearer how to extend the result to arbitrary k-cycles. Also note that, by (4.6), the coupling of the processes $(Z_t)_{t\geq 0}$ with $(X_t)_{t\geq 0}$ is also a coupling of $(Y_t)_{t\geq 0}$ with $(X_t)_{t\geq 0}$.

Now consider $k \geq 4$. Let $\mathbf{Y}_t^{(k)}$ denote $(Y_{t,3}, Y_{t,4}, \ldots, Y_{t,k})$, where as before $Y_{t,i}$ is the number of i-cycles in G_t. Also let $\mathbf{e}_i$ denote the elementary vector with 1 in the coordinate referring to $Y_{t,i}$ (i.e. the $(i-2)$th coordinate) and 0 elsewhere. We use the same convention for a vector $\mathbf{X}_t^{(k)} = (X_{t,3}, X_{t,4}, \ldots, X_{t,k})$, and such vectors will be written as $\mathbf{Y}_t$ etc. as long as the dependence on k does not need to be emphasised. We will extend the argument used for $k = 3$ to couple the sequence of vectors $(\mathbf{Y}_t)_{t\geq 0}$ with $(\mathbf{X}_t)_{t\geq 0}$ for any $k \geq 3$, where $(\mathbf{X}_t)_{t\geq 0}$ is a Markov Chain with the stationary distribution as that of independent Poisson variables.

We have shown in the proof of Lemma 3.2, that the expected number of i-cycles created in going from G_t to G_{t+1} by pegging the end edges of i-paths is

$$\frac{3^{i-1}}{n_t} + O\left(\frac{1+\|\mathbf{Y}_t^{(i)}\|}{n_t^2}\right). \tag{4.21}$$

We also saw that the expected number of i-cycles created from pegging an edge in a $(i-1)$-cycle is

$$\frac{(i-1)Y_{t,i-1}}{n_t} + O\left(\frac{Y_{t,i-1}}{n_t^2}\right), \tag{4.22}$$

where the error term comes from the case that two edges in the same cycle are pegged together (which could not happen in the case $i = 3$). When an edge in an $(i-1)$-cycle is pegged, it usually results in one more i-cycle but one less $(i-1)$-cycle, and so we expect $\mathbf{Y}_{t+1}$ to equal $\mathbf{Y}_t + \mathbf{e}_i - \mathbf{e}_{i-1}$. In fact, a decrease in $Y_{t,i-1}$ will normally force an increase $Y_{t,i}$.

Let $\mathbf{n} := \{n_3, n_4, \ldots, n_k\}$ be a vector of integers, and let $T_{\mathbf{n},t}$ denote the event that the number of i-cycles changes by n_i from G_t to G_{t+1} for any $3 \le i \le k$. The integers n_i can be positive, negative or 0, indicating increasing, decreasing or no change.

As we have seen earlier, the event that more than one i-cycle is created or destroyed is rare, and simultaneous creation and destruction is rare (except for the transition $T_{\mathbf{e}_i-\mathbf{e}_{i-1},t}$ mentioned above). More precisely, by (4.21), (4.22), Corollary 3.2, and Corollary 3.4,

$$\begin{aligned}
\mathbf{P}(T_{\mathbf{e}_i,t} \mid \mathbf{Y}_t) &= 3^{i-1}/n_t + \mathbf{E}(b(\mathbf{e}_i, t, G_t) \mid \mathbf{Y}_t) \qquad (3 \le i \le k),\\
\mathbf{P}(T_{\mathbf{e}_{i+1}-\mathbf{e}_i,t} \mid \mathbf{Y}_t) &= iY_{t,i}/n_t + \mathbf{E}(b(\mathbf{e}_{i+1}-\mathbf{e}_i, t, G_t) \mid \mathbf{Y}_t) \qquad (3 \le i \le k-1),\\
\mathbf{P}(T_{-\mathbf{e}_k,t} \mid \mathbf{Y}_t) &= kY_{t,k}/n_t + \mathbf{E}(b(-\mathbf{e}_k, t, G_t) \mid \mathbf{Y}_t),\\
\mathbf{P}(T_{\mathbf{n},t} \mid \mathbf{Y}_t) &= \mathbf{E}(b(\mathbf{n}, t, G_t) \mid \mathbf{Y}_t) \qquad \text{(for all other non-zero values of } \mathbf{n}\text{)}\\
\mathbf{P}(T_{\mathbf{0},t} \mid \mathbf{Y}_t) &= 1 - \sum_{i=3}^{k} \left(3^{i-1}/n_t + iY_{t,i}/n_t\right) - \sum_{\mathbf{n}\ne\mathbf{0}} \mathbf{E}(b(\mathbf{n}, t, G_t) \mid \mathbf{Y}_t)
\end{aligned}$$

where the values of the (error) terms satisfy $b(\mathbf{n}, t, G_t) = O\big((1 + \|\mathbf{Y}_t^{(2k)}(G_t)\|^2)/n_t^2 + W_{t,k}^*(G_t)/n_t\big)$ for all $\mathbf{n} \ne 0$.

It is now clear that the significant transitions of the random vector process $(\mathbf{Y}_t)_{t\ge0} = (\mathbf{Y}_t^{(k)})_{t\ge0}$ are

$$T_{\mathbf{e}_i,t},\ T_{\mathbf{e}_{i+1}-\mathbf{e}_i,t},\ T_{-\mathbf{e}_k,t},\ T_{\mathbf{0},t}.$$

We call these transitions the *main transitions* for the process $(\mathbf{Y}_t)_{t\ge0}$.

For convenience, let $a(\mathbf{n}, t, \mathbf{Y}_t) = \mathbf{E}(b(\mathbf{n}, t, G_t) \mid \mathbf{Y}_t)$, and define

$$\bar{a}(\mathbf{n}, t, \mathbf{Y}_t) = \begin{cases} \max\{0, -a(\mathbf{n}, t, \mathbf{Y}_t)\} & \text{if } \mathbf{n} \in \{\mathbf{e}_i, \mathbf{e}_{i+1}-\mathbf{e}_i, -\mathbf{e}_k\}\\ a(\mathbf{n}, t, \mathbf{Y}_t) & \text{for all other values of } \mathbf{n}\text{, such that } \mathbf{n} \ne \mathbf{0}. \end{cases} \tag{4.23}$$

Note that $\bar{a}(\mathbf{n}, t, \mathbf{Y}_t) \ge 0$, and is an extension of the functions $\bar{a}_t(j)$ and $\bar{b}_t(j)$ in the case $k = 3$. Since $\mathbf{E}\|\mathbf{Y}_t^{(2k)}\|^2 = O(1)$ and $\mathbf{E}W_{t,k}^* = O(n_t^{-1})$ from Lemma 3.2 and Corollary 3.1, we have

$$\mathbf{E}\, a(\mathbf{n}, t, \mathbf{Y}_t) = \mathbf{E}\, b(\mathbf{n}, t, G_t) = O(n_t^{-2}) \text{ for all } \mathbf{n} \ne 0. \tag{4.24}$$

This indicates that the final contribution of $\bar{a}(\mathbf{n}, t, \mathbf{Y}_t)$ will be $O(n_t^{-2})$ and hence negligible for our argument. The proof of this is basically the same as the analogous statement regarding $\bar{a}_t(j)$ and $\bar{b}_t(j)$ in the case $k = 3$.

Similar to the case of $k = 3$, we define

$$\begin{aligned}
\mathbf{B}_{t,k} &:= \{\mathbf{x} = (x_3, x_4, \ldots, x_k) \in \mathbf{Z}_+^{k-2} : \sum_{i=3}^{k} 3^{i-1} + k \sum_{i=3}^{k} x_i \le n_t\}, \\
\partial \mathbf{B}_{t,k} &:= \{\mathbf{x} = (x_3, x_4, \ldots, x_k) \in \mathbf{B}_{t,k} : \exists i,\, 3 \le i \le k,\, \mathbf{x} + \mathbf{e}_i \notin \mathbf{B}_{t,k}\}.
\end{aligned}$$

It is immediate by definition that for any $\mathbf{x} \in \mathbf{B}_{t,k}$, we have

$$\frac{1}{n_t} \sum_{i=3}^{k} \left(3^{i-1} + i x_i\right) \le 1.$$

We define the random walk $(\mathbf{X}_t)_{t \ge 0}$ as follows. For $\mathbf{X}_t \in \mathbf{B}_{t,k} \setminus \partial \mathbf{B}_{t,k}$,

$$\mathbf{X}_{t+1} = \begin{cases} \mathbf{X}_t + \mathbf{e}_i & \text{w.p. } 3^{i-1}/n_t \ (3 \le i \le k), \\ \mathbf{X}_t - \mathbf{e}_i + \mathbf{e}_{i+1} & \text{w.p. } i X_{t,i}/n_t \ (3 \le i \le k-1), \\ \mathbf{X}_t - \mathbf{e}_k & \text{w.p. } k X_{t,k}/n_t, \\ \mathbf{X}_t & \text{w.p. } 1 - \sum_{i=3}^{k} \left(3^{i-1} + i X_{t,i}\right)/n_t. \end{cases} \tag{4.25}$$

For $\mathbf{X}_t \in \partial \mathbf{B}_{t,k}$,

$$\mathbf{X}_{t+1} = \begin{cases} \mathbf{X}_t - \mathbf{e}_i + \mathbf{e}_{i+1} & \text{w.p. } i X_{t,i}/n_t \ (3 \le i \le k-1), \\ \mathbf{X}_t - \mathbf{e}_k & \text{w.p. } k X_{t,k}/n_t, \\ \mathbf{X}_t & \text{w.p. } 1 - \sum_{i=3}^{k} i X_{t,i}/n_t. \end{cases} \tag{4.26}$$

For $\mathbf{X}_t \notin \mathbf{B}_{t,k}$,

$$\mathbf{X}_{t+1} = \mathbf{X}_t \quad (\text{w.p. } 1). \tag{4.27}$$

We declare the initial distribution of $\mathbf{X}_0$ to be that of independent Poisson variables with means $\mu_3, \ldots, \mu_k$ respectively, where $\mu_i = (3^i - 9)/2i$, for all $3 \le i \le k$. That is,

$$\mathbf{P}\big(\mathbf{X}_0 = (x_3, \ldots, x_k)\big) = \exp\left(-\sum_{i=3}^{k} \mu_i\right) \prod_{i=3}^{k} \frac{\mu_i^{x_i}}{x_i!}. \tag{4.28}$$

We next verify that the independent Poisson distribution is invariant in one step of the Markov chain. Assuming the distribution holds at time t, for $\mathbf{x} \in \mathbf{B}_{t,k} \setminus \partial \mathbf{B}_{t,k}$, the (unconditional) probability that the random walk moves from $\mathbf{x}$ elsewhere in the next step is equal to the probability of moving from elsewhere to $\mathbf{x}$ (the details of this are straightforward and so omitted). See Figure 3, which illustrates the case $k = 4$. Here, the numbers shown on the arrows should be divided by n_t to obtain probabilities. The probability of making a transition from $\mathbf{x}$ to a state with a larger value of $\sum x_i$ (i.e. the grey region in the figure) is

$$\frac{1}{n_t} \sum_{i=3}^{k} 3^{i-1} \mathbf{P}(\mathbf{X}_t = \mathbf{x}).$$

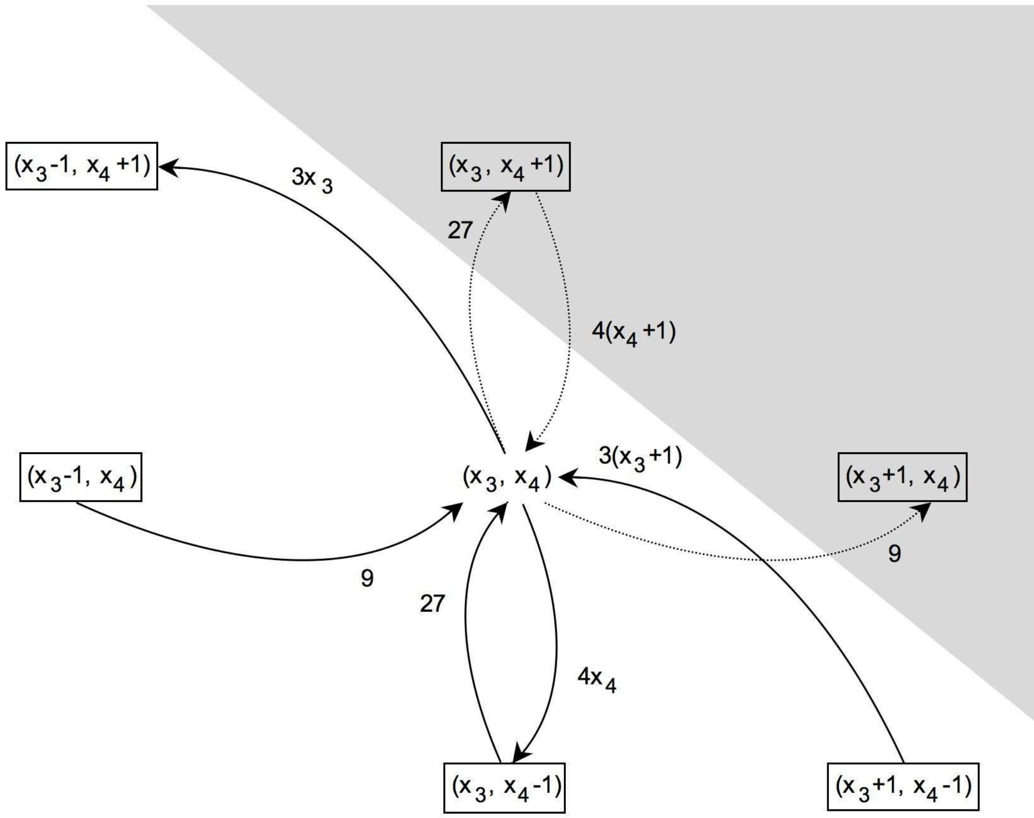

Figure 3: *When* $(X_3, X_4) \in \partial \mathbf{B}_{t,4}$*, the shaded region is outside* $\mathbf{B}_{t,4}$.

On the other hand, the probability of moving from one of these states to $\mathbf{x}$ is, assuming $\mathbf{X}_t$ has the same independent Poisson distribution as $\mathbf{X}_0$ in (4.28),

$$\frac{1}{n_t} k(x_k + 1)\mathbf{P}(\mathbf{X}_t = \mathbf{x} + \mathbf{e_k}) = \frac{1}{n_t} k \frac{3^k - 9}{2k} \mathbf{P}(\mathbf{X}_t = \mathbf{x})$$

which works out to be the same. For $\mathbf{x} \in \partial \mathbf{B}_{t,k}$, exactly these transitions to and from $\mathbf{x}$ are prevented, and hence these states maintain the Poisson distribution as well. Finally, for $\mathbf{x} \notin \mathbf{B}_{t,k}$, the probability of changing state is 0, thus the required relation holds for such $\mathbf{x}$. Hence, the joint distribution of $\mathbf{X}_t$ remains the same as that of $\mathbf{X}_0$ for all $t \geq 0$.

The coupling definition is similar to the case $k = 3$. In order to present the proof in a less cluttered fashion, we will first give a coupling between $\mathbf{X}_t$ and $\mathbf{Z}_t$ for all $t \geq 0$, where $(\mathbf{X}_t)_{t \geq 0}$ has initial distribution equal to its stationary distribution, and $(\mathbf{Z}_t)_{t \geq 0}$ is an independent copy of the same Markov chain, but with an arbitrary initial distribution. Then only the main transitions will occur with nonzero probability. Later, we will define a modified version of $(\mathbf{Z}_t)_{t \geq 0}$ that is a Markov chain in which $\mathbf{Z}_t$ has the same distribution as $\mathbf{Y}_t$.

For convenience, let the notation $T^{\mathbf{X}}_{\mathbf{n},t}$ denote the event that $\mathbf{X}$ takes the transition $T_{\mathbf{n},t}$, and similarly for $\mathbf{Z}$. Define

$$\mathcal{C}_t = \{(\mathbf{Z}_t, \mathbf{X}_t) : \mathbf{Z}_t \in B_t \text{ or } \ \mathbf{X}_t \in B_t\}, \tag{4.29}$$

where

$$B_t = \{\mathbf{j} \notin \mathbf{B}_{t,k}, \ \text{ or } \ \mathbf{j} \in \mathbf{B}_{t,k} : \mathbf{P}(\mathbf{X}_{t+1} = \mathbf{X}_t \mid \mathbf{X}_t = \mathbf{j}) < \frac{1}{2}\}.$$

If $(\mathbf{Z}_t, \mathbf{X}_t) \in \mathcal{C}_t$, the coupling of $\mathbf{Z}_t$ and $\mathbf{X}_t$ is defined by letting the two Markov chains take an independent step. Namely,

$$\mathbf{P}\big((\mathbf{Z}_{t+1}, \mathbf{X}_{t+1}) = (\mathbf{Z}_t + \mathbf{i}, \mathbf{X}_t + \mathbf{j}) \mid (\mathbf{Z}_t, \mathbf{X}_t)\big) = \mathbf{P}(\mathbf{Z}_{t+1} = \mathbf{Z}_t + \mathbf{i} \mid \mathbf{Z}_t)\mathbf{P}(\mathbf{X}_{t+1} = \mathbf{X}_t + \mathbf{j} \mid \mathbf{X}_t).$$

Now we consider $(\mathbf{Z}_t, \mathbf{X}_t)$ such that $(\mathbf{Z}_t, \mathbf{X}_t) \notin \mathcal{C}_t$. If $\mathbf{X}_t = \mathbf{Z}_t$, then $\mathbf{X}_t$ and $\mathbf{Z}_t$ take any of the main transitions simultaneously with probability $\mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t) = \mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t} \mid \mathbf{Z}_t)$.

If $\mathbf{X}_t \neq \mathbf{Z}_t$, let M be the smallest number such that $X_{t,M} \neq Z_{t,M}$. For $3 \leq i \leq M-1$, let $(T^{\mathbf{X}}_{\mathbf{n},t}, T^{\mathbf{Z}}_{\mathbf{n},t})$ have probability $\mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t) = \mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t} \mid \mathbf{Z}_t)$. For $i > M$, let $(T^{\mathbf{X}}_{\mathbf{n},t}, T^{\mathbf{Z}}_{\mathbf{0},\mathbf{t}})$ have probability $\mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t)$, and $(T^{\mathbf{X}}_{\mathbf{0},t}, T^{\mathbf{Z}}_{\mathbf{n},t})$ have probability $\mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t} \mid \mathbf{Z}_t)$, when $\mathbf{n} \in \{\mathbf{e}_i, \mathbf{e}_{i+1} - \mathbf{e}_i\}$ (if $i < k$), or $\mathbf{n} \in \{\mathbf{e}_k, -\mathbf{e}_k\}$ (if $i = k$). Namely, we let either $\mathbf{X}_t$ take one of its main transitions and $\mathbf{Z}_t$ stay unchanged, or vice versa. The only transitions of $(\mathbf{X}_t, \mathbf{Z}_t)$ left to be defined are those concerning the M-th coordinate of $\mathbf{X}_t$ and $\mathbf{Z}_t$, namely, $T_{\mathbf{e}_M - \mathbf{e}_{M-1},t}$, $T_{\mathbf{e}_M,t}$ and $T_{\mathbf{e}_{M+1} - \mathbf{e}_M,t}$ (or $T_{-\mathbf{e}_k,t}$ when $M = k$).

Let

$$(T^{\mathbf{X}}_{\mathbf{e}_M - \mathbf{e}_{M-1},t}, T^{\mathbf{Z}}_{\mathbf{e}_M - \mathbf{e}_{M-1},t}) \quad \text{have probability} \quad \mathbf{P}(T^{\mathbf{X}}_{\mathbf{e}_M - \mathbf{e}_{M-1},t} \mid \mathbf{X}_t),$$

which is equal to $\mathbf{P}(T^{\mathbf{Z}}_{\mathbf{e}_M - \mathbf{e}_{M-1},t} \mid \mathbf{Z}_t)$ because $X_{M-1} = Z_{M-1}$, and let

$$(T^{\mathbf{X}}_{\mathbf{e}_M,t}, T^{\mathbf{Z}}_{\mathbf{0},t}) \quad \text{and} \quad (T^{\mathbf{X}}_{\mathbf{0},t}, T^{\mathbf{Z}}_{\mathbf{e}_M,t}) \quad \text{each have probability} \quad \mathbf{P}(T^{\mathbf{X}}_{\mathbf{e}_M,t} \mid \mathbf{X}_t),$$

which is equal to $\mathbf{P}(T^{\mathbf{Z}}_{\mathbf{e}_M,t} \mid \mathbf{Z}_t)$ by the first line of (4.25).

On the other hand, since $X_{t,M} \neq Z_{t,M}$, for $\mathbf{n} = \mathbf{e}_{M+1} - \mathbf{e}_M$ (when $M < k$) or $\mathbf{n} = -\mathbf{e}_k$ (when $M = k$), let the transitions

$$(T^{\mathbf{X}}_{\mathbf{n},t}, T^{\mathbf{Z}}_{\mathbf{0},\mathbf{t}}) \text{ and } (T^{\mathbf{X}}_{\mathbf{0},\mathbf{t}}, T^{\mathbf{Z}}_{\mathbf{n},t}) \quad \text{have probabilities} \quad \mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t) \text{ and } \mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t} \mid \mathbf{Z}_t),$$

respectively.

The remaining probability is assigned to the transition $(T^{\mathbf{X}}_{\mathbf{0},\mathbf{t}}, T^{\mathbf{Z}}_{\mathbf{0},t})$. These probabilities are well defined, and in particular the last one is nonnegative, because $(\mathbf{Z}_t, \mathbf{X}_t) \notin \mathcal{C}_t$.

At this point it should be clear that the marginal transition probabilities from $\mathbf{X}_t$ to $\mathbf{X}_{t+1}$ satisfy (4.25). Also $\mathbf{Z}_t$ obeys the same rules. Hence the marginal distributions are the same for $\mathbf{Z}_t$ and $\mathbf{X}_t$.

We now make an adjustment to the definition of $\mathbf{Z}_t$. Let $\mathbf{Z}_0 = \mathbf{Y}_0$, and for $t \geq 0$, define a Markov chain in which

$$\mathbf{P}(\mathbf{Z}_{t+1} = \mathbf{j} \mid \mathbf{Z}_t = \mathbf{i}) = \mathbf{E}(\mathbf{P}(\mathbf{Y}_{t+1} = \mathbf{j} \mid G_t) \mid \mathbf{Y}_t = \mathbf{i}).$$

Then $\mathbf{Z}_t$ has the same distribution as $\mathbf{Y}_t$.

Now redefine

$$\mathcal{C}_t = \{(\mathbf{Z}_t, \mathbf{X}_t) : \mathbf{Z}_t \in A_t \text{ or } \ \mathbf{X}_t \in B_t\}, \tag{4.30}$$

where

$$A_t = \left\{\mathbf{i} : \mathbf{P}(\mathbf{Y}_{t+1} = \mathbf{Y}_t \mid \mathbf{Y}_t = \mathbf{i}) < \frac{1}{2}\right\},$$
$$B_t = \left\{\mathbf{j} \notin \mathbf{B}_{t,k}, \ \text{ or } \ \mathbf{j} \in \mathbf{B}_{t,k} : \mathbf{P}(\mathbf{X}_{t+1} = \mathbf{X}_t \mid \mathbf{X}_t = \mathbf{j}) < \frac{1}{2}\right\}.$$

An argument similar to the case $k = 3$ shows that $\mathbf{P}(\mathcal{C}_t) = O(n_t^{-2})$.

We now define the coupling of $\mathbf{X}_t$ and this new $\mathbf{Z}_t$. For $(\mathbf{Z}_t, \mathbf{X}_t)$ such that $(\mathbf{Z}_t, \mathbf{X}_t) \in \mathcal{C}_t$, again let the two Markov chains take an independent step. For $(\mathbf{Z}_t, \mathbf{X}_t)$ such that $(\mathbf{Z}_t, \mathbf{X}_t) \notin \mathcal{C}_t$, the same type of adjustment of transition probabilities can be made as in the case of $k = 3$, by adding and subtracting various a and $\bar{a}$ terms. We omit most of the details, but include discussion of the case $\mathbf{X}_t = \mathbf{Z}_t$. Here, we define $\mathbf{X}_t$ and $\mathbf{Z}_t$ to take the main nonzero transitions simultaneously with probability $\min\{\mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t), \mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t} \mid \mathbf{Z}_t)\}$. For $\mathbf{n} \in \{\mathbf{e}_i, \mathbf{e}_{i+1} - \mathbf{e}_i, -\mathbf{e}_k\}$,

$$\min\{\mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t), \mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t} \mid \mathbf{Z}_t)\} = \begin{cases} \mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t) & \text{if } a(\mathbf{n}, t, \mathbf{Z}_t) \geq 0 \\ \mathbf{P}(T^{\mathbf{X}}_{\mathbf{n},t} \mid \mathbf{X}_t) + a(\mathbf{n}, t, \mathbf{Z}_t) & \text{if } a(\mathbf{n}, t, \mathbf{Z}_t) < 0, \end{cases} \tag{4.31}$$

by the definition of $b(\mathbf{n}, t, G_t)$ and $a(\mathbf{n}, t, \mathbf{Z}_t)$, which occur just before (4.23), and

$$\bar{a}(\mathbf{n}, t, \mathbf{Z}_t) = \begin{cases} 0 & \text{if } a(\mathbf{n}, t, \mathbf{Z}_t) \geq 0 \\ -a(\mathbf{n}, t, \mathbf{Z}_t) & \text{if } a(\mathbf{n}, t, \mathbf{Z}_t) < 0, \end{cases} \tag{4.32}$$

by the definition of $\bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$ in (4.23). Assign probability $\bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$ to the transition $(T^{\mathbf{X}}_{\mathbf{n},t}, T^{\mathbf{Z}}_{\mathbf{0},t})$, and probability $a(\mathbf{n}, t, \mathbf{Z}_t) + \bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$ to the transition $(T^{\mathbf{X}}_{\mathbf{0},t}, T^{\mathbf{Z}}_{\mathbf{n},t})$, when $\mathbf{n} \in \{\mathbf{e}_i, \mathbf{e}_{i+1} - \mathbf{e}_i, -\mathbf{e}_k\}$. Note this is simply an adjustment of probabilities so that the marginals of $\mathbf{X}_t$ and $\mathbf{Z}_t$ are satisfied.

In the remainder of the proof, we show that for fixed k,

$$\mathbf{P}(\mathbf{X}_t^{(k)} \neq \mathbf{Z}_t^{(k)}) = O(n_t^{-1}), \tag{4.33}$$

by induction on k. We have already shown that it is true for $k = 3$. Now assume it is true for $k - 1$ where $k \geq 4$.

For any $\tau > 0$ and $\tau \leq j \leq 2\tau$, let $H_j = H_j(\tau)$ be the event that $\mathbf{Z}_t^{(k-1)} = \mathbf{X}_t^{(k-1)}$ for all $\tau \leq t \leq j$. In the following, for any statements about H_i we assume $\tau \leq i \leq 2\tau$, sometimes implicitly. Clearly $H_{j+1} \subset H_j$ for all relevant j. Let $D_{t,i} = |X_{t,i} - Z_{t,i}|$, for $3 \leq i \leq k$, and $\mathbf{D}_t^{(j)} = \sum_{i=3}^{j} D_{t,i}$, a simple extension of D_t. Of course, H_t implies $\mathbf{Z}_t^{(k-1)} = \mathbf{X}_t^{(k-1)}$. By the definition of the coupling of $(\mathbf{Z}_t^{(k)}, \mathbf{X}_t^{(k)})$ above, conditional on $\mathbf{Z}_t^{(k)}$ (actually $\mathbf{Z}_t^{(k-1)}$ is enough) and H_t, $\mathbf{Z}_t^{(k-1)}$ and $\mathbf{X}_t^{(k-1)}$ take transitions simultaneously in going from G_t to G_{t+1}, except for adjustments of probabilities of the form $\bar{a}(\cdot, t, \mathbf{Z}_t)$, or $a(\cdot, t, \mathbf{Z}_t) + \bar{a}(\cdot, t, \mathbf{Z}_t)$. Hence, this situation is analogous the the case $Z_t = X_t$ for $k = 3$, with H_t playing the role of the event $Z_t = X_t$ in this case.

Let T_τ be chosen such that $\mathbf{P}(H_{T_\tau}) > 1/2$. Then

$$\mathbf{P}(\bar{H}_t) < \frac{1}{2} \qquad (\tau \leq t \leq T_\tau). \tag{4.34}$$

Define $\widehat{\mathbf{E}}(\cdot) = \mathbf{E}(\cdot \mid \bar{\mathcal{C}}_t)$. As in the case $k = 3$, if U_t is some bounded random variable, then $\mathbf{E}(U_t)$ and $\widehat{\mathbf{E}}(U_t)$ differ only by $O(n_t^{-2})$, for the reason that $\mathbf{P}(\mathcal{C}) = O(n_t^{-2})$. It is straightforward to obtain $\widehat{\mathbf{E}}(\mathbf{D}_{t+1} - \mathbf{D}_t \mid \mathbf{Z}_t, H_t)$ in a form analogous to (4.15), with $L_{i,t}$ and $R_{i,t}$ being replaced by the more general $T_{\mathbf{n}}$, and $\bar{a}_t(Z_t)$, $\bar{b}_t(Z_t)$ being replaced by $\bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$ or $a(\mathbf{n}, t, \mathbf{Z}_t) + \bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$. This argument uses Corollary 3.2, and Corollary 3.4 to bound the effect of the transitions of $\mathbf{Z}_t$ that do not occur for $\mathbf{X}_t$. Then, taking expectation conditional on H_t (only), we obtain analogous to (4.16),

$$\widehat{\mathbf{E}}(\mathbf{D}_{t+1} - \mathbf{D}_t \mid H_t) = O\left(\frac{1 + \widehat{\mathbf{E}}(\|\mathbf{Z}_t\|^2 \mid H_t)}{n_t^2} + \frac{\widehat{\mathbf{E}}(\xi(\mathbf{Z}_t) \mid H_t)}{n_t}\right) + O\left(\sum \mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t} \mid H_t)\right),$$

where $\xi(\mathbf{Z}_t) = \mathbf{E}(W^*_{t,k} \mid \mathbf{Z}_t)$ (analogous to $\xi(Y_t)$ in the case $k = 3$), and the sum in the last term is taken over all $\mathbf{n}$ such that $T_{\mathbf{n},t}$ is not a main transition.

Note that $\mathbf{E}(\|\mathbf{Z}_t\|^2) = O(1)$ by Theorem 2.1, $\mathbf{E}(\xi(\mathbf{Z}_t)) = O(n_t^{-1})$ by Lemma 3.3, and $\mathbf{P}(T^{\mathbf{Z}}_{\mathbf{n},t}) = \mathbf{E}\,\bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$ (or $\mathbf{E}a(\mathbf{n}, t, \mathbf{Z}_t)$) $= O(n_t^{-2})$ by (4.24), for any non-main transition $T_{\mathbf{n}_t}$. These also hold for the conditional expectation $\widehat{\mathbf{E}}(\cdot)$ by an argument analogous to the comments after (4.17), recalling the bound on $\mathbf{P}(\mathcal{C})$ and also (4.34). That is, $\widehat{\mathbf{E}}(\|\mathbf{Z}_t\|^2 \mid H_t) = O(n_t^{-2})$ and so on (with the constant implicit in $O()$ independent of $t \le 2\tau$). Thus

$$\mathbf{E}(\mathbf{D}_{t+1} - \mathbf{D}_t \mid H_t) = O(n_t^{-2}), \tag{4.35}$$

and thus, since $\mathbf{P}(\bar{H}_{t+1} \mid H_t) \le \mathbf{E}(\mathbf{D}_{t+1} - \mathbf{D}_t \mid H_t)$, we have

$$\mathbf{P}(\bar{H}_{t+1} \mid H_t) = O(n_t^{-2}). \tag{4.36}$$

We now have

$$\mathbf{P}(\bar{H}_{t+1}) \le \mathbf{P}(\bar{H}_{t+1} \mid H_t) + \mathbf{P}(\bar{H}_t) \le O(n_t^{-2}) + \mathbf{P}(\bar{H}_t)$$

where of course the constant implicit in $O()$ is absolute. Iterating this, beginning with (4.33), we obtain $\mathbf{P}(\bar{H}_{t+1}) = O((t-\tau)n_\tau^{-2}) = O(n_\tau^{-1})$ for all $t \le \min\{2\tau, T_\tau\}$. For τ sufficiently large, this is at most $1/2$, and so T_τ can be chosen as 2τ. Thus

$$\mathbf{P}(\bar{H}_t) = O(n_\tau^{-1}) \qquad (\tau \le t \le 2\tau), \tag{4.37}$$

and (4.36) is valid for the same range of t. Hence

$$\begin{aligned}
\mathbf{P}(\mathbf{X}_t^{(k)} \ne \mathbf{Z}_t^{(k)}) &\le \mathbf{P}(\mathbf{X}_t^{(k-1)} \ne \mathbf{Z}_t^{(k-1)}) + \mathbf{P}(X_{t,k} \ne Z_{t,k}) \\
&\le \mathbf{P}(\bar{H}_t) + \mathbf{P}(X_{t,k} \ne Z_{t,k} \mid H_t) + \mathbf{P}(\bar{H}_t) \\
&\le \mathbf{E}(D_{t,k} \mid H_t) + O(n_\tau^{-1}).
\end{aligned}$$

We will shortly show that

$$\mathbf{E}(D_{t,k} \mid H_t) = O((t-\tau)^{-1}). \tag{4.38}$$

It then follows that for all t between say $3\tau/2$ and 2τ, $\mathbf{P}(\mathbf{X}_t^{(k)} \ne \mathbf{Z}_t^{(k)}) = O(n_\tau^{-1})$. For sufficiently large t, suitable τ can be chosen, and n_τ will be arbitrarily large. This establishes the inductive step (4.33), and the theorem follows.

It only remains to show (4.38). We first consider $\widehat{\mathbf{E}}(D_{t+1,k} - D_{t,k} \mid H_t)$, focussing on the transitions of $(Z_{t,k}, X_{t,k})$ induced by the transitions of $(\mathbf{Z}_t, \mathbf{X}_t)$. If $Z_{t,k} \ne X_{t,k}$, the transition probabilities in the coupling definition cause $(Z_{t,k}, X_{t,k})$ to have transition probabilities very similar to what is shown in Table 1 for $k = 3$, but with an extra entry in the $(j+1, m+1)$ position due to the transition $(T^{\mathbf{Z}}_{\mathbf{e}_k-\mathbf{e}_{k-1},t}, T^{\mathbf{X}}_{\mathbf{e}_k-\mathbf{e}_{k-1},t})$. The other significant effects for $(Z_{t,k}, X_{t,k})$ arise from the transitions

$$(T^{\mathbf{X}}_{\mathbf{e}_k,t}, T^{\mathbf{Z}}_{\mathbf{0},t}),\ (T^{\mathbf{X}}_{\mathbf{0},t}, T^{\mathbf{Z}}_{\mathbf{e}_k,t}),\ (T^{\mathbf{X}}_{-\mathbf{e}_k,t}, T^{\mathbf{Z}}_{\mathbf{0},t}),\ (T^{\mathbf{X}}_{\mathbf{0},t}, T^{\mathbf{Z}}_{-\mathbf{e}_k,t}),\ (T^{\mathbf{X}}_{\mathbf{0},t}, T^{\mathbf{Z}}_{\mathbf{0},t})$$

for $(\mathbf{Z}_t, \mathbf{X}_t)$.

Instead of the factors $\mathbf{P}(L_{i,t} \mid Y_t)$ and $\mathbf{P}(R_{i,t} \mid Y_t)$ that occurred in the case $k = 3$, we now have $\mathbf{P}(T_{\mathbf{n},t} \mid \mathbf{Z}_t)$ involved. Analogous to the argument concerning the equations (4.12), (4.14) and (4.16), apart from a or $\bar{a}$ terms, $\mathbf{P}(T_{\mathbf{e}_k,t} \mid \mathbf{Z}_{t,k})$ cancels $\mathbf{P}(\mathbf{X}_{t+1,k} = \mathbf{X}_{t,k} + \mathbf{e}_k) = 3^{k-1}/n_t$ by the definition of the coupling of $(\mathbf{X}_t, \mathbf{Z}_t)$. Similarly, the difference between $\mathbf{P}(T_{-\mathbf{e}_k,t} \mid \mathbf{Z}_{t,k})$ and $kX_{t,k}/n_t$ gives essential contribution $(-k/n_t)\widehat{\mathbf{E}}D_{t,k}$ to $\widehat{\mathbf{E}}(D_{t+1,k} - D_{t,k} \mid H_t)$. On the other hand, the (new) transition $(T^{\mathbf{Z}}_{\mathbf{e}_k-\mathbf{e}_{k-1},t}, T^{\mathbf{X}}_{\mathbf{e}_k-\mathbf{e}_{k-1},t})$ causes $Z_{t,k}$ and $X_{t,k}$ to move in the same direction, thus giving no contribution.

The non-main transitions $T_{\mathbf{n},t}$ with $n_k \neq 0$ contribute at most $\bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$ or $a(\mathbf{n}, t, \mathbf{Z}_t) + \bar{a}(\mathbf{n}, t, \mathbf{Z}_t)$, as with $\mathbf{P}(L_{i,t})$ and $\mathbf{P}(R_{i,t})$ for $i > 1$ in the case of $k = 3$.

In view of the bounds on a and $\bar{a}$ terms implied by (4.24), an argument very similar to that leading to (4.18) now gives

$$\widehat{\mathbf{E}}(D_{t+1,k} - D_{t,k} \mid H_t) = -\frac{k}{n_t}\widehat{\mathbf{E}}(D_{t,k} \mid H_t) + O(n_t^{-2}).$$

Then, similar to the steps leading to (4.35) (see also (4.19) to (4.20)), we obtain

$$\mathbf{E}(D_{t+1,k} \mid H_t) = \left(1 - \frac{k}{n_t}\right)\mathbf{E}(D_{t,k} \mid H_t) + O(n_t^{-2}). \tag{4.39}$$

In order to get a recursive equation, we would like $\mathbf{E}(D_{t+1,k} \mid H_{t+1})$ to appear in the left hand side of this equation. However, we have

$$\begin{aligned}\mathbf{E}(D_{t+1,k} \mid H_{t+1}) &= \frac{\mathbf{E}(D_{t+1,k} I_{H_{t+1}})}{\mathbf{P}(H_{t+1})} \\ &= \frac{\mathbf{E}(D_{t+1,k} I_{H_t}) - \mathbf{E}(D_{t+1,k} I_{H_t \setminus H_{t+1}})}{\mathbf{P}(H_t) - \mathbf{P}(H_t \setminus H_{t+1})} \\ &\leq \frac{\mathbf{E}(D_{t+1,k} I_{H_t})}{\mathbf{P}(H_t) - \mathbf{P}(H_t \setminus H_{t+1})}\end{aligned}$$

because $D_{t+1,k} \geq 0$. By (4.37) and (4.36), this gives

$$\mathbf{E}(D_{t+1,k} \mid H_{t+1}) \leq (1 + O(n_t^{-2}))\mathbf{E}(D_{t+1,k} \mid H_t).$$

So from (4.39) we obtain

$$\mathbf{E}(D_{t+1,k} \mid H_{t+1}) = \left(1 - \frac{k}{n_t} + O(n_t^{-2})\right)\mathbf{E}(D_{t,k} \mid H_t) + O(n_t^{-2}) \quad \text{for all } \tau \leq t \leq 2\tau.$$

Note from (4.37) and the fact that the expected values of $\mathbf{X}_t^{(k)}$ and $\mathbf{Z}_t^{(k)}$ are bounded, $\mathbf{E}(D_{\tau,k} \mid H_\tau) = O(1)$. Hence we obtain (4.38) by Lemma 3.1, as required.

5 Arbitrary $d \geq 3$

In this section we prove Theorem 2.3.

We begin with the case that d is even. This is a natural generalisation of the case $d = 4$. The strategy used in Theorem 2.1 and 2.2 to estimate the creation and destruction of short cycles applies in exactly the same way for the more general case. So we sketch here only the main steps of the first moment calculation generalising Lemma 3.2, and error terms that are eventually seen to be negligible will be ignored.

Let G_t be a random d-regular graph generated by pegging operations, for even d. Then G_t contains $n_t = n_0 + t$ vertices, and $m_t = dn_t/2$ edges. The number, N_t, of ways to do a pegging operation, i.e. the number of ways to choose $d/2$ non-adjacent edges, is asymptotically $\binom{dn_t/2}{d/2}$. For any fixed $k \geq 3$, there are two ways a k-cycle can be created in a pegging operation. One occurs if two of the $d/2$ edges that are pegged are the end edges of a k-path. The other occurs if one of the pegged edges is contained in a $(k-1)$-cycle.

The number of k-paths in G_t is asymptotically $d(d-1)^{k-1}n_t/2$, so the number of ways to form a k-cycle in the first way is asymptotically

$$\frac{d(d-1)^{k-1}n_t}{2}\binom{\frac{dn}{2}}{\frac{d}{2}-2} \sim \frac{d(d-1)^{k-1}n_t(dn_t/2)^{d/2-2}}{2(d/2-2)!}$$

and clearly for the second way it is asymptotically

$$(k-1)Y_{t,d,k-1}\binom{\frac{dn_t}{2}}{\frac{d}{2}-1} \sim \frac{(k-1)Y_{t,d,k-1}(dn_t/2)^{d/2-1}}{(d/2-1)!}$$

where $Y_{t,d,i}$ is the number of i-cycles in G_t. (Note that we let $Y_{t,d,2} = 0$.) To destroy an existing k-cycle, the algorithm pegs an edge contained in a k-cycle, together with another $d/2-1$ non-adjacent edges. The number of ways to destroy an existing k-cycle is thus asymptotically

$$kY_{t,d,k}\binom{\frac{dn_t}{2}}{\frac{d}{2}-1} \sim \frac{kY_{t,d,k}(dn_t/2)^{d/2-1}}{(d/2-1)!}.$$

Thus

$$\mathbf{E}(Y_{t+1,d,k} - Y_{t,d,k} \mid Y_{t,d,k}) = \frac{d(d-1)^{k-1}n_t(dn/2)^{d/2-2}}{2(d/2-2)!N_t} + \frac{(k-1)Y_{t,d,k-1}(dn_t/2)^{d/2-1}}{(d/2-1)!N_t} - \frac{kY_{t,d,k}(dn_t/2)^{d/2-1}}{(d/2-1)!N_t}.$$

Taking expectation of both sides,

$$\begin{aligned}
&\mathbf{E}(Y_{t+1,d,k} - Y_{t,d,k})\\
&\sim \frac{d(d-1)^{k-1}n_t(dn/2)^{d/2-2}}{2(d/2-2)!N_t} + \frac{(k-1)\mathbf{E}Y_{t,d,k-1}(dn_t/2)^{d/2-1}}{(d/2-1)!N_t} - \frac{k\mathbf{E}Y_{t,d,k}(dn_t/2)^{d/2-1}}{(d/2-1)!N_t}\\
&= \frac{(dn_t/2)^{d/2-2}dn_t}{(d/2-2)!N_t}\left(\frac{(d-1)^{k-1}}{2} + \frac{(k-1)\mathbf{E}Y_{t,d,k-1}}{d-2} - \frac{k\mathbf{E}Y_{t,d,k}}{d-2}\right)\\
&\sim \frac{d-2}{n_t}\left(\frac{(d-1)^{k-1}}{2} + \frac{(k-1)\mathbf{E}Y_{t,d,k-1}}{d-2} - \frac{k\mathbf{E}Y_{t,d,k}}{d-2}\right)\\
&= \frac{(d-2)(d-1)^{k-1}}{2n_t} + \frac{(k-1)\mathbf{E}Y_{t,d,k-1}}{n_t} - \frac{k\mathbf{E}Y_{t,d,k}}{n_t}.
\end{aligned}$$

Similar to the calculations in Lemma 3.2, we obtain by induction on k, starting with $\mathbf{E}Y_{t,d,2} = 0$, that

$$\mathbf{E}Y_{t,d,k} = \frac{(d-1)^k - (d-1)^2}{2k} + O\left(\frac{1}{n_t}\right).$$

The joint moments as in Theorem 2.1 and ϵ-mixing time as in Theorem 2.2 can also be estimated in a similar fashion to complete the proof of Theorem 2.3 for d even.

For d odd, the proof is again quite analogous, so here we only point out the salient features of computing the first moment of $Y_{t,d,k}$, ignoring the error terms as in the proof for arbitrary even d.

It is easy to see that $n_t = n_0 + 2t$, and $m_t = dn_t/2$. There are asymptotically $\binom{m_t}{(d-1)/2}$ ways to choose the non-adjacent edges in E_1 in the first step of the algorithm, which determine neighbours of the new vertex u, and then a similar number of ways to choose the edges in E_2. Hence, the total number of ways to perform a pegging on G_t is

$$N_t \sim \binom{m_t}{(d-1)/2}\binom{m_t}{(d-1)/2}/2,$$

where the factor 1/2 accounts for the double counting caused by the symmetry resulting from interchanging u and v together with E_1 and E_2.

There are three ways to create a new k-cycle:

Case 1: Two of the pegged edges (one from E_1, say e_1, and the other from E_2, say e_2, in the algorithm respectively) are end edges of a $(k-1)$-path. The expected number of k-cycles created of this type is

$$\big(d(d-1)^{k-2}n_t/2\big)\binom{m_t}{(d-1)/2-1}\binom{m_t}{(d-1)/2-1}/N_t.$$

Case 2: Two of the edges e_1 and e_2 (both from E_1 or both from E_2) chosen are end edges of a k-path. This is like the even degree case, so only one new vertex is contained in the new cycle that is created. The number of ways to choose such two edges is asymptotically $d \cdot (d-1)^{k-1}n_t/2$, and the expected number of k-cycles created of this type is

$$\big(d(d-1)^{k-1}n_t/2\big)\binom{m_t}{(d-1)/2-2}\binom{m_t}{(d-1)/2}/N_t.$$

Case 3: A new k-cycle is created by pegging an edge in a $(k-1)$-cycle. Clearly, the contribution from this case is

$$(k-1)Y_{t,d,k-1}\binom{m_t}{(d-1)/2-1}\binom{m_t}{(d-1)/2}/N_t.$$

The expected number of k-cycles destroyed in one step is asymptotically

$$kY_{t,d,k}\binom{dn_t/2}{(d-1)/2-1}\binom{dn_t/2}{(d-1)/2}/N_t.$$

So

$$
\begin{aligned}
&\mathbf{E}(Y_{t+1,d,k} - Y_{t,k} \mid Y_{t,d,k}, Y_{t,d,k-1}) \\
&\qquad \sim \Bigg\{ \frac{d \cdot (d-1)^{k-2} n_t}{2} \binom{dn_t/2}{(d-1)/2-1} \binom{dn_t/2}{(d-1)/2-1} \\
&\qquad + \frac{d \cdot (d-1)^{k-1} n_t}{2} \binom{dn_t/2}{(d-1)/2-2} \binom{dn_t/2}{(d-1)/2} \\
&\qquad + (k-1) Y_{t,d,k-1} \binom{dn_t/2}{(d-1)/2-1} \binom{dn_t/2}{(d-1)/2} - k Y_{t,d,k} \binom{dn_t/2}{(d-1)/2-1} \binom{dn_t/2}{(d-1)/2} \Bigg\} \frac{1}{N_t}.
\end{aligned}
$$

It is now easy to obtain

$$
\mathbf{E}Y_{t,d,k} \to \frac{(d-1)^k - (d-1)^2}{2k}.
$$

by applying induction on the value of k, and using Lemma 3.1. ∎

6 Concluding remarks

We believe that the upper bound on ϵ-mixing time in Theorem 2.2 is essentially tight, that is, there is a matching lower bound of the form $\tau_\epsilon^*\big((\sigma_t)_{t\geq 0}\big) \geq c/n_t$. We will pursue this issue elsewhere.

Theorem 2.3 implies that the random d-regular graphs generated by the pegging algorithm are not uniformly distributed, since, in the uniform distribution, the expected number of k-cycles is asymptotic to $(d-1)^k/2k$. Nonetheless, the theorem indicates the possibility that the pegging model and the uniform model might be close in the sense of contiguity. Let $\mathcal{PG}(G_0, t, d)$ be the probability space of all random d-regular graphs generated by pegging algorithm at time t, starting with graph G_0. Let $\mathcal{G}_{n,d}$ be the probability space of all random d-regular graphs with uniform distribution. Let $n = n_0 + t$, and $\hat{G}$ be an arbitrary d-regular graph from $\mathcal{G}_{n,d}$. Let $Y_t(\hat{G})$ be the number of ways that $\hat{G}$ could be obtained in the pegging algorithm, i.e. $Y_t(\hat{G})$ is the total probability of all sequences $(G_0, G_1, \dots, G_t) \in \mathcal{P}(G_0, d)$ such that $G_t = \hat{G}$. If we can show that $Y_t/\mathbf{E}Y_t$ converges in distribution to some random variable W as $t \to \infty$, and $W > 0$ a.s., it follows that $\mathcal{PG}(G_0, t, d)$ is contiguous with $\mathcal{G}_{n,d}$, where $n = n_0 + t$ (see Janson, Łuczak and Rucinński [4, P. 266] for a discussion of this). This relationship is denoted by $\mathcal{PG}(G_0, t, d) \approx \mathcal{G}_{n,d}$. If true, it means that properties a.a.s. true in one model are also true in the other model.

The small subgraph conditioning method of Robinson and Wormald gives a way of proving the convergence of $Y_t/\mathbf{E}Y_t$ (see [11] for example). It is well known that for any integer $m \geq 3$, $X_3, X_4, \dots, X_m$ are independent Poisson random variables with means $\lambda_i = (d-1)^i/(2i)$, where X_i is the number of i-cycles in a graph $G \in \mathcal{G}_{n,d}$. Theorem 2.3 shows that for any integer $d \geq 3$, the number of cycles in random d-regular graphs generated by pegging are asymptotically independent Poisson random variables with means $\mu_i = ((d-1)^i - (d-1)^2)/(2i)$. To use the small subgraph conditioning method, one computes

$$
\delta_i = \frac{\mu_i}{\lambda_i} - 1 = -\frac{(d-1)^2}{(d-1)^i},
$$

and then it is easy to check that

$$\sum_{i=3}^{k} \lambda_i \delta_i^2 < \infty.$$

It then becomes conceivable that the variation in probabilities in the pegging model is strongly associated with the varying numbers of short cycles (see [10] and [4]). According to the method, this would be proved if we could show that $\mathbf{E}Y_t^2/(\mathbf{E}Y_t)^2 \to \exp\left(\sum_{i=3}^{k} \lambda_i \delta_i^2\right)$. Since $Y_t(G)$ is a probability, its expected value is the reciprocal of the number of d-regular graphs on n vertices, a well understood quantity. However, the estimation of $\mathbf{E}Y_t^2$ seems to be out of reach at present. Nevertheless, the finiteness of the above summation seems too lucky to happen by chance, and we make the following conjecture.

Conjecture 6.1 *For all fixed $d \geq 3$ and every fixed d-regular graph G_0, $\mathcal{PG}(G_0, t, d) \approx \mathcal{G}_{n,d}$, where $t = n - n_0$.*

References

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

[2] V. Bourassa and F. Holt, SWAN: Small-world wide area networks, in *Proceeding of International Conference on Advances in Infrastructures* (SSGRR 2003w), L'Aquila, Italy, 2003, paper# 64.

[3] C. Cooper, M. Dyer and C. Greenhill, Sampling regular graphs and a peer-to-peer network, *Proceedings of the sixteenth annual ACM-SIAM Symposium on Discrete Algorithms* (2005), 980–988.

[4] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs.* Wiley, New York, 2000.

[5] T. Lindvall, *Lectures on the Coupling Method.* Wiley, New York, 1992.

[6] B.D. McKay and N.C. Wormald, Uniform generation of random regular graphs of moderate degree, *J. Algorithms* **11** (1990), 52–67.

[7] H. Robalewska and N. C. Wormald, Random star d-processes, *Combinatorics, Probability and Computing* **9** (2000), 33–43.

[8] H. Robalewska, Random graphs generated by the star 2-process, *Random Structures and Algorithms* **11** (1997), 125–149.

[9] A. Ruciński and N.C. Wormald, Random graphs processes with degree restrictions, *Combinatorics, Probability and Computing* **1** (1992), 169–180.

[10] N.C. Wormald, The asymptotic distribution of short cycles in random regular graphs, *J. Combinatorial Theory. Ser. B* **31** (1981), 168–182.

[11] N.C. Wormald, Models of random regular graphs, *Surveys in Combinatorics, 1999*, London Mathematical Society Lecture Note Series **267** (J.D. Lamb and D.A. Preece, eds) Cambridge University Press, Cambridge, pp. 239–298, 1999.