The expected degree distribution in transient duplication divergence models

A. D. Barbour and Tiffany Y. Y. Lo

Institut für Mathematik, Universität Zürich, Winterthurertrasse 190, 8057 Zürich, Switzerland
E-mail address: `a.d.barbour@math.uzh.ch`

School of Mathematics and Statistics, University of Melbourne, Parkville, VIC 3010, Australia
E-mail address: `tiffanyloyinyuan@gmail.com`

Abstract. We study the degree distribution of a randomly chosen vertex in a duplication–divergence graph, under a variety of different generalizations of the basic model of Bhan et al. (2002) and Vázquez et al. (2003). We pay particular attention to what happens when a non-trivial proportion of the vertices have large degrees, establishing a central limit theorem for the logarithm of the degree distribution. Our approach, as in Jordan (2018) and Hermann and Pfaffelhuber (2021), relies heavily on the analysis of related birth–catastrophe processes, and couplings are used to show that a number of different formulations of the process have asymptotically similar expected degree distributions.

1. Introduction

The duplication–divergence (DD) model introduced in Bhan et al. (2002), Vázquez et al. (2003) and Chung et al. (2003), which we refer to as the basic model, can be described as follows. An initial network with m_0 vertices is given. At each discrete time step, a vertex in the network is randomly chosen, and a new vertex having an identical set of neighbours is added. With probability $1 - q$, the new set of edges is then independently thinned, with each edge having probability p of being retained. Thus the new vertex has as neighbours a subset of the neighbours of the originally chosen vertex, and hence has smaller (or equal) degree; at the same time, the degrees of some of its neighbours are increased by 1. We henceforth refer to this model as the basic DD model. There are other variants; for instance, a link may be allowed between the originally chosen vertex and its copy, and the new vertex may be allowed a random number of extra links to other, randomly chosen vertices. Relatively little is known about the properties of the models, though simulations suggest that they generate networks looking more like protein interaction networks than the standard models; see, for example, Vázquez et al. (2003), Middendorf et al. (2005) and Gibson and Goldberg (2011). A closely related model, introduced by Jordan (2018), is obtained by supposing that vertices are

duplicated at random times, chosen in such a way that each vertex has an exponentially distributed 'lifetime' before duplication, independently of all others. The original model can then be deduced as the sequence of states of this model at the successive duplication times. An advantage of the model in Jordan (2018) is that other events, such as removal of edges, can more naturally be introduced, as in Hermann and Pfaffelhuber (2021); in the original model, there is no obvious way of incorporating events that happen between duplications.

The information most easily deduced from the basic DD model is the distribution of the degree of a randomly chosen vertex, though other features have been addressed; for example, the numbers of cliques and stars of different sizes in Hermann and Pfaffelhuber (2016), and the evolution of the degree of a given vertex in Hermann and Pfaffelhuber (2016) and Frieze et al. (2020). Starting from a configuration on m_0 vertices, let $N_{m,k}$ denote the number of vertices of degree k in the graph when the total number of vertices is m, and let $n_{m,k} := \mathbb{E}N_{m,k}$ denote its expectation. Then, for each $m \geq m_0$, the quantities $\mathfrak{p}_{m,k} := m^{-1}n_{m,k}$, $k \geq 0$, define a probability distribution $\mathfrak{p}_m$ on the non-negative integers $\mathbb{Z}_+$. Furthermore, $\mathfrak{p}_m$ can be interpreted as the probability distribution at time m of a *time inhomogeneous* Markov chain Y on $\mathbb{Z}_+$. The state Y_m of the chain at time m represents the degree of a 'tagged' vertex, changing in response to the evolution of the DD model, with the tag switching to the *copy* whenever the tagged vertex is chosen for duplication. This discrete time inhomogeneous chain can be analyzed directly. However, it was observed in Jordan (2018) that there is a continuous time *homogeneous* pure jump Markov process X that apparently has very similar behaviour; indeed, X and Y have the same jump chain. The process X is a birth–catastrophe process of a particular kind, and this enables well known probabilistic techniques to be applied. In particular, X has an absorbing state at zero, with the set $\mathbb{N}$ of positive integers forming a single (transient) communicating class, and a criterion in terms of the parameters p and q can be established, to determine whether or not the probability of absorption in zero is equal to or less than 1. If the probability of absorption in zero is 1, the expected proportion of vertices that have degree zero tends to 1 as $t \to \infty$, so that the DD model generates a graph, most of whose vertices are isolated. If the probability of absorption in zero is less than 1, then there is an asymptotically non-trivial proportion of the vertices that have large degrees, of magnitude growing with time. These results, first established in Hermann and Pfaffelhuber (2016) in the case $q = 0$, are rather disappointing for the practical application of the DD model, where a limiting 'steady state' degree distribution would be much preferable.

However, the degree weighted distribution, with elements

$$\tilde{\mathfrak{p}}_{m,k} := \frac{k\mathfrak{p}_{m,k}}{\sum_{l \geq 1} l\mathfrak{p}_{m,l}},$$

can be represented as the probability distribution at time m of another inhomogeneous Markov chain, to which there is also a closely related continuous time homogeneous pure jump Markov process, studied in Jordan (2018). This process evolves on $\mathbb{N}$, which is a single communicating class. Thus, when the degree weighted process is positive recurrent, it has a non-trivial steady state distribution, that can be easily translated (by removing the weights) into the distribution of the original process, conditional on non-absorption. This, in turn, can be used as a template for the degree distributions of networks observed in practice. Again, a criterion for positive recurrence in terms of p and q can be explicitly given. What is more, a power law decay of the tail of this distribution, with an explicit exponent, was presaged for $q = 0$ in Jordan (2018), and made precise in Jacquet et al. (2020).

In this paper, we have two main aims. The first is to show that the results above are relatively robust, in that the rigid prescription of the DD model can be relaxed somewhat, without changing the qualitative behaviour. For instance, the thinning of edges in the copied vertex need not be binomial. It is relatively easy to modify the 'tagged' processes, but not all such modifications have obvious counterparts in the original DD process. The second aim is to examine the distribution

of the large degrees when the original process is transient. In practice, it need not be the case that an observed network is in equilibrium, and the degree distribution of a transient process at a particular time m may still provide a reasonable template. Here, we prove a central limit theorem for the *logarithm* of the degree of a randomly chosen vertex, if the degree is positive, implying an asymptotically log-normal distribution for the degree itself, together with a point mass on zero. We also make the connection between the discrete and continuous time models explicit, showing that the discrete time model can be represented as a random non-linear time change of the continuous time model, where the magnitude of the time change, in the time scale of the continuous process, converges almost surely as $t \to \infty$. We conduct our analysis in the setting in which q is not necessarily equal to zero, and this also reveals interesting behaviour in the degree weighted process. If $q = 0$, the criterion for positive recurrence of the degree weighted process takes the form $p < p_o := e^{-1}$, as in Jordan (2018). If $0 < q < q_0 := 1/(1 + e^2)$, the criterion is of the form $c_1(q) < p < c_2(q)$, where, somewhat surprisingly, $c_1(q) > 0$; above the value q_0, the degree weighted process is transient for all values of p.

1.1. *Model and results.* We work primarily with the following generalization of the DD model. If a vertex chosen for duplication has degree zero, then so has its copy. If a vertex chosen for duplication has degree $k \geq 1$ and set of neighbours $\mathcal{S}$, we suppose that the new vertex has the same set of neighbours $\mathcal{S}$ with probability $q_k + (1 - q_k)\pi_{kk}$, where $0 \leq q_k < 1$, and that its neighbours consist of a random subset of $\mathcal{S}$ of size j with probability $(1 - q_k)\pi_{kj}$, $0 \leq j < k$. Here, for each $k \geq 1$, $\Pi_k := \{\pi_{kj}, 0 \leq j \leq k\}$ is a probability distribution on $\{0, 1, \ldots, k\}$ such that $\pi_{kk} < 1$; for $k = 0$, $\pi_{00} = 1$. Given that the size of the subset is j, the set of neighbours is chosen uniformly among all j-subsets of $\mathcal{S}$. In the basic model, $q_k = q$ for all k, and Π_k is the binomial distribution $\mathrm{Bi}\,(k, p)$. We then assume that the parameters are such that, at least asymptotically as $n \to \infty$, the general model resembles the basic model. Specifically, we assume that, for $0 < p < 1$ and $0 \leq q < 1$, and for positive constants c_l and γ_l, $1 \leq l \leq 3$,

$$|q_k - q| \leq c_1 k^{-\gamma_1}; \quad |p_k - p| \leq c_2 k^{-\gamma_2} \quad \text{and} \quad k^{-2}\sigma_k^2 \leq c_3 k^{-\gamma_3}, \tag{1.1}$$

where kp_k and σ_k^2 are the mean and variance of Π_k. Note that, in order to obtain a satisfactory 'tagged' process, there needs to be a relationship between p_l and q_l, in that $\alpha_l := q_l + p_l(1 - q_l)$ has to be equal to a fixed value α for all l. This is because the probability that the new vertex retains a given link needs to have the same value α, whatever the degree of the vertex being copied. In the basic model, $c_1 = c_2 = 0$, $\gamma_3 = 1$ and $c_3 = p(1 - p)$.

Let $N_{m,k}$ and $n_{m,k}$ be as above. At time $m + 1$, a vertex v is chosen at random to be copied; with probability $m^{-1}N_{m,j}$, it has degree j. Hence the probability that the new vertex has degree k, conditional on the values $(N_{m,l}, l \geq 0)$, is given by

$$\frac{1}{m}\left\{q_k N_{m,k} + \sum_{j \geq k} N_{m,j}(1 - q_j)\pi_{jk}\right\},$$

and, if the new vertex has degree k, this adds one to $N_{m,k}$ to be included in the count $N_{m+1,k}$. However, for any j, $N_{m+1,j}$ may also differ from $N_{m,j}$ because vertices of degree j at time m may become neighbours of the new vertex, each removing one from the count of j-vertices at time $m+1$, and vertices of degree $j - 1$ may also become neighbours of the new vertex, each adding one to the count of j-vertices at time $m + 1$. The probability that one of the neighbours of a vertex of degree j is copied is j/m, and the probability that this event leads to a new link to the vertex is α (irrespective of the degree of the neighbour). Hence, and setting $n_{m,-1} = 0$ for all m, it follows that

$$n_{m+1,k} = n_{m,k} - \frac{k\alpha}{m}n_{m,k} + \frac{(k-1)\alpha}{m}n_{m,k-1} + \frac{q_k}{m}n_{m,k} + \frac{1}{m}\sum_{j \geq k}n_{m,j}(1 - q_j)\pi_{jk}, \tag{1.2}$$

for $k = 0, 1, \ldots$ and $m \geq m_0$.

Define Q to be the matrix with elements

$$
\begin{aligned}
Q_{k,k+1} &= \alpha k; & Q_{k,k} &= -\{\alpha k + (1 - q_k)(1 - \pi_{kk})\}; \\
Q_{k,j} &= (1 - q_k)\pi_{kj}, & 0 \leq j \leq k - 1; & \quad Q_{k,j} = 0, \quad j \geq k + 2,
\end{aligned}
\tag{1.3}
$$

for each $k \geq 1$, and with $Q_{0,j} = 0$ for all j; it is a Q-matrix, in the usual Markovian sense. Writing $\mathfrak{p}_{m,k} := m^{-1}n_{m,k}$, Equation (1.2) gives

$$
\begin{aligned}
\mathfrak{p}_{m+1,k} = \mathfrak{p}_{m,k} + \frac{1}{m+1}\Big\{ &-\mathfrak{p}_{m,k}(1 + \alpha k) + \alpha(k-1)\mathfrak{p}_{m,k-1} \\
&+ q_k\mathfrak{p}_{m,k} + \sum_{j \geq k}(1 - q_j)\mathfrak{p}_{m,j}\pi_{jk}\Big\},
\end{aligned}
\tag{1.4}
$$

or, in vector notation,

$$
\mathfrak{p}_{m+1}^\top = \mathfrak{p}_m^\top\{I + (m+1)^{-1}[Q]_{m+1}\}, \quad \text{and so} \quad \mathfrak{p}_m^\top = \mathfrak{p}_{m_0}^\top \prod_{j=m_0+1}^{m} \{I + j^{-1}[Q]_j\},
\tag{1.5}
$$

where the matrix $[Q]_j$ consists of the rows of Q with indices $0, 1, \ldots, j-1$, with all remaining rows set to be identically zero. The truncation of Q has no effect in (1.5), because the vector $\mathfrak{p}_m$ can only have positive mass on the elements $0, 1, \ldots, m$, but it is convenient, because the matrix $P^{(j)} := \{I + j^{-1}[Q]_j\}$ is a stochastic matrix for all j; the row sums are all 1, and the diagonal elements are non-negative, because of the truncation. This implies that the vector of expected proportions $(\mathfrak{p}_m, m \geq m_0)$ evolves as the state probabilities of a discrete time, inhomogeneous Markov chain, as in (1.5): $\mathfrak{p}_{m_0+s}^\top = \mathfrak{p}_{m_0}^\top \prod_{j=1}^{s} P^{(m_0+j)}$.

Rather than work immediately with this discrete time chain, to determine the properties of $\mathfrak{p}_m$, we observe that, in view of Equation (1.5), $\mathfrak{p}_m^\top$ should look much like $\mathfrak{p}_{m_0}^T e^{Q \log(m/m_0)}$, or, alternatively, like $\mathfrak{p}_{m_0}^T P(\log m - \log m_0)$, where $P(\cdot)$ denotes the semigroup of transition matrices generated by Q. Thus it is reasonable to suppose that the state probabilities at time $\log(m/m_0)$ of the pure jump Markov process X corresponding to P, with initial distribution $\mathfrak{p}_{m_0}$, should approximate the vector $\mathfrak{p}_m$. This continuous time chain was introduced in Jordan (2018), to describe the evolution of the degree of a 'tagged' vertex, in a model analogous to the discrete model, in which each vertex independently has an exponential waiting time before duplication, with mean 1. In this paper, we concentrate on the analysis of the Markov process X, which, being time homogeneous, is easier to work with than the inhomogeneous discrete time process. We then derive results for the discrete process as a consequence, in Section 6.

Since $Q_{0,0} = 0$ and, for all $k \geq 1$, $Q_{k,k+1} > 0$ and $\sum_{j=0}^{k-1} Q_{k,j} > 0$, the state zero is absorbing for X, and $\mathbb{N}$ is a single class, which is transient, because $Q_{1,0} > 0$. Hence X is either absorbed in zero, or drifts to infinity. Letting $\mathcal{X}$ denote the event $\{X_t \to \infty\}$, we give conditions in Section 2 that determine whether $\mathbb{P}[\mathcal{X}]$ is equal to or greater than 0; see Corollary 2.3. Our arguments are based on the Foster–Lyapunov–Tweedie approach, and have the flavour of those in Jordan (2018). Since the Foster–Lyapunov–Tweedie theorems are stated for *irreducible* chains, we modify X to a process X^* which has transitions out of the state zero, but which is recurrent exactly when $\mathbb{P}[\mathcal{X}] = 0$. Such processes are also useful when discussing the degree weighted process in Section 5 and other variants of the basic model in Section 7. Criteria for determining whether X^* is positive recurrent, null recurrent or transient are given in Theorem 2.1.

In Section 3, when $\mathbb{P}[\mathcal{X}] > 0$, we investigate how the process X approaches infinity on the event $\mathcal{X}$. The main result is an almost sure convergence theorem, Theorem 3.1, reminiscent of that for supercritical branching processes. The behaviour is one of exponential growth, interspersed with large downward jumps, approximately at the times of a Poisson process of constant rate. Under these circumstances, a central limit theorem, Theorem 4.1, analogous to that in Bansaye et al. (2013), is

established for $\log X_t$ in Section 4. This indicates that the expected proportions of vertices of large degree in the corresponding DD model follow an approximately log-normal distribution.

If $\mathbb{P}[\mathcal{X}] = 0$, the *conditional* distribution of X_t, given $\{X_t > 0\}$, may still have a non-trivial limit, and this is investigated using a continuous analogue of the degree weighted Markov chain in Section 5. This process is equivalent to a process of the form X^*, but with different parameters, and the conclusions of Theorem 2.1 can be carried over without difficulty. In particular, as mentioned above, the degree weighted process is only positive recurrent for a very restricted set of parameters, and is always transient if $q > 1/(1 + e^2)$. The corresponding results in the case $q = 0$, and some more detailed conclusions, can be found in Hermann and Pfaffelhuber (2016) and Jordan (2018). In particular, in Hermann and Pfaffelhuber (2016), the rate at which $\mathbb{P}[X_t = 0]$ converges to its limit is investigated, also in the case where $\mathbb{P}[\mathcal{X}] > 0$.

In Section 6, we show that the original, discrete time model, after a logarithmic time change, can be closely coupled to the continuous time model. This enables the results for the continuous time process X to be translated into corresponding conclusions for the discrete time process, both for the original and for the degree–weighted processes. In particular, for the original process, there are analogues of the almost sure convergence theorem (Theorem 6.4) and the central limit theorem (Theorem 6.5). For the degree–weighted process, the limiting distribution for the discrete time process is the same as that for the continuous time process.

Finally, in Section 7, we discuss some other variants of the models. In Hermann and Pfaffelhuber (2021), random deletion of edges is introduced into a model analogous to that of Jordan (2018), in which vertices are duplicated independently of one another, each with a rate $(m + 1)/m$ if the current number of vertices is m. This results in transitions corresponding to deaths in the related (inhomogeneous) Markov process describing the tagged particle. Here, we investigate the corresponding modification of the homogeneous Markov process X, and establish behaviour analogous to that of the original process X. We also allow multiple births to occur simultaneously, so that, between catastrophes, the process X resembles a Markov branching process. This may not seem natural in the DD context, though it turns out to be a useful generalization when considering random re-wiring variants. These models, such as that introduced in Pastor-Satorras et al. (2003), involve changes in the duplication rules that often prevent absorption in zero. This allows X, for appropriate parameter values, to have a non-trivial *unconditional* limiting distribution, rather than being absorbed in zero. Once again, our framework is sufficiently flexible to accommodate the analysis of such models.

Markov branching catastrophe processes of the kind discussed above are examined in their own right in Hermann and Pfaffelhuber (2020). Their analysis is based on a delightful duality between branching catastrophe processes and a piecewise deterministic Markov process. We use rather more elementary techniques, which are nonetheless able to cope with the extra variation in the basic model that we allow. There is also a large literature on birth, death and catastrophe processes, conducted under different sets of assumptions (see Kaplan et al. (1975), Pakes (1986), Brockwell (1986) and Bartoszyński et al. (1989), for example), none of which fit our setting; instead, we have used our own arguments.

2. Absorption in zero

In the next three sections, we study the behaviour of the pure jump Markov process $X := (X_t, t \in \mathbb{R}_+)$ on the non-negative integers $\mathbb{Z}_+$ that has Q-matrix

$$
\begin{aligned}
Q_{k,k+1} &= k\alpha_k; & Q_{k,k} &= -\{k\alpha_k + \beta_k(1 - \pi_{kk})\}; \\
Q_{k,j} &= \beta_k \pi_{kj}, \quad 0 \le j \le k-1; & Q_{k,j} &= 0, \quad j \ge k+2,
\end{aligned} \tag{2.1}
$$

for each $k \ge 1$, with $\alpha_k, \beta_k > 0$, and with $\Pi_k := \{\pi_{kj}, 0 \le j \le k\}$ a probability distribution on $\{0, 1, \ldots, k\}$ such that $\pi_{kk} < 1$. For $k = 0$, we define $Q_{0,j} = 0$ for all j, so that the state 0 is

absorbing. We assume, as before, that kp_k and σ_k^2 are the mean and variance of Π_k, and that, as in (1.1),

$$|p_k - p| \leq c_2 k^{-\gamma_2} \quad \text{and} \quad k^{-2}\sigma_k^2 \leq c_3 k^{-\gamma_3}; \tag{2.2}$$

we also suppose that

$$|\beta_k - \beta| \leq c_1 k^{-\gamma_1} \quad \text{and} \quad |\alpha_k - \alpha| \leq c_4 k^{-\gamma_4}, \tag{2.3}$$

for positive constants $\alpha, \beta, c_1, c_4, \gamma_1$ and γ_4, and we define

$$\gamma := \{\min_{1 \leq l \leq 4} \gamma_l\} \wedge 1. \tag{2.4}$$

The process that we study here is rather more general than the one with Q-matrix as in (1.3), whose parameters are constrained to have $\alpha_k := 1 - (1 - p_k)\beta_k = \alpha > 0$ constant, where $\beta_k := 1 - q_k$; for that process, we can then take $\gamma_4 = \min\{\gamma_1, \gamma_2\}$ and $c_4 = c_1 + c_2$. Because X is dominated by a Yule process with rate $\max_{k \geq 1} \alpha_k$, it is non-explosive.

We also extend X to a bivariate Markov jump process $(X, Z) := ((X_t, Z_t), t \in \mathbb{R}_+)$, whose Q-matrix has elements

$$\begin{aligned}
Q_{(k,l),(k+1,l)} &= k\alpha_k; \qquad Q_{(k,l),(k,l)} = -\{k\alpha_k + \beta_k\}; \\
Q_{(k,l),(j,l+1)} &= \beta_k \pi_{kj}, \quad 0 \leq j \leq k,
\end{aligned} \tag{2.5}$$

for all $k \geq 0$, with $\beta_0 := 0$; for all other pairs (k', l'), $Q_{(k,l),(k',l')} := 0$. The X-component yields the same process as before, and the Z-component is the counting process of downward jumps, including (for neatness) a 'jump' at rate $\beta_k \pi_{kk}$ at which the value of X does not change. Thus, in state (k, l), the jump rate for Z is β_k, indicating that Z is not too far from being a homogeneous Poisson process with rate β, which is exactly the case in the basic model. We let $\mathcal{F}_t := \sigma((X_s, Z_s), 0 \leq s \leq t)$.

The state 0 is absorbing for X. Because of the assumptions that $\beta_k > 0$ and $\pi_{kk} < 1$ for all $k \geq 1$, it follows that the set $\mathbb{N} := \{1, 2, \ldots\}$ forms a single class for X, which is not closed, because it is possible to reach state 0 from state 1, and hence is transient. As before, let $\mathcal{X}$ denote the event $\{X_t \to \infty\}$. The main result of this section is a criterion for determining whether $\mathbb{P}[\mathcal{X}]$ is equal to or greater than zero. The arguments that we use, as in Jordan (2018), are based on Foster–Lyapunov–Tweedie theorems for irreducible Markov processes, which are sufficiently robust to accommodate our more general model. However, since X is not irreducible, we first modify it to an irreducible process X^*, by adding a transition from state 0 to state 1 with positive rate α_0. All states in $\mathbb{Z}_+$ are now either positive recurrent, null recurrent or transient for X^*, since it is irreducible; $\mathbb{P}[\mathcal{X}] = 0$ if X^* is recurrent, and $\mathbb{P}[\mathcal{X}] > 0$ if X^* is transient. A criterion for deciding which is the case is established in the following theorem.

Theorem 2.1. *For the process X^*, we have the following behaviour:*

$$\begin{aligned}
(1): \quad &\text{if } \alpha < \beta \log(1/p), \text{ then } X^* \text{ is geometrically ergodic;} \\
(2): \quad &\text{if } \alpha = \beta \log(1/p), \text{ then } X^* \text{ is null recurrent;} \\
(3): \quad &\text{if } \alpha > \beta \log(1/p), \text{ then } X^* \text{ is transient.}
\end{aligned} \tag{2.6}$$

Proof: To prove (1), we apply Meyn and Tweedie (1993, Theorem 6.1) with $f(j) := (j + 1)^\eta$ for $\eta \in (0, 1]$ suitably chosen. It is enough then to show that $(Qf)(j) \leq -cf(j)$ for all j sufficiently large, for some $c > 0$, since, for all $j \geq 0$, $\sum_{l \neq j} Q_{j,l}(l + 1)^\eta < \infty$. Letting $Y_j \sim \Pi_j$, we have

$$(Qf)(j) = j\alpha_j[f(j + 1) - f(j)] + \beta_j(\mathbb{E}\{f(Y_j)\} - f(j)), \qquad j \geq 1. \tag{2.7}$$

Now

$$j[f(j + 1) - f(j)] = \eta(j + 1)^\eta(1 + O((j + 1)^{-1})),$$

and f is concave, so that

$$\mathbb{E}\{f(Y_j)\} - f(j) \leq f(\mathbb{E}Y_j) - f(j) = (jp_j + 1)^\eta - (j + 1)^\eta = (j + 1)^\eta(p^\eta - 1 + O(j^{-\gamma})), \tag{2.8}$$

where γ is as defined in (2.4). Hence it follows that

$$(Qf)(j) \le f(j)\{\alpha\eta + \beta(p^\eta - 1) + O(j^{-\gamma})\}.$$

Now

$$\frac{d}{d\eta}\{\alpha\eta + \beta(p^\eta - 1)\} = \alpha - \beta p^\eta \log(1/p) < 0$$

for all $\eta > 0$ small enough, if $\alpha < \beta\log(1/p)$, so that there are values of $\eta \in (0,1]$ such that $\alpha\eta < \beta(1 - p^\eta)$. Then, for any such η,

$$(Qf)(j) \le -c_\eta f(j)$$

for all j large enough, where $c_\eta := \frac{1}{2}\{\beta(1 - p^\eta) - \alpha\eta\} > 0$.   This proves (1).

To prove that X^* is recurrent when $\alpha = \beta\log(1/p)$, we apply Tweedie (1975, Theorem 3.3) to the jump chain of X^*. We take $f(j) := \log(j+1) + 1 - (j+1)^{-\zeta}$ with $0 < \zeta < \gamma$, where γ is as defined in (2.4); then $\lim_{j\to\infty} f(j) = \infty$, and we show that $Qf(j) \le 0$ for all j sufficiently large, if $\alpha = \beta\log(1/p)$. It is immediate that

$$j[f(j+1) - f(j)] = (1 + \zeta(j+1)^{-\zeta}) + O((j+1)^{-1}) \tag{2.9}$$

is bounded for all j, and that

$$f'(x) = \frac{1}{x+1} + \frac{\zeta}{(x+1)^{1+\zeta}}$$

is decreasing in $x > 0$, so that f is concave. Hence, applying Jensen's inequality, it follows that

$$\mathbb{E}\{f(Y_j)\} - f(j) \le f(\mathbb{E}Y_j) - f(j) = \log p + j^{-\zeta}(1 - p^{-\zeta}) + O(j^{-\gamma}).$$

Hence, from (2.7),

$$(Qf)(j) = j\alpha_j[f(j+1) - f(j)] + \beta_j\{\mathbb{E}\{f(Y_j)\} - f(j)\} \tag{2.10}$$

$$\le j\alpha[f(j+1) - f(j)] + \beta\{f(\mathbb{E}Y_j) - f(j)\} + O(j^{-\gamma}) \tag{2.11}$$

$$= \alpha(1 + \zeta j^{-\zeta}) + \beta\{\log p + j^{-\zeta}(1 - p^{-\zeta})\} + O(j^{-\gamma}) \tag{2.12}$$

$$= j^{-\zeta}\{\alpha\zeta + \beta(1 - p^{-\zeta})\} + O(j^{-\gamma}),$$

because $\alpha - \beta\log(1/p) = 0$. Since $e^x > 1 + x$ in $x > 0$, it follows that $p^{-\zeta} > 1 + \zeta\log(1/p)$, and hence that

$$\alpha\zeta + \beta(1 - p^{-\zeta}) < \zeta(\alpha - \beta\log(1/p)) = 0;$$

thus, because $\zeta < \gamma$, $(Qf)(j) \le 0$ for all j large enough.

To show that X^* is null, if $\alpha = \beta\log(1/p)$, we apply a theorem of Barbour et al. (1996, Appendix), which is just the continuous time analogue of Tweedie (1976, Theorem 9.1(ii)). Now taking $f(j) := \log(j+1) + (j+1)^{-\zeta}$, with $0 < \zeta < \gamma$ as before, we show that $(Qf)(j) \ge 0$ for all j large enough, and also that $\sum_{k\ge 0} Q_{jk}|f(j) - f(k)|$ is uniformly bounded for all k. As above, the key step is to approximate the expectation $\mathbb{E}\{f(Y_j)\}$, now with a *lower* bound. For this we use the lower bound

$$f(y) \ge f(x) + (y - x)f'(x) + (y - x)^2 f''(x), \tag{2.13}$$

vaild for y in some neighbourhood of x, in view of Taylor's theorem, and because, since $0 < \zeta < 1$, $f''(x) < 0$ for all $x > 0$. Indeed, because f, f' and f'' are uniformly close to $\log x$ and its derivatives for large x, it is easy to check that (2.13) is valid for all $y \ge x/2$, provided that $x \ge x_\zeta$, for some fixed $x_\zeta > 0$. For j such that $jp_j > x_\zeta$, we thus have

$$\mathbb{E}\{f(Y_j)\} \ge \mathbb{E}\Big\{\big(f(\mathbb{E}Y_j) + (Y_j - \mathbb{E}Y_j)f'(\mathbb{E}Y_j) + (Y_j - \mathbb{E}Y_j)^2 f''(\mathbb{E}Y_j)\big)\mathbf{1}_{[\frac{1}{2}\mathbb{E}Y_j,\infty)}(Y_j)\Big\}$$

$$\ge f(\mathbb{E}Y_j)(1 - \mathbb{P}[Y_j < \tfrac{1}{2}\mathbb{E}Y_j]) + f''(\mathbb{E}Y_j)\mathrm{Var}\,(Y_j). \tag{2.14}$$

From (1.1) and because of the definition of the function f, this immediately implies that, for all j large enough,

$$\mathbb{E}\{f(Y_j)\} \ \geq \ f(\mathbb{E}Y_j) + O(j^{-\gamma}\log j), \tag{2.15}$$

where $\mathbb{P}[Y_j < \frac{1}{2}jp_j]$ has been bounded using Chebyshev's inequality. Now, arguing exactly as for (2.10) and (2.11), but with the current definition of f, it follows that, for all j large enough,

$$(Qf)(j) \ \geq \ -j^{-\zeta}(\alpha\zeta + \beta(1 - p^{-\zeta})) + O(j^{-\gamma}\log j),$$

and hence that $(Qf)(j) \geq 0$ for all j large enough. It is easier to show that $\sum_{k\geq 0} Q_{jk}|f(j) - f(k)|$ is uniformly bounded by $2(\alpha + \beta\log(1/p))$, for all j large enough; the necessary estimates are essentially already in (2.12), and the extension to all j is immediate. This completes the proof of (2).

For part (3), we use Tweedie (1976, Theorem 11.3(i)) applied to the jump chain of X^*, with the function $f(j) := (j + 1)^{-\eta}$, for suitably chosen $\eta > 0$; since f is a decreasing function, we simply need to show that $(Qf)(j) \leq 0$ for all j sufficiently large. Now the key step is to upper bound the expectation $\mathbb{E}\{f(Y_j)\}$ using

$$f(y) \ \leq \ f(x) + (y - x)f'(x) + (y - x)^2 f''(x), \tag{2.16}$$

since $f''(x) > 0$ for all $x > 0$, valid as for (2.13) for all $y \geq \theta x$, where $\theta = \theta(\eta) < 1$. Indeed, by substituting $y = \theta x$ into (2.16), with $f(j) = (j + 1)^{-\eta}$, it can be seen that (2.16) is satisfied for all $y \geq \theta x$, provided that x is large enough, if θ is chosen close to 1 in such a way that

$$\theta^{-\eta} \ < \ 1 + \eta(1 - \theta) + \eta(1 + \eta)(1 - \theta)^2. \tag{2.17}$$

Since the left hand side of (2.17) can be written as $1 + \eta(1 - \theta) + \frac{1}{2}\eta(1 + \eta)(1 - \theta)^2 + O(|1 - \theta|^3)$, there are such values of θ to be chosen. Now, arguing much as for (2.14) and (2.15), with the current definition of f and with $\frac{1}{2}$ replaced by θ, it follows that

$$\mathbb{E}\{f(Y_j)\} - f(j) \ \leq \ f(\mathbb{E}Y_j) + O(j^{-\gamma}) - f(j) \ = \ j^{-\eta}(p^{-\eta} - 1) + O(j^{-\gamma}), \tag{2.18}$$

with γ as before. On the other hand,

$$j\alpha_j[f(j + 1) - f(j)] \ = \ -\alpha\eta j^{-\eta}(1 + O(j^{-\gamma})),$$

and hence, for all j large enough,

$$\begin{aligned} (Qf)(j) \ &= \ \{-\alpha\eta + \beta(p^{-\eta} - 1)\}j^{-\eta} + O(j^{-\gamma}) \\ &= \ \{-\alpha\eta + \beta(e^{\eta\log(1/p)} - 1)\}j^{-\eta} + O(j^{-\gamma}). \end{aligned}$$

Thus, if $\alpha > \beta\log(1/p)$, it is possible to choose $\eta > 0$ small enough that $\eta < \gamma$ and that

$$-\alpha\eta + \beta(e^{\eta\log(1/p)} - 1) \ < \ 0,$$

in which case $(Qf)(j) \leq 0$ for all j sufficiently large, establishing (3). This completes the proof of the theorem. ∎

Remark 2.2. Note that, if $\pi_{11} = 1$ and $\pi_{k0} = 0$ for all $k \geq 1$, then the process X is irreducible on $\mathbb{N}$. For such a process, we can take $X^* = X$, and the conclusions of Theorem 2.1 remain valid, either using the same proof, or by re-labelling the states by subtracting 1 from each label. The degree weighted process of Section 5 comes into this category.

As a consequence of Theorem 2.1, we have the following criterion determining whether $\mathbb{P}[\mathcal{X}]$ is zero or positive, valid because $\mathbb{P}[\mathcal{X}] = 0$ if X^* is recurrent, and $\mathbb{P}[\mathcal{X}] > 0$ if X^* is transient.

Corollary 2.3. *If $\alpha \leq \beta\log(1/p)$, then $\mathbb{P}[\mathcal{X}] = 0$; if $\alpha > \beta\log(1/p)$, then $\mathbb{P}[\mathcal{X}] > 0$.*

Remark 2.4. In the basic model, with $\beta = (1-q)$, the condition $\alpha \leq \beta \log(1/p)$ is equivalent to

$$pe^p \leq e^{-q/(1-q)}, \tag{2.19}$$

and is thus satisfied for $p \leq p_*(q)$, where $p_*(q)$ is the unique solution of (2.19) in $[0,1]$. As in Hermann and Pfaffelhuber (2016) and Jordan (2018), $p_*(0) \approx 0.5671$ is the solution to $pe^p = 1$; p_* is a decreasing function, and $p_*(1) = 0$.

In Jordan (2018), conditions are given as to which moments of the degree weighted process are finite. Here, we use similar arguments to make comparable statements about the moments of X^*, when it is positive recurrent. For any $u > 1$, let $x^*(u)$ denote the positive solution to the equation $x = u(1 - e^{-x})$; note that $x^*(u) < u$.

Theorem 2.5. *If $\alpha < \beta \log(1/p)$, the process X^* has a stationary distribution, whose η-th moment is finite for all $\eta < \eta^* := x^*(\beta \log(1/p)/\alpha)/\log(1/p)$, and is infinite for all $\eta > \eta^*$. In particular, $\eta^* < \beta/\alpha$.*

Proof: If $\alpha < \beta \log(1/p)$, $\eta \leq 1$ and $\alpha\eta < \beta(1-p^\eta)$, the existence of an η-th moment of the stationary distribution of X^* follows from the proof of case (1) in Theorem 2.1, by Meyn and Tweedie (1993, Theorem 4.2). If $\eta > 1$ and $\alpha\eta < \beta(1 - p^\eta)$, the inequality (2.8) cannot be established as before, because $f(j) := (j+1)^\eta$ is no longer concave. However, as for (2.18), we can use the bound (2.16), valid for this choice of f for $|y - x| \leq c(\eta)x$ and for all x large enough, for some $c(\eta) > 0$. Using Chebyshev's inequality to bound any contribution from $|Y_j - jp_j| > c(\eta)jp_j$, taking expectations in (2.16) results in the bound

$$\mathbb{E}\{f(Y_j)\} - f(j) \leq f(\mathbb{E}Y_j) - f(j) + O(j^{\eta-\gamma}) = (j+1)^\eta(p^\eta - 1 + O(j^{-\gamma})),$$

which is the final inequality in (2.8) once again. The remainder of the argument is as before. Note that, if $\eta < \eta^*$, then $\alpha\eta < \beta(1 - p^\eta)$.

For $\eta > \eta^*$, so that $\alpha\eta > \beta(1 - p^\eta)$, entirely similar arguments can be used to show that

$$(Qf)(j) = f(j)\{\alpha\eta + \beta(p^\eta - 1) + O(j^{-\gamma})\},$$

and hence that

$$(Qf)(j) \geq c'_\eta f(j), \quad \text{where } c'_\eta := \tfrac{1}{2}\{\alpha\eta - \beta(1 - p^\eta)\} > 0,$$

for all $j \geq j_0$, for some j_0 large enough. Defining $f_0(j) := \max\{f(j_0), f(j)\}$, it thus follows that

$$(Qf_0)(j) \geq c'_\eta f_0(j), \quad j \geq j_0; \qquad (Qf_0)(j) \geq 0, \quad \text{otherwise.}$$

Now, with $f(j) := (j+1)^\eta$, the conditions of Hamza and Klebaner (1995, Theorem 2) are satisfied by f_0, implying that $f_0(X_t^*) - \int_{t_0}^t (Qf_0)(X_s^*)\, ds$ is a martingale in $t \geq t_0$, for any t_0. If $\mathbb{E}f_0(X_{t_0}^*) < \infty$, this in turn implies that $\mathbb{E}f_0(X_t^*)$ is non-decreasing in $t \geq t_0$, and that $\mathbb{E}f_0(X_t^*) > \mathbb{E}f_0(X_{t_0}^*)$ for $t > t_0$ if, with positive probability, $X_s^* \geq j_0$ on a subset of $[t_0, t]$ of positive Lebesgue measure; and this is the case if $\mathbb{P}[X_{t_0}^* \geq j_0] > 0$, because X^* is a pure jump Markov process. Now the stationary distribution of X^* assigns positive probability to all states, so that this condition is satisfied if $X_{t_0}^*$ has the stationary distribution, in which case X_t^* also has the stationary distribution, and $\mathbb{E}f_0(X_t^*) = \mathbb{E}f_0(X_{t_0}^*)$ if either one is finite. Thus $\mathbb{E}f_0(X_{t_0}^*) < \infty$ is not possible, and the η-th moment of the stationary distribution of X^* is infinite. ∎

3. The path to infinity; almost sure convergence

In view of Corollary 2.3, the event $\mathcal{X}$, on which $X_t \to \infty$, has positive probability if $\alpha > \beta \log(1/p)$. We now investigate how X_t approaches infinity, when it does so. If β_k took the value zero for all k, and if $\alpha_k = \alpha$ were the same for all k, then the process X would be a Yule process with *per capita* birth rate α. In this case, $e^{-\alpha t}X_t$ would converge almost surely to a positive limit W. The presence of catastrophes, when $\beta > 0$, changes the behaviour of X substantially. Nonetheless, there is an

analogous almost sure convergence theorem for the bivariate process (X, Z), which, for constant α and β and binomial thinning, would be a consequence of Kaplan et al. (1975, Theorem 4.1) and Bühler and Puri (1989, Theorem 4.1). The theorem shows that, on $\{\lim_{t\to\infty} X_t = \infty\}$, X_t behaves like $e^{\alpha t} p^{Z'_t} W$, where W is a positive random variable and Z' is a Poisson process of rate β. Throughout the section, we assume that

$$\nu := \alpha - \beta \log(1/p) > 0; \tag{3.1}$$

for the basic process, $\nu = \alpha - (1 - q) \log(1/p)$.

Theorem 3.1. *If $\nu > 0$, then $W_t := e^{-\alpha t} p^{-Z_t} X_t$ has an almost sure limit as $t \to \infty$, which we denote by W. Furthermore, $\mathcal{X} = \{W > 0\}$ almost surely.*

The process $(W_t, t \geq 0)$ is the counterpart, in the process with catastrophes, of the usual non-negative martingale appearing in the Markov branching process, though, in our setting, it is only approximately a martingale. Just as for branching processes, the random variable W is essentially determined by the early fluctuations of the process. It is easier to understand the behaviour by writing the convergence in the form

$$\log X_t - \alpha t \approx -Z_t \log(1/p) + \log W_s, \tag{3.2}$$

for s and t large, but $s \ll t$. This is an approximation only because $\log W_s$ is not exactly equal to $\log W_t$, but the two are close, for s and t large, if $W > 0$. Thus the process $\log X_t - \alpha t$, instead of being asymptotically constant on $\{X_t \to \infty\}$, as in the case of the Yule process, has downward jumps of magnitude $\log(1/p)$, approximately according to a Poisson process of rate β, staying very close to the lattice $\log W + \log(1/p)\mathbb{Z}$. The origin of the lattice is essentially determined by W_s, and the development after s is approximately given by

$$(\log X_t - \alpha t) - (\log X_s - \alpha s) \approx -(Z_t - Z_s) \log(1/p), \tag{3.3}$$

where $Z_t - Z_s$ is close to a Poisson process of rate β that is independent of X_s.

The proofs of the counterparts of Theorem 3.1 in Kaplan et al. (1975, Theorem 4.1) and Bühler and Puri (1989, Theorem 4.1) use the fact that, under their assumptions, there is a Galton–Watson process with random environments embedded in the process, enabling techniques from branching processes to be invoked. In our setting, this is no longer the case. Instead, we proceed in a succession of intermediate steps, establishing that $\log X_t$ develops in almost deterministically linear fashion between catastrophes, provided that X_t is large enough. This is already close to (3.2). The remainder of the argument consists of tidying up the details.

We begin by considering the bivariate process (X, Z) starting a time t_0 in an arbitrary state (j_0, k_0). Define

$$T_0 := t_0; \qquad T_n := \inf\{t > t_0 \colon Z_t = k_0 + n\}, \tag{3.4}$$

so that T_n is the time at which Z makes its n-th jump after t_0. Define $X_{[i]} := X_{T_i}$. With γ as in (2.4), and for $i \geq 1$, define the following events:

$$A_i := A_i(j_0, t_0) := \left\{ \sup_{T_{i-1} \leq t < T_i} \left| \log X_t - \alpha(t - T_{i-1}) - \log X_{T_{i-1}} \right| \leq X_{T_{i-1}}^{-\gamma/4} \right\};$$

$$B_i := B_i(j_0, t_0) := \{ |\log X_{T_i} - \log p - \log X_{T_i-}| \leq X_{T_i-}^{-\gamma/4} \}. \tag{3.5}$$

We begin by showing that the events A_i and B_i typically hold, if $X_{[i-1]}$ is large, as is eventually the case if $X_t \to \infty$.

Lemma 3.2. *With γ as in (2.4), and for $i \geq 1$, the conditional probabilities of A_i^c and B_i^c can be bounded as follows:*

$$(1): \quad \mathbb{P}[A_i^c(j_0, t_0) \,|\, \mathcal{F}_{T_{i-1}}] \leq C_A X_{[i-1]}^{\gamma/2-1} \leq C_A X_{[i-1]}^{-\gamma/2};$$

$$(2): \quad \mathbb{P}[B_i^c(j_0, t_0) \,|\, \sigma(T_i \vee \mathcal{F}_{T_i-})] \leq C_B X_{[i-1]}^{-\gamma/2}.$$

Here, C_A and C_B are (finite) constants depending only on the sequences $(\alpha_j, p_j, \sigma_j^2, j \geq 1)$.

Proof: From time T_{i-1} to time T_i, X evolves as a pure birth process, with jump rate $j\alpha_j$ when in state j. Define $Y_t^{(i)} = X_t$ for $T_{i-1} \leq t < T_i$, set $Y_{T_i}^{(i)} := Y_{T_i-}^{(i)}$, and continue $Y_t^{(i)}$ for $t > T_i$ as a pure birth process with the same distribution, independently of the further evolution of X. With $S_0 := T_{i-1}$, let S_l be the time of the l-th jump of $Y^{(i)}$, and let

$$h(0) := 0; \qquad h(j) := \sum_{l=1}^{j} \frac{1}{l}, \qquad j \geq 1, \tag{3.6}$$

denote the harmonic numbers. Then, conditional on $\mathcal{F}_{T_{i-1}} \cap \{X_{[i-1]} = j\}$, it follows that

$$M_l := h(j+l-1) - h(j-1) - \sum_{k=1}^{l} \alpha_{j+k-1}(S_k - S_{k-1}), \qquad l \in \mathbb{N}_0, \tag{3.7}$$

can be written as the sum $\sum_{k=1}^{l} E_k'$, where

$$E_k' := (j+k-1)^{-1}(1 - E_k),$$

and $(E_k, k \in \mathbb{N})$ are independent standard exponential random variables. Note also that M_l can be written in the form

$$M_l = h(Y_{S_l}^{(i)} - 1) - h(Y_{T_{i-1}}^{(i)} - 1) - \int_{T_{i-1}}^{S_l} \alpha_{Y_u^{(i)}} \, du, \tag{3.8}$$

linking the development of the process $h(Y_t^{(i)})$ with time, since, in view of (2.3), the integral is close to $\alpha(S_l - T_{i-1})$ if $Y_{T_{i-1}}^{(i)}$ is large. Since $\mathrm{Var}\,(E_k') = (j+k-1)^{-2}$, it follows that, for each $l \geq 1$,

$$\mathrm{Var}\,(M_l \,|\, \mathcal{F}_{T_{i-1}}) = \sum_{k=1}^{l} \mathrm{Var}\,(E_k') \leq 2X_{[i-1]}^{-1}. \tag{3.9}$$

Hence, for any $a > 0$, by Kolmogorov's inequality,

$$\mathbb{P}[\sup_{l \geq 0} |M_l| > a \,|\, \mathcal{F}_{T_{i-1}}] \leq \frac{2}{a^2 X_{[i-1]}}. \tag{3.10}$$

This inequality controls the fluctuations of $(h(Y_{S_l}^{(i)} - 1) - h(Y_{T_{i-1}}^{(i)} - 1), l \geq 1)$ around the values defined by the integrals $(\int_{T_{i-1}}^{S_l} \alpha_{Y_u^{(i)}} \, du, l \geq 1)$.

The next step is to bound the differences

$$\Delta_l := \int_{T_{i-1}}^{S_l} \alpha_{Y_u^{(i)}} \, du - \alpha(S_l - T_{i-1}), \qquad l \geq 1.$$

We have, for any $l \geq 1$,

$$|\mathbb{E}\{\Delta_l \,|\, \mathcal{F}_{T_{i-1}}\}| = \left| \sum_{j=X_{[i-1]}}^{X_{[i-1]}+l-1} j^{-1}(1 - \alpha/\alpha_j) \right| \leq C_1(\Delta) X_{[i-1]}^{-\gamma};$$

$$\mathrm{Var}\,(\Delta_l \,|\, \mathcal{F}_{T_{i-1}}) = \sum_{j=X_{[i-1]}}^{X_{[i-1]}+l-1} j^{-2}(1 - \alpha/\alpha_j)^2 = O(X_{[i-1]}^{-1-2\gamma}), \tag{3.11}$$

from which it also follows, using Kolmogorov's inequality, that

$$\mathbb{P}\left[\sup_{l \geq 0} |\Delta_l - \mathbb{E}\Delta_l| > a \,|\, \mathcal{F}_{T_{i-1}}\right] \leq \frac{C_2(\Delta)}{a^2 X_{[i-1]}^{1+2\gamma}}, \tag{3.12}$$

for any $a > 0$, where $C_l(\Delta)$, $l = 1, 2$, are constants depending on $(\alpha_j, j \geq 1)$. Hence, for any $a > 2C_1(\Delta)X_{T_{i-1}}^{-\gamma}$,

$$\mathbb{P}\left[\sup_{l \geq 0}\left|\int_{T_{i-1}}^{S_l} \alpha_{Y_u^{(i)}}\,du - \alpha(S_l - T_{i-1})\right| > a \,\big|\, \mathcal{F}_{T_{i-1}}\right] \leq \frac{4C_2(\Delta)}{a^2 X_{[i-1]}^{1+2\gamma}}. \tag{3.13}$$

In order to recover a bound for the probability of the event A_i from (3.10) and (3.13), we first observe that, by a simple integral comparison, for $l > k$, we have

$$h(l) - h(k) \leq \int_k^l x^{-1}\,dx = \log l - \log k \leq h(l-1) - h(k-1),$$

implying that $|(h(l-1) - h(k-1)) - (\log l - \log k)| \leq k^{-1}$. Then, because $Y^{(i)}$ is a step function,

$$\sup_{T_{i-1} \leq t \leq S_l} |h(Y_t^{(i)} - 1) - h(Y_{T_{i-1}}^{(i)} - 1) - \alpha(t - T_{i-1})|$$

is attained at one of the points $\{S_k-, S_k, \, 1 \leq k \leq l\}$, and $|h(Y_{S_k}^{(i)}) - h(Y_{S_k-}^{(i)})| \leq X_{[i-1]}^{-1}$. Hence

$$\sup_{t \geq t_0} |\log(Y_t^{(i)}) - \log(X_{[i-1]}) - \alpha(t - T_{i-1})| \leq \sup_{l \geq 0}|M_l + \Delta_l| + 2X_{[i-1]}^{-1}. \tag{3.14}$$

Combining (3.10), (3.13) and (3.14), and because $X_t = Y_t^{(i)}$ for $T_{i-1} \leq t < T_i$, it follows that, if $X_{[i-1]} \geq y_0$, where $y_0^{3\gamma/4} := 6\max\{C_1(\Delta), 1\}$, then

$$\mathbb{P}[A_i^c \,|\, \mathcal{F}_{T_{i-1}}] \leq \frac{2 + 4C_2(\Delta)}{X_{[i-1]}^{1-\gamma/2}}.$$

This establishes the first inequality, with $C_A := \max\{2 + 4C_2(\Delta), j_0^{1-\gamma/2}\}$.

The inequality (2) is much easier. We first note that

$$\mathbb{P}[B_i^c \,|\, \sigma(T_i \vee \mathcal{F}_{T_i-}) \cap \{X_{T_i-} = j\}] = \mathbb{P}[|\log Y_j - \log p - \log j| > j^{-\gamma/4}],$$

where $Y_j \sim \Pi_j$, as before. Now, by (1.1),

$$\mathbb{P}[|\log Y_j - \log p - \log j| > j^{-\gamma/4}] \leq \mathbb{P}[|\log Y_j - \log p_j - \log j| > \tfrac{1}{2}j^{-\gamma/4}]$$

if $\frac{1}{2}j^{-\gamma/4} > c_2 j^{-\gamma}$, and, since $|\log(1+x)| \leq 2|x|$ in $x \geq -1/2$,

$$\mathbb{P}[|\log Y_j - \log p_j - \log j| > \tfrac{1}{2}j^{-\gamma/4}] \leq \mathbb{P}[|(Y_j/jp_j) - 1| > \tfrac{1}{4}j^{-\gamma/4}] \leq \frac{4c_3}{p_j^2}\,j^{-\gamma/2},$$

where the last inequality follows by Chebyshev's inequality, and because of (1.1). Inequality (2) now follows directly. ∎

The implication of Lemma 3.2 is that, if $X_{[i-1]}$ is large, then $A_i \cap B_i$ occurs with high probability, in which case $\log X_t$ stays close to $\alpha(t - T_{i-1}) + \log X_{[i-1]}$ for $T_{i-1} \leq t < T_i$, and $\log X_{[i]}$ is close to $\alpha(T_i - T_{i-1}) + \log X_{[i-1]} - \log(1/p)$. If this is the case for all $1 \leq j \leq i$, then $\log X_t$ stays close to $\alpha(t - T_0) - (Z_t - Z_{T_0})\log(1/p) + \log X_{[0]}$ for $0 \leq t \leq T_i$. Now $\mathbb{E}\{Z_t - Z_{T_0}\} \approx \beta(t - T_0)$, and so $\alpha(t - T_0) - \mathbb{E}\{Z_t - Z_{T_0}\}\log(1/p) \approx \nu(t - T_0)$, where ν is as in (3.1). If $\nu > 0$, this suggests that $\log X_t$ grows linearly with t, or, equivalently, that X_t grows exponentially with t, on the event

$$H := H(j_0, t_0) := \bigcap_{i \geq 1}\{A_i(j_0, t_0) \cap B_i(j_0, t_0)\}.$$

The remainder of the proof consists of making this heuristic precise.

We begin by investigating the times T_j defined in (3.4), again for the process (X, Z) starting a time t_0 in an arbitrary state (j_0, k_0). Let j' be such that, for all $j \geq j'$,

$$\alpha - \beta_j \log(1/p) \; > \; \nu/2 \; > \; 0; \tag{3.15}$$

write $\beta' := \sup_{k \geq j'} \beta_k$. Observe that the process $(T_j, \, j \geq 1)$ is a point process adapted to the filtration $(\mathcal{F}_t, \, t \geq t_0)$, with compensator $A(t) := \int_{t_0}^t \beta_{X_u} \, du$. Hence, as long as $X_u \geq j'$ for all u, it can be coupled to a point process $(T_j', \, j \geq 1)$ with compensator $A'(t) := \beta'(t - t_0)$, hence a Poisson process with rate β', in such a way that $T_j - T_{j-1} \geq T_j' - T_{j-1}'$ a.s. for all $j \geq 1$, where we take $T_0 := T_0' := t_0$. More precisely, taking a sequence $(U_j, \, j \geq 1)$ of independent uniform $U[0, 1]$ random variables, independent of everything else, define

$$T_j' - T_{j-1}' \; := \; -\frac{1}{\beta'} \log U_j \; = \; \inf\{t \geq 0 \colon \beta' t \geq -\log U_j\};$$

$$T_j - T_{j-1} \; := \; \inf\left\{ t \geq 0 \colon \int_{T_{j-1}}^t \beta_{X_u} \, du \geq -\log U_j \right\}, \tag{3.16}$$

so that the times of the Z-jumps in (X, Z) are defined in terms of the sequence $(U_j, \, j \geq 1)$; the times of the remaining transitions in (X, Z) are defined using an independent sequence $(U_j', \, j \geq 1)$ of uniform random variables. On the event $X_{[j-1]} \geq j'$, we then have $T_j - T_{j-1} \geq T_j' - T_{j-1}'$, since X is non-decreasing in $[T_{j-1}, T_j)$.

For any $\eta > 0$ and with γ as in (2.4), define the events

$$\mathcal{E}_i^\eta \; := \; \mathcal{E}_i^\eta(j_0, t_0) \; := \; \left\{ \min_{1 \leq j \leq i} \{\alpha(T_j' - t_0) - j(\log(1/p) + \eta) + \log j_0\} > (4/\gamma) \log(2/\eta) \right\};$$

$$H_i \; := \; H_i(j_0, t_0) \; := \; \bigcap_{j=1}^i \big(A_j(j_0, t_0) \cap B_j(j_0, t_0) \big). \tag{3.17}$$

The next step is to show that, on the event $\mathcal{E}_i^\eta(j_0, t_0) \cap H_i(j_0, t_0)$, the values of $X_{[i]} = X_{T_i}$ can be bounded below in terms of the times T_i'.

Lemma 3.3. *If $\nu > 0$, choose any $\eta > 0$ such that $(2/\eta)^{4/\gamma} > j'$, where j' is as defined in (3.15). Then, for any $j_0 > (2/\eta)^{4/\gamma}$, on the event $\mathcal{E}_i^\eta(j_0, t_0) \cap H_i(j_0, t_0)$, it follows that*

$$2X_{[j]}^{-\gamma/4} \; < \; \eta \quad \text{and} \quad \log X_{[j]} \; \geq \; \alpha(T_j' - t_0) - j(\log(1/p) + \eta) + \log j_0 \tag{3.18}$$

for all $0 \leq j \leq i$.

Proof: The proof runs by induction on j. By assumption, $2X_{[0]}^{-\gamma/4} = 2j_0^{-\gamma/4} < \eta$, and the second inequality is trivial for $j = 0$. If (3.18), with $j - 1$ for j, is true for some j, then, from the definitions (3.5) of $A_j(j_0, t_0)$ and $B_j(j_0, t_0)$,

$$\log X_{[j]} - \log X_{[j-1]} \; \geq \; \alpha(T_j - T_{j-1}) - \log(1/p) - 2X_{[j-1]}^{-\gamma/4}$$
$$\geq \; \alpha(T_j' - T_{j-1}') - (\log(1/p) + \eta);$$

here, we have used $2X_{[j-1]}^{-\gamma/4} < \eta$, so that $X_{[j-1]} \geq j'$; and then also the fact that X is non-decreasing on $[T_{j-1}, T_j)$, so that $\beta_{X_u} \leq \beta'$ for all $T_{j-1} \leq u < T_j$. Adding this inequality to the second induction hypothesis gives

$$\log X_{[j]} \; \geq \; \alpha(T_j' - t_0) - j(\log(1/p) + \eta) + \log j_0 \; > \; (4/\gamma) \log(2/\eta),$$

in view of $\mathcal{E}_i^\eta(j_0, t_0)$, implying that $2X_{[j]}^{-\gamma/4} < \eta$ also. ∎

The lower bound established in Lemma 3.3 is useful for establishing that the values $X_{[j]}$ grow geometrically fast with j for $1 \leq j \leq i$, on the event $\mathcal{E}_i^\eta(j_0, t_0) \cap H_i(j_0, t_0)$, provided that the distribution of the times $(T_j', j \geq 1)$ can be controlled. This is made possible by the next lemma.

Lemma 3.4. *Let $(E_i, i \geq 1)$ be independent standard exponential random variables, and let $V_i := \sum_{j=1}^i E_j$.*

(i). *For any $0 < \phi < 1$, let $u(\phi)$ be the positive solution to $u\phi = \log(1 + u)$. Then, for any $a > 0$,*

$$\mathbb{P}[\min_{j \geq 1}\{V_j - j\phi\} \leq -a] \leq e^{-au(\phi)}.$$

(ii). *For any $0 < \phi < 1$ and $c, a > 0$, there exist $k = k(\phi, c) > 0$ and $0 < \psi = \psi(\phi, c) < 1$ such that*

$$\sum_{i \geq i_0} \mathbb{E}\{\min(1, e^{-c(V_i - i\phi + a)})\} \leq k\psi^{\max\{i_0, a\}}, \quad i_0 \geq 1.$$

Hence, in particular,

$$\sum_{i \geq 1} \min(1, e^{-c(V_i - i\phi + a)}) < \infty \text{ a.s.}$$

Proof: For part (i), it is immediate that the sequence $\widehat{M}_j^u := e^{-uV_j + j\log(1+u)}$ is a martingale for any $u > -1$. Writing $\sigma(a) := \inf\{j \geq 0 : V_j - j\phi \leq -a\}$, it thus follows that, for any $n \geq 1$,

$$1 = \mathbb{E}\{\widehat{M}_{\sigma(a) \wedge n}^{u(\phi)}\} = \mathbb{E}\{e^{-u(\phi)(V_{\sigma(a) \wedge n} - (\sigma(a) \wedge n)\phi)}\} \geq \mathbb{P}[\sigma(a) \leq n]e^{au(\phi)},$$

from which part (i) follows.

For part (ii), observe first that, for any $\phi' \in (\phi, 1)$, we have

$$\mathbb{E}\{\min(1, e^{-c(V_i - i\phi + a)})\} \leq \mathbb{P}[V_i \leq i\phi' - a/2] + \mathbb{E}\{e^{-c(V_i - i\phi + a)}I[V_i > i\phi' - a/2]\}$$
$$\leq \mathbb{P}[V_i \leq i\phi' - a/2] + e^{-ci(\phi' - \phi) - ca/2}. \tag{3.19}$$

If $i\phi' \leq a/2$, $\mathbb{P}[V_i \leq i\phi' - a/2] = 0$. For $i\phi' > a/2$, take $u_2(\phi') := \frac{1}{2}u(\phi')$, for which

$$v(\phi') := \log(1 + u_2(\phi')) - \phi' u_2(\phi') > 0,$$

beacuse the function $\log(1 + u)$ is strictly concave, and $\log(1 + u) = u\phi'$ for $u = 0$ and $u(\phi')$. Then

$$1 = \mathbb{E}\{\widehat{M}_i^{u_2(\phi')}\} \geq \mathbb{E}\{\exp\{-u_2(\phi')V_i + i\log(1 + u_2(\phi'))\}I[V_i \leq i\phi' - a/2]\}$$
$$\geq e^{au_2(\phi')/2 + iv(\phi')}\mathbb{P}[V_i \leq i\phi' - a/2].$$

Hence

$$\sum_{i \geq i_0} \mathbb{P}[V_i \leq i\phi' - a/2] \leq e^{-au_2(\phi')/2} \frac{e^{-i_0 v(\phi')}}{1 - e^{-v(\phi')}}. \tag{3.20}$$

Since also

$$\sum_{i \geq i_0} e^{-ci(\phi' - \phi) - ca/2} \leq e^{-ca/2} \frac{e^{-i_0 c(\phi' - \phi)}}{1 - e^{-c(\phi' - \phi)}},$$

the second part follows from (3.19). ∎

The next step is to use Lemmas 3.2–3.4 to conclude that, by choosing j_0 large enough, the probability $\mathbb{P}[H(j_0, t_0) \cap \mathcal{E}^\eta(j_0, t_0)]$ can be made arbitrarily close to 1; here, $\mathcal{E}^\eta(j_0, t_0) := \bigcap_{i \geq 1} \mathcal{E}_i^\eta(j_0, t_0)$.

Lemma 3.5. *Assume that $\nu := \alpha - \beta\log(1/p) > 0$, and choose any $0 < \theta < 1$ and $\eta > 0$ such that*

$$\eta \leq \frac{\theta\nu'}{\beta'} \quad \text{and} \quad \left(\frac{2}{\eta}\right)^{4/\gamma} > j', \tag{3.21}$$

where $\nu' := \alpha - \beta' \log(1/p) > 0$, and β' and j' are as in (3.15). Then

$$\mathbb{P}[H(j_0, t_0) \cap \mathcal{E}^\eta(j_0, t_0)] \geq 1 - C_0 j_0^{-\delta_0},$$

for $0 < \delta_0, C_0 < \infty$ depending only on the sequences $(\alpha_j, \beta_j, p_j, \sigma_j^2)$ and θ.

Proof: First, we show that $\mathbb{P}[(\mathcal{E}^\zeta(j_0, t_0))^c]$ is small if j_0 is large, for all $\zeta \leq \eta$. Writing $\phi_\zeta := \beta'(\log(1/p) + \zeta)/\alpha \leq \phi_\eta < 1$, the process $(\alpha(T'_j - t_0) - j(\log(1/p) + \zeta), j \geq 0)$ is that of the partial sums of independent random variables $(\alpha/\beta')(E_j - \phi_\zeta)$, where $(E_j, j \geq 1)$ is a sequence of independent exponential random variables with mean 1. Hence, by Lemma 3.4 (i), for any $a > 0$,

$$\mathbb{P}[\widetilde{M}_\zeta \leq -a] \leq e^{-au(\phi_\zeta)\beta'/\alpha}, \quad \text{where} \quad \widetilde{M}_\zeta := \min_{j \geq 1}\{\alpha(T'_j - t_0) - j(\log(1/p) + \zeta)\}.$$

Hence

$$\mathbb{P}[(\mathcal{E}^\zeta(j_0, t_0))^c] = \mathbb{P}[\widetilde{M}_\zeta \leq (4/\gamma)\log(2/\zeta) - \log j_0] \leq \left\{\left(\frac{\zeta}{2}\right)^{4/\gamma} j_0\right\}^{-u(\phi_\zeta)\beta'/\alpha}. \quad (3.22)$$

Next, we express $\mathbb{P}[H(j_0, t_0)^c \cap \mathcal{E}^\eta(j_0, t_0)]$ in the form

$$\mathbb{P}[H^c \cap \mathcal{E}^\eta] = \mathbb{P}\left[\left(\bigcap_{i \geq 1} H_i\right)^c \cap \mathcal{E}^\eta\right]$$

$$= \mathbb{P}[(A_1 \cap B_1)^c \cap \mathcal{E}^\eta] + \sum_{i \geq 1} \mathbb{P}[H_i \cap (A_{i+1} \cap B_{i+1})^c \cap \mathcal{E}^\eta]. \quad (3.23)$$

Now, because $\mathcal{E}^\eta \subset \mathcal{E}_i^\eta$ for each i, we have

$$\mathbb{P}[H_i \cap (A_{i+1} \cap B_{i+1})^c \cap \mathcal{E}^\eta] \leq \mathbb{P}[H_i \cap (A_{i+1} \cap B_{i+1})^c \cap \mathcal{E}_i^\eta]$$

$$= \mathbb{E}\{\mathbb{P}[H_i \cap (A_{i+1} \cap B_{i+1})^c \cap \mathcal{E}_i^\eta \,|\, \mathcal{F}_{T_i}]\}, \quad (3.24)$$

and it is immediate that

$$\mathbb{P}[H_i \cap (A_{i+1} \cap B_{i+1})^c \cap \mathcal{E}_i^\eta \,|\, \mathcal{F}_{T_i}] = \mathbb{P}[(A_{i+1} \cap B_{i+1})^c \,|\, \mathcal{F}_{T_i}]I[H_i \cap \mathcal{E}_i^\eta]. \quad (3.25)$$

From Lemma 3.2 and both inequalities in Lemma 3.3, it thus follows that

$$\mathbb{P}[H_i \cap (A_{i+1} \cap B_{i+1})^c \cap \mathcal{E}_i^\eta \,|\, \mathcal{F}_{T_i}] \leq (C_A + C_B)X_{[i]}^{-(\gamma/2)}I[H_i \cap \mathcal{E}_i^\eta]$$

$$\leq (C_A + C_B)\min\{1, \exp\{-(\gamma/2)(\alpha(T'_i - t_0) - i(\log(1/p) + \eta) + \log j_0)\}\}, \quad (3.26)$$

where we have used $j' \geq 1$. Combining (3.24) and (3.26), it follows that

$$\mathbb{P}[H_i \cap (A_{i+1} \cap B_{i+1})^c \cap \mathcal{E}^\eta] \leq (C_A + C_B)\mathbb{E}\{1 \wedge e^{-c(V_i - i\phi_\eta + a_0)}\}, \quad (3.27)$$

where V_i and ϕ_a are as before,

$$c := \frac{\alpha\gamma}{2\beta'} > 0 \quad \text{and} \quad a_0 := \frac{\beta' \log j_0}{\alpha}.$$

Applying Lemma 3.4 (ii), it follows from (3.23) and (3.27) that

$$\mathbb{P}[H^c \cap \mathcal{E}^\eta] \leq (C_A + C_B)k(\phi_\eta, c)\{\psi(\phi_\eta, c)\}^{a_0}; \quad (3.28)$$

taking $\zeta = \eta$ in (3.22), together with (3.28) and the definition of a_0, proves the lemma. ∎

With these preparations, it is now possible to prove Theorem 3.1.

Proof of Theorem 3.1. For any $j \in \mathbb{N}$, let $T_0 = \tau(j) := \inf\{t > 0 \colon X_t = j\} \leq \infty$. Then, defining $\mathcal{X}_j := \{\tau(j) < \infty\}$, we have $\mathbb{P}[H(j, \tau(j)) \cap \mathcal{E}^\eta(j, \tau(j)) \,|\, \mathcal{X}_j] \geq 1 - C_0 j^{-\delta_0}$, by Lemma 3.5 and the strong Markov property. On the event $H(j, \tau(j)) \cap \mathcal{E}^\eta(j, \tau(j))$, for any $r \geq 0$, we have

$$\sup_{t \geq T_r} |\log X_t - \alpha(t - T_r) + (Z_t - Z_{T_r})\log(1/p) - \log X_{[r]}| \;\leq\; 2\sum_{l \geq r} X_{[l]}^{-\gamma/4}, \qquad (3.29)$$

and hence, recalling that $W_t := e^{-\alpha t} p^{-Z_t} X_t$, it follows that

$$\sup_{s,t \geq T_r} |\log W_t - \log W_s| \;\leq\; 4\sum_{l \geq r} X_{[l]}^{-\gamma/4}. \qquad (3.30)$$

Then, by Lemma 3.3, on the event $H(j, \tau(j)) \cap \mathcal{E}^\eta(j, \tau(j))$, we have

$$X_{[i]} \;\geq\; \max\{1, \, j e^{\alpha(T_i' - T_0) - i(\log(1/p) + \eta)}\} \quad \text{for all } i \geq 0, \qquad (3.31)$$

and this, in turn, from Lemma 3.4, implies that $\sum_{i \geq 1} X_{[i]}^{-\gamma/4} < \infty$ a.s., and hence that, on the event $H(j, \tau(j)) \cap \mathcal{E}^\eta(j, \tau(j))$,

$$\lim_{r \to \infty} \sum_{i \geq r} X_{[i]}^{-\gamma/4} \;=\; 0 \quad \text{a.s.} \qquad (3.32)$$

Thus, on the event

$$A(j) \;:=\; \mathcal{X}_j \cap H(j, \tau(j)) \cap \mathcal{E}^\eta(j, \tau(j)),$$

we conclude from (3.30) and (3.32), using the Cauchy criterion, that $\log W_t$ converges a.s. to a finite limit, implying that W_t converges a.s. to a strictly positive limit W.

Now, for $X_0 = x_0$, $\mathcal{X} = \bigcap_{j > x_0} \mathcal{X}_j$. Writing $\mathbb{P}_j[\cdot]$ for $\mathbb{P}[\cdot \,|\, X_0 = j]$, this implies that $\mathbb{P}_{x_0}[\mathcal{X}] = \lim_{j \to \infty} \mathbb{P}_{x_0}[\mathcal{X}_j]$, and, as above, $A(j) \subset \mathcal{X}_j$ and, using Lemma 3.5,

$$0 \;\leq\; \mathbb{P}_{x_0}[\mathcal{X}_j] - \mathbb{P}_{x_0}[A(j)] \;\leq\; \mathbb{P}_{x_0}[\mathcal{X}_j] C_0 j^{-\delta_0} \;\to\; 0 \quad \text{as } j \to \infty.$$

Hence, if $C := \{W_t \text{ converges to a positive limit}\}$, we have $A(j) \subset C$ for each j, and

$$\limsup_{j \to \infty} \mathbb{P}_{x_0}[A(j)] \;=\; \lim_{j \to \infty} \mathbb{P}_{x_0}[\mathcal{X}_j] \;=\; \mathbb{P}_{x_0}[\mathcal{X}].$$

Hence $\mathbb{P}_{x_0}[C] \geq \mathbb{P}_{x_0}[\mathcal{X}]$. However, $\mathcal{X}^c = \{X_t = 0 \text{ eventually}\}$, on which event $W_t \to 0$ a.s. Hence $C = \mathcal{X}$ a.s., and the theorem is proved.

4. The path to infinity: the central limit theorem

In view of Theorem 3.1, and its interpretation in (3.3), the main fluctuations in $\log X_t$, for large t, are governed by those of a Poisson process. In keeping with this interpretation, we now deduce a central limit theorem for $\log X_t$, which is based on the central limit theorem for the Poisson distribution. As a consequence, for large t, that part of the distribution of X_t that is not the mass on zero has an approximately log-normal distribution.

Theorem 4.1. *Under the assumptions of Theorem 3.1, for any $x_0 \in \mathbb{N}$, the process X with $X_0 = x_0$ satisfies*

$$\lim_{t \to \infty} \mathbb{P}_{x_0}[t^{-1/2}\{\log X_t - \nu t\} \geq y] \;=\; \mathbb{P}_{x_0}[\mathcal{X}](1 - \Phi(y/v)),$$

where $v^2 := \beta\{\log(1/p)\}^2$ and ν is as in (3.1).

Proof: From the definition of W_t, we have

$$\log X_t \;=\; \alpha t - Z_t \log(1/p) - \log W_t,$$

so that, for any $0 < s < t$,

$$\log X_t - \nu t = -(Z_t - \beta t)\log(1/p) - \log W_t \tag{4.1}$$
$$= -(Z_t - Z_s - \beta(t-s))\log(1/p) - (Z_s - \beta s)\log(1/p) - \log W_t.$$

The remainder of the proof consists of showing that, for suitable choice of s, the final two terms are asymptotically negligible, and that $Z_t - Z_s$ is approximately Poisson distributed with mean $\beta(t-s)$.

The last two terms in (4.1)
We start by showing that, if $t \gg s \gg 1$, the final two terms of (4.1) are small compared to $\sqrt{t}$ on a 'good event' that has probability close enough to $\mathbb{P}_{x_0}[\mathcal{X}]$. In preparation, choose any $0 < \varepsilon < \frac{1}{2}\beta$. With j' is as defined in (3.15), choose $j_0 := j_0(\varepsilon) \geq j'$ such that $c_1 j_0^{-\gamma_1} < \varepsilon$. Then choose $j_1 := j_1(\varepsilon) \geq (j_0(\varepsilon))^2$ so that, with $\mathcal{X}_j := \{\tau(j) < \infty\}$,

$$0 < \mathbb{P}_{x_0}[\mathcal{X}] - \mathbb{P}_{x_0}[\mathcal{X}_{j_1}] < \varepsilon, \tag{4.2}$$

this is possible, since $\mathcal{X} = \bigcap_{j > x_0} \mathcal{X}_j$. Choose

$$s := s(\varepsilon) > 0 \quad \text{such that} \quad \mathbb{P}_{x_0}[\tau(j_1(\varepsilon)) > s \mid \mathcal{X}_{j_1}] < \varepsilon;$$
$$z := z(\varepsilon) > 0 \quad \text{such that} \quad \mathbb{P}_{x_0}[Z_{s(\varepsilon)} > z \mid \tau(j_1) \leq s(\varepsilon)] < \varepsilon; \tag{4.3}$$
$$w := w(\varepsilon) > 0 \quad \text{such that} \quad \mathbb{P}_{x_0}[|\log W_{\tau(j_1)}| > w \mid \tau(j_1) \leq s(\varepsilon)] < \varepsilon,$$

noting that the quantities j_0, j_1, s, z and w depend on the choice of ε, but are then fixed. Let $A_\varepsilon^{(1)} := \{\tau(j_1(\varepsilon)) \leq s(\varepsilon)\}$, and define part of the 'good event' by

$$E_\varepsilon^{(1)} := A_\varepsilon^{(1)} \cap \{Z_{s(\varepsilon)}) \leq z(\varepsilon)\} \cap \{|\log W_{\tau(j_1(\varepsilon))}| \leq w(\varepsilon)\}; \tag{4.4}$$

on $E_\varepsilon^{(1)}$, the values of the process up to the time of first hitting the state $j_1(\varepsilon)$ are not extreme.

Take $T_0 := T_0(\varepsilon) := \tau(j_1(\varepsilon))$, and let the subsequent jumps of Z be denoted by $(T_i, i \geq 1)$; as before, note that T_0 itself need not be a jump time of Z. Couple them to the jump times $(T_i', i \geq 1)$ of a Poisson process Z' of rate β' as in (3.16), and, fixing $\eta > 0$ to satisfy (3.21), define

$$\widehat{\mathcal{E}}_i^\eta := \widehat{\mathcal{E}}_i^\eta(j_1, T_0) := \left\{ \min_{1 \leq j \leq i} \{\alpha(T_j' - T_0) - j(\log(1/p) + 3\eta/2)\} > -\tfrac{1}{2}\log j_1 \right\};$$
$$\widehat{\mathcal{E}}^\eta := \widehat{\mathcal{E}}^\eta(j_1, T_0) := \bigcap_{i \geq 1} \widehat{\mathcal{E}}_i^\eta(j_1, T_0). \tag{4.5}$$

The event

$$E_\varepsilon^{(2)} := H(j_1(\varepsilon), \tau(j_1(\varepsilon))) \cap \mathcal{E}^\eta(j_1(\varepsilon), \tau(j_1(\varepsilon))) \cap \widehat{\mathcal{E}}^\eta(j_1(\varepsilon), \tau(j_1(\varepsilon))), \tag{4.6}$$

is to be the second element of the 'good event'. On this event, using (3.31) and the definition of $\widehat{\mathcal{E}}^\eta$, it follows that

$$X_{[i]} \geq j_0 e^{i\eta/2} \quad \text{for all } i \geq 0. \tag{4.7}$$

Thus, on the event $\mathcal{X}_{j_1(\varepsilon)} \cap E_\varepsilon^{(2)}$, from (3.29) and (4.7),

$$\sup_{t \geq \tau(j_1(\varepsilon))} |\log X_t - \log j_1(\varepsilon) - \alpha(t - \tau(j_1(\varepsilon))) + (Z_t - Z_{\tau(j_1(\varepsilon))})\log(1/p)|$$

$$= \sup_{t \geq \tau(j_1(\varepsilon))} |\log W_t - \log W_{\tau(j_1(\varepsilon))}| \leq \frac{2}{(j_0(\varepsilon))^{\gamma/4}(1 - e^{-\eta\gamma/8})}, \tag{4.8}$$

implying that, on the event $\mathcal{X}_{j_1(\varepsilon)} \cap E_\varepsilon^{(1)} \cap E_\varepsilon^{(2)}$,

$$\sup_{t \geq \tau(j_1(\varepsilon))} |\log W_t| \leq w(\varepsilon) + \frac{2}{(j_0(\varepsilon))^{\gamma/4}(1 - e^{-\eta\gamma/8})}. \tag{4.9}$$

Hence, on the same event, for all $t \geq s(\varepsilon)$,

$$|Z_{s(\varepsilon)} - \beta s(\varepsilon)| \log(1/p) + |\log W_t| \leq (z(\varepsilon) + \beta s(\varepsilon)) \log(1/p) + w(\varepsilon) + \frac{2}{(j_0(\varepsilon))^{\gamma/4}(1 - e^{-\eta\gamma/8})} ; \quad (4.10)$$

thus, as $t \to \infty$ and with $s = s(\varepsilon)$, the contribution to (4.1) from the last two terms is asymptotically negligible compared to $\sqrt{t}$.

Note also that, from (4.8), we have

$$\sup_{s,t \geq \tau(j_1(\varepsilon))} |\log X_t - \log X_s - \alpha(t - s) + (Z_t - Z_s) \log(1/p)|$$

$$\leq 2 \sup_{t \geq \tau(j_1(\varepsilon))} |\log W_t - \log W_{\tau(j_1(\varepsilon))}| \leq \frac{4}{(j_0(\varepsilon))^{\gamma/4}(1 - e^{-\eta\gamma/8})}. \quad (4.11)$$

This shows that the development of $\log X_t - \alpha t$ after the time $\tau(j_1(\varepsilon))$ is mirrored in the development of $-Z_t$. Furthermore, on the event $\mathcal{X}_{j_1(\varepsilon)} \cap E_\varepsilon^{(2)}$ and from (1.1), for $i \geq 1$ and for $T_{i-1} \leq t < T_i$,

$$|\beta_{X_t} - \beta| \leq c_1(j_0(\varepsilon)e^{(i-1)\eta/2})^{-\gamma_1} \leq \varepsilon e^{-(i-1)\eta\gamma_1/2} =: \varepsilon_\beta(i). \quad (4.12)$$

For large times, Z is close to a Poisson process with constant rate

The next step is to show that, to approximate the development of $\log X_t - \alpha t$ after time $\tau(j_1(\varepsilon))$, the process Z can be replaced by a Poisson process of rate β, with only small error. Suppressing the dependence on ε in the notation, where possible, we set $T_0 := \tau(j_1)$, as before, and set $\widetilde{T}_0 := T_0$. We then use the independent uniform random variables $(U_i, i \geq 1)$ introduced in (3.16) to define the jump times $(\widetilde{T}_i, i \geq 1)$ of a further Poisson process $\widetilde{Z}$, of rate β, on the time interval (T_0, ∞), coupled to the jump times $(T_i, i \geq 1)$ of Z on this time interval. The coupling is achieved by defining $\widetilde{T}_i - \widetilde{T}_{i-1} := -\frac{1}{\beta} \log U_i$, with $T_i - T_{i-1}$ defined as in (3.16). It then follows from (4.12) that, on the event $\mathcal{X}_{j_1} \cap E^{(2)}(j_1)$,

$$|(\widetilde{T}_i - \widetilde{T}_{i-1}) - (T_i - T_{i-1})| \leq \frac{\varepsilon_\beta(i)}{\beta - \varepsilon_\beta(i)} (\widetilde{T}_i - \widetilde{T}_{i-1}) \leq \frac{2\varepsilon_\beta(i)}{\beta} (\widetilde{T}_i - \widetilde{T}_{i-1}),$$

because $\varepsilon < \frac{1}{2}\beta$, and hence that

$$|\widetilde{T}_i - T_i| \leq \sum_{l=1}^{i} \frac{2\varepsilon_\beta(l)}{\beta^2} (-\log U_l).$$

Because $\sum_{i \geq 1} \varepsilon_\beta(i) < \infty$, it follows that

$$T^* := \sum_{l \geq 1} \frac{2\varepsilon_\beta(l)}{\beta^2} (-\log U_l)$$

is an a.s. finite random variable, and that

$$\sup_{i \geq 1} |\widetilde{T}_i - T_i| \leq T^* \quad \text{a.s.;}$$

indeed,

$$\mathbb{E}T^* = \sum_{l \geq 1} \frac{2\varepsilon_\beta(l)}{\beta^2} \leq \frac{2\varepsilon}{\beta^2(1 - e^{-\eta\gamma/2})} =: \varepsilon t^*, \quad (4.13)$$

also. Thus, on the event $\mathcal{X}_{j_1} \cap E_\varepsilon^{(2)}$, for all $t \geq T^* + T_0 = T^* + \tau(j_1)$, we have

$$\widetilde{Z}_{t-T^*} - \widetilde{Z}_{\tau(j_1)} \leq Z_t - Z_{\tau(j_1)} \leq \widetilde{Z}_{t+T^*} - \widetilde{Z}_{\tau(j_1)}. \quad (4.14)$$

This in turn implies that

$$|(Z_t - Z_{\tau(j_1)}) - (\widetilde{Z}_t - \widetilde{Z}_{\tau(j_1)})| \leq \widetilde{Z}_{t+T^*} - \widetilde{Z}_{t-T^*} =: \widetilde{Z}_t^*, \quad (4.15)$$

bounding the error involved at time t, if Z is replaced by $\widetilde{Z}$. Note that $\mathbb{P}[T^* > 2t^*] \leq \frac{1}{2}\varepsilon$, by Markov's inequality, and that, on $\{T^* \leq 2t^*\}$,

$$\widetilde{Z}_t^* \leq \widetilde{Z}_{t+2t^*} - \widetilde{Z}_{t-2t^*} \sim \text{Po}\,(4t^*\beta) \quad \text{if} \quad t \geq 2t^* + \tau(j_1).$$

Hence, choosing $\tilde{z} := \tilde{z}(\varepsilon)$ in such a way that $\text{Po}\,(4t^*\beta)\{[\tilde{z}, \infty)\} \leq \frac{1}{2}\varepsilon$, it follows that, for all $t \geq s(\varepsilon) + 2t^*$,

$$\mathbb{P}_{x_0}[\widetilde{Z}_t^* > \tilde{z} \,|\, A_\varepsilon^{(1)}] < \varepsilon, \tag{4.16}$$

quantifying the error (4.15).

Approximation using Poisson probabilities
Let $\widehat{Z}_t := -(\widetilde{Z}_t - \beta t)$ denote minus the centred version of $\widetilde{Z}_t$. Then, for $t > s(\varepsilon)$ and on the 'good event'

$$A_{t,\varepsilon} := E_\varepsilon^{(1)} \cap E_\varepsilon^{(2)} \cap \{\widetilde{Z}_t^* \leq \tilde{z}(\varepsilon)\}, \tag{4.17}$$

we can use (4.3), (4.15) and (4.11) to give the following summary of the approximation of $\log X_t - \nu t$ using the centred Poisson random variable $\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)}$:

$$|(\log X_t - \nu t) - (\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p)|\, I[A_{t,\varepsilon}] \tag{4.18}$$

$$\leq \{z(\varepsilon) + s(\varepsilon)\beta + \tilde{z}(\varepsilon)\}\log(1/p) + w(\varepsilon) + \frac{4}{j_0(\varepsilon)^{\gamma/4}(1 - e^{-\eta/2})} =: r(\varepsilon).$$

Note that the 'natural' approximation to $\log X_t - \nu t$ would be $\widehat{Z}_t \log(1/p)$, but for (4.20) below we need the fact that, conditional on any event in $\mathcal{F}_{s(\varepsilon)} \cap A_\varepsilon^{(1)}$, the increment $\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)}$ has the centred version of the Poisson distribution with mean $\beta(t - s(\varepsilon))$; using this increment in the approximation leads to the term $\{z(\varepsilon) + s(\varepsilon)\beta\}\log(1/p)$ in the error bound $r(\varepsilon)$.

Using (4.18), it follows that, for any $y \in \mathbb{R}$, we have

$$\left(\{(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq y\sqrt{t} + r(\varepsilon)\} \cap A_\varepsilon^{(1)}\right) \setminus (A_{t,\varepsilon}^c \cap A_\varepsilon^{(1)})$$

$$\subset \{\log X_t - \nu t \geq y\sqrt{t}\} \cap A_{t,\varepsilon}$$

$$\subset \{(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq y\sqrt{t} - r(\varepsilon)\} \cap A_\varepsilon^{(1)},$$

and hence that

$$\mathbb{P}_{x_0}[\{(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq y\sqrt{t} + r(\varepsilon)\} \cap A_\varepsilon^{(1)}] - \mathbb{P}_{x_0}[A_{t,\varepsilon}^c \,|\, A_\varepsilon^{(1)}]$$

$$\leq \mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap A_{t,\varepsilon}] \tag{4.19}$$

$$\leq \mathbb{P}_{x_0}[\{(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq y\sqrt{t} - r(\varepsilon)\} \cap A_\varepsilon^{(1)}].$$

Because the increment $\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)}$ has the Poisson distribution $\text{Po}\,\big(\beta(t - s(\varepsilon))\big)$, conditional on any event in $\mathcal{F}_{s(\varepsilon)} \cap A_\varepsilon^{(1)}$, it also follows that, for any $x \in \mathbb{R}$,

$$\mathbb{P}_{x_0}[\{(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq x\} \cap A_\varepsilon^{(1)}] = \mathbb{P}_{x_0}[A_\varepsilon^{(1)}]\text{Po}\,\big(\beta(t - s(\varepsilon))\big)[\beta(t - s(\varepsilon)) + x/\log(1/p), \infty), \tag{4.20}$$

and this can be used on both the left and the right hand sides of (4.19), with only slightly different values of x.

A first normal approximation
We now convert the inequalities (4.19) into an explicit normal approximation to the probability $\mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap \mathcal{X}]$. First, by the Berry–Esseen theorem, if $t > 2s(\varepsilon)$, we have

$$|\mathbb{P}[(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq x\sqrt{t - s(\varepsilon)}] - (1 - \Phi(x/v))| \leq \frac{2C_\beta}{\sqrt{t}}, \quad x \in \mathbb{R},$$

where C_β is a finite constant and v^2 is as in the statement of the theorem. Thus

$$\mathbb{P}[(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq y\sqrt{t} - r(\varepsilon)] \ \leq \ 1 - \Phi\Big(\frac{y\sqrt{t} - r(\varepsilon)}{v\sqrt{t - s(\varepsilon)}}\Big) + \frac{2C_\beta}{\sqrt{t}}.$$

Then, from the properties of the normal distribution, and using $t > 2s(\varepsilon)$,

$$0 \ < \ \Phi\Big(\frac{y\sqrt{t}}{v\sqrt{t - s(\varepsilon)}}\Big) - \Phi\Big(\frac{y\sqrt{t} - r(\varepsilon)}{v\sqrt{t - s(\varepsilon)}}\Big) \ \leq \ \frac{r(\varepsilon)}{v\sqrt{\pi t}},$$

and

$$\Big|\Phi\Big(\frac{y\sqrt{t}}{v\sqrt{t - s(\varepsilon)}}\Big) - \Phi\Big(\frac{y}{v}\Big)\Big| \ \leq \ \frac{|y|e^{-(y/v)^2/2}}{v\sqrt{2\pi}}\Big\{\sqrt{\frac{t}{t - s(\varepsilon)}} - 1\Big\} \ \leq \ \frac{1}{\sqrt{2\pi e}}\frac{s(\varepsilon)}{t},$$

where we have used $0 \leq (1-x)^{-1/2} - 1 \leq x$ in $0 \leq x \leq 1/2$. Collecting these bounds, we have shown that

$$\mathbb{P}[(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq y\sqrt{t} - r(\varepsilon)] \ \leq \ (1 - \Phi(y/v)) + \delta(t, \varepsilon), \tag{4.21}$$

where

$$\delta(t, \varepsilon) \ := \ \frac{2C_\beta}{\sqrt{t}} + \frac{r(\varepsilon)}{v\sqrt{\pi t}} + \frac{1}{\sqrt{2\pi e}}\frac{s(\varepsilon)}{t}; \tag{4.22}$$

note that $\lim_{t\to\infty}\delta(t, \varepsilon) = 0$ for each $\varepsilon > 0$. An entirely similar argument shows that

$$\mathbb{P}[(\widehat{Z}_t - \widehat{Z}_{s(\varepsilon)})\log(1/p) \geq y\sqrt{t} + r(\varepsilon)] \ \geq \ (1 - \Phi(y/v)) - \delta(t, \varepsilon). \tag{4.23}$$

Hence, in view of (4.19), we have

$$|\mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap A_{t,\varepsilon}] - (1 - \Phi(y/v))\mathbb{P}_{x_0}[A_\varepsilon^{(1)}]| \ \leq \ \mathbb{P}_{x_0}[A_{t,\varepsilon}^c \,|\, A_\varepsilon^{(1)}] + \delta(t, \varepsilon). \tag{4.24}$$

The final normal approximation

It now remains to tidy up, replacing $A_{t,\varepsilon}$ and $A_\varepsilon^{(1)}$ by $\mathcal{X}$ in (4.24), and evaluating the various error terms. First, observe that

$$\mathcal{X} \ \subset \ \mathcal{X}_{j_1(\varepsilon)} \quad \text{and} \quad A_\varepsilon^{(1)} \ = \ \{\tau(j_1(\varepsilon)) \leq s(\varepsilon)\} \ \subset \ \mathcal{X}_{j_1(\varepsilon)},$$

so that, for any event B,

$$|\mathbb{P}_{x_0}[B \cap \mathcal{X}] - \mathbb{P}_{x_0}[B \cap A_\varepsilon^{(1)}]|$$
$$\leq \ (\mathbb{P}_{x_0}[\mathcal{X}_{j_1(\varepsilon)}] - \mathbb{P}_{x_0}[\mathcal{X}]) + (\mathbb{P}_{x_0}[\mathcal{X}_{j_1(\varepsilon)}] - \mathbb{P}_{x_0}[\tau(j_1(\varepsilon)) < s(\varepsilon)]) \ < \ 2\varepsilon, \tag{4.25}$$

where the last inequality follows from (4.2) and from the definition of $s(\varepsilon)$. In particular, taking B to be the certain event, this implies that

$$|\mathbb{P}_{x_0}[\mathcal{X}] - \mathbb{P}_{x_0}[A_\varepsilon^{(1)}]| \ < \ 2\varepsilon. \tag{4.26}$$

Thus, again using (4.25),

$$|\mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap A_{t,\varepsilon}] - \mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap \mathcal{X}]|$$
$$\leq \ |\mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap A_\varepsilon^{(1)}] - \mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap \mathcal{X}]|$$
$$+ \mathbb{P}_{x_0}[A_{t,\varepsilon}^c \,|\, A_\varepsilon^{(1)}]$$
$$\leq \ 2\varepsilon + \mathbb{P}_{x_0}[A_{t,\varepsilon}^c \,|\, A_\varepsilon^{(1)}].$$

Hence, and from (4.24) and (4.26),

$$|\mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap \mathcal{X}] - (1 - \Phi(y/v))\mathbb{P}_{x_0}[\mathcal{X}]| \ \leq \ 2\mathbb{P}_{x_0}[A_{t,\varepsilon}^c \,|\, A_\varepsilon^{(1)}] + \delta(t, \varepsilon) + 4\varepsilon. \tag{4.27}$$

To bound $\mathbb{P}_{x_0}[A_{t,\varepsilon}^c \,|\, A_\varepsilon^{(1)}]$, note that, from (4.17),

$$\mathbb{P}_{x_0}[A_{t,\varepsilon}^c \,|\, A_\varepsilon^{(1)}] \ \leq \ \mathbb{P}_{x_0}[(E_\varepsilon^{(1)})^c \,|\, A_\varepsilon^{(1)}] + \mathbb{P}_{x_0}[(E_\varepsilon^{(2)})^c \,|\, A_\varepsilon^{(1)}] + \mathbb{P}_{x_0}[\widetilde{Z}_t^* > \tilde{z}(\varepsilon) \,|\, A_\varepsilon^{(1)}].$$

From (4.4), we have $\mathbb{P}_{x_0}[(E_\varepsilon^{(1)})^c \,|\, A_\varepsilon^{(1)}] \leq 2\varepsilon$, and $\mathbb{P}_{x_0}[\widetilde{Z}_t^* > \tilde{z}(\varepsilon) \,|\, A_\varepsilon^{(1)}] < \varepsilon$, from (4.16), if $t > 2t^* + s(\varepsilon)$. It remains to bound $\mathbb{P}_{x_0}[(E_\varepsilon^{(2)})^c \,|\, A_\varepsilon^{(1)}]$. From (4.6), we have

$$\mathbb{P}_{x_0}[(E_\varepsilon^{(2)})^c \,|\, A_\varepsilon^{(1)}] \;\leq\; \mathbb{P}_{x_0}[(H(j_1(\varepsilon), \tau(j_1(\varepsilon))) \cap \mathcal{E}^\eta(j_1(\varepsilon), \tau(j_1(\varepsilon))))^c \,|\, A_\varepsilon^{(1)}]$$
$$+ \mathbb{P}_{x_0}[\{\widehat{\mathcal{E}}^\eta(j_1(\varepsilon), \tau(j_1(\varepsilon)))\}^c \,|\, A_\varepsilon^{(1)}],$$

and, from Lemma 3.5, we have

$$\mathbb{P}_{x_0}\big[\{H(j_1(\varepsilon), \tau(j_1(\varepsilon))) \cap \mathcal{E}^\eta(j_1(\varepsilon), \tau(j_1(\varepsilon)))\}^c \,|\, A_\varepsilon^{(1)}\big] \;\leq\; C_0 j_1^{-\delta_0}.$$

From the definition (4.5) of $\widehat{\mathcal{E}}^\eta(j_1, \tau(j_1))$,

$$\mathbb{P}_{x_0}[\{\widehat{\mathcal{E}}^\eta(j_1(\varepsilon), \tau(j_1(\varepsilon)))\}^c \,|\, A_\varepsilon^{(1)}] \;=\; \mathbb{P}\big[(\alpha/\beta')\inf_{i \geq 1}(V_i - i\phi') > -\tfrac{1}{2}\log j_1(\varepsilon)\big],$$

where $(V_i, \, i \geq 1)$ are as in Lemma 3.4, and where

$$\phi' \;:=\; \beta'(\log(1/p) + 3\eta/2)/\alpha \;<\; 1.$$

By an argument as for Lemma 3.4 (i), it follows that

$$\mathbb{P}[\{\widehat{\mathcal{E}}^\eta(j_1(\varepsilon), \tau(j_1(\varepsilon)))\}^c \,|\, A_\varepsilon^{(1)}] \;\leq\; \exp\{-(\tfrac{1}{2}\log j_1(\varepsilon))\beta' u(\phi')/\alpha\} \;=\; (j_1(\varepsilon))^{-\delta_1},$$

where $\delta_1 > 0$. Combining these bounds, we conclude that

$$\mathbb{P}_{x_0}[(E_\varepsilon^{(2)})^c \,|\, A_\varepsilon^{(1)}] \;\leq\; C_1(j_1(\varepsilon))^{-(\delta_0 \wedge \delta_1)}. \tag{4.28}$$

Hence

$$\mathbb{P}_{x_0}[(A_{t,\varepsilon})^c \,|\, A_\varepsilon^{(1)}] \;\leq\; 3\varepsilon + C_1(j_1(\varepsilon))^{-(\delta_0 \wedge \delta_1)} \qquad \text{if} \quad t > 2t^* + s(\varepsilon).$$

Thus, and from (4.27) and (4.22), it follows that

$$\limsup_{t \to \infty} |\mathbb{P}_{x_0}[\{\log X_t - \nu t \geq y\sqrt{t}\} \cap \mathcal{X}] - (1 - \Phi(y/v))\mathbb{P}_{x_0}[\mathcal{X}]| \tag{4.29}$$

can be made arbitrarily small by choice of ε, and is therefore equal to zero. Finally, if Y_t is any stochastic process such that $Y_t \to \infty$ a.s., then $\lim_{t \to \infty} \mathbb{P}[Y_t \leq M] = 0$ for all $M \in \mathbb{R}$. Applying this observation using $\mathbb{P}_{x_0}[\cdot \,|\, \mathcal{X}^c]$ as probability measure shows that, since $\log X_t \to -\infty$ on $\mathcal{X}^c$,

$$\limsup_{t \to \infty} \mathbb{P}_{x_0}[\{\log X_t \geq \nu t + y\sqrt{t}\} \cap \mathcal{X}^c] \;\leq\; \lim_{t \to \infty} \mathbb{P}_{x_0}[\log X_t \geq M \,|\, \mathcal{X}^c] \;=\; 0$$

for any $y, M \in \mathbb{R}$, completing the proof of the theorem. ∎

5. The degree weighted distribution

Throughout this section, we work with the process X under the restriction

$$\alpha_k \;=\; 1 - (1 - p_k)\beta_k \;=\; \alpha, \quad \text{for all } k \geq 1,$$

for some fixed $\alpha > 0$, which includes the generalized DD model of (1.3). If $\mathbf{e}$ is defined to be the vector with $e_k = k$ for $k \geq 0$, then $Q\mathbf{e} = (2\alpha - 1)\mathbf{e}$, implying that $\mathbf{e}$ is the $(1 - 2\alpha)$-invariant vector for Q on $\mathbb{N}$. Because of this, the matrix $\widetilde{Q}$ on $\mathbb{N}$ having elements

$$\widetilde{Q}_{k,k+1} \;=\; \alpha(k+1); \qquad \widetilde{Q}_{k,k} \;=\; -\{-1 + k\alpha + \beta_k(1 - \pi_{kk}) + 2\alpha\};$$
$$\widetilde{Q}_{k,j} \;=\; p_k \beta_k \frac{j\pi_{kj}}{kp_k}, \quad 1 \leq j \leq k-1; \qquad \widetilde{Q}_{k,j} \;=\; 0, \quad j \geq k+2, \tag{5.1}$$

for each $k \geq 2$, and with $\widetilde{Q}_{1,2} = -\widetilde{Q}_{1,1} = 2\alpha$, is a Q-matrix; indeed,

$$\widetilde{Q}_{i,j} = (Q_{i,j} + (1 - 2\alpha)\delta_{ij})j/i \quad \text{for } i, j \in \mathbb{N}. \tag{5.2}$$

Note that, from the definition of p_k, $\pi_{11} = p_1$, and $1 - (1 - p_1)\beta_1 = \alpha$, so that the general formula for $\widetilde{Q}_{k,k}$ in (5.1), when evaluated for $k = 1$, still gives the correct value -2α. The form of the

matrix $\widetilde{Q}$ shows that the pure jump Markov process $\widetilde{X} := (\widetilde{X}_t, t \in \mathbb{R}_+)$ on $\mathbb{N}$ corresponding to $\widetilde{Q}$ is dominated by a Yule process with rate 2α, and is thus non-explosive. Hence we can deduce the following proposition; as observed in Jordan (2018), in the case where $\alpha = p_k = p$ and $\Pi_k = \mathrm{Bi}\,(k, p)$, this follows from Pollett (1988, Lemma 3.3) when $2\alpha < 1$.

Proposition 5.1. *The point probabilities for the processes X and $\widetilde{X}$ are related by*

$$j\mathbb{P}_i[X_t = j] \;=\; e^{-(1-2\alpha)t}\, i\, \mathbb{P}_i[\widetilde{X}_t = j] \quad \text{for all } i, j \in \mathbb{N} \text{ and } t \geq 0. \tag{5.3}$$

Proof: Because X is a non-explosive pure jump Markov process, the Kolmogorov forward equations

$$\frac{d}{dt}\mathbb{P}_i[X_t = j] \;=\; \sum_{l \geq 0} \mathbb{P}_i[X_t = l]Q_{l,j}$$

are satisfied. Hence, writing $x_t^{(1)}(j) := e^{(1-2\alpha)t}\,(j/i)\,\mathbb{P}_i[X_t = j]$, it follows that

$$\frac{d}{dt}x_t^{(1)}(j) \;=\; \sum_{l \geq 0} x_t^{(1)}(l)(j/l)Q_{l,j} + (1-2\alpha)x_t^{(1)}(j) \;=\; \sum_{l \geq 0} x_t^{(1)}(l)\widetilde{Q}_{l,j},$$

this last from (5.2). Hence $x_t^{(1)}$ satisfies the forward equations for the process $\widetilde{X}$. On the other hand, by Hamza and Klebaner (1995, Theorem 2), the process

$$\mathbf{e}(X_t) - \int_0^t (Q\mathbf{e})(X_u)\,du \;=\; X_t - \int_0^t (2\alpha - 1)X_u\,du$$

is a martingale, implying that $m^{(1)}(t) := \mathbb{E}\{X_t \mid X_0 = i\}$ satisfies

$$m^{(1)}(t) \;=\; i + (2\alpha - 1)\int_0^t m^{(1)}(u)\,du, \qquad \text{or} \qquad m^{(1)}(t) \;=\; ie^{(2\alpha-1)t}.$$

Hence, for all $t \geq 0$,

$$\sum_{l \geq 0} x_t^{(1)}(l) \;=\; e^{(1-2\alpha)t}\sum_{l \geq 0}(l/i)\mathbb{P}_i[X_t = l] \;=\; e^{(1-2\alpha)t}i^{-1}m^{(1)}(t) \;=\; 1.$$

Thus $x_t^{(1)}$ is a probability distribution for each t, and hence is the unique probability solution to the forward equations for $\widetilde{X}$ having initial condition the point mass on i; that is,

$$x_t^{(1)}(j) \;=\; e^{(1-2\alpha)t}\,(j/i)\,\mathbb{P}_i[X_t = j] \;=\; \mathbb{P}_i[\widetilde{X}_t = j],$$

proving the proposition. ∎

The matrix $\widetilde{Q}$ has the same form as that given in (2.1), with $\tilde{\alpha}_k := \alpha(k+1)/k$ and $\tilde{\beta}_k := p_k\beta_k$, and with $\widetilde{\Pi}_k$ defined by $\tilde{\pi}_{kj} := j\pi_{kj}/kp_k$, and in particular with $\tilde{\pi}_{11} = 1$. Note that the quantities $\tilde{\alpha}_k$ are not the same for each k, but that differing α_k were allowed for in (2.1). The only effect of not having $\tilde{\pi}_{11} < 1$ is to make $\mathbb{N}$ a closed class, and to remove the absorbing state at 0. The mean and variance of $\widetilde{\Pi}_k$ are given by

$$k\tilde{p}_k \;:=\; \sum_{j=1}^k j\tilde{\pi}_{kj} \;=\; \sum_{j=1}^k \frac{j^2\pi_{kj}}{kp_k} \quad \text{and} \quad \tilde{\sigma}_k \;:=\; \sum_{j=1}^k \frac{j^3\pi_{kj}}{kp_k} - (k\tilde{p}_k)^2.$$

If the assumptions in (2.2) and (2.3) are satisfied for X (the condition on α_k is automatic, since in this section $\alpha_k = \alpha$ for all k), then

$$|\tilde{\beta}_k - p\beta| \;\leq\; \tilde{c}_1 k^{-\gamma} \quad \text{and} \quad |\tilde{\alpha}_k - \alpha| \;\leq\; \tilde{c}_4 k^{-1}, \tag{5.4}$$

so that the assumptions (2.3) are satisfied for $\widetilde{X}$ with $\tilde{\beta} := p\beta$ and $\tilde{\alpha} = \alpha$. That assumptions (2.2) are satisfied for $\widetilde{X}$, with $\gamma/4$ for γ, is established in the following lemma.

Lemma 5.2. *Let $\widetilde{Y}_k \sim \widetilde{\Pi}_k$. Then there are positive constants $\tilde{c}_2$ and $\tilde{c}_3$*

$$|\tilde{p}_k - p| \;\leq\; \tilde{c}_2 k^{-\gamma} \quad and \quad k^{-2}\tilde{\sigma}_k^2 \;\leq\; \tilde{c}_3 k^{-\gamma/4}, \tag{5.5}$$

Proof: It follows by direct calculation that

$$k\tilde{p}_k \;=\; \mathbb{E}\widetilde{Y}_k \;=\; kp_k + \sigma_k^2/(kp_k),$$

from which, and from (2.2), the first assertion follows. To prove the second assertion, it is enough to show that $(kp_k)^{-2}\mathbb{E}\widetilde{Y}_k^2 \leq 1 + O(k^{-\gamma/4})$, because $\tilde{\sigma}_k^2 = \mathbb{E}(\widetilde{Y}_k^2) - (\mathbb{E}\widetilde{Y}_k)^2$, and $\mathbb{E}\widetilde{Y}_k = kp_k + O(k^{1-\gamma})$, from (2.2). Letting $\varepsilon_k := k^{-\gamma/4}$ and $Y_k \sim \Pi_k$, it follows from Chebyshev's inequality that

$$
\begin{aligned}
(kp_k)^{-2}\mathbb{E}\widetilde{Y}_k^2 \;&=\; (kp_k)^{-3}\sum_{j=1}^{k} j^3 \pi_{kj} \\
&\leq\; (kp_k)^{-3}\big\{(1+\varepsilon_k)^3(kp_k)^3 + k^3\mathbb{P}[Y_k > (1+\varepsilon_k)kp_k]\big\} \\
&\leq\; (1+\varepsilon_k)^3 + O(k^{-\gamma/2}) \\
&=\; 1 + O(k^{-\gamma/4}),
\end{aligned}
$$

which completes the proof. ∎

As a result of the inequalities (5.4) and of Lemma 5.2, we can now apply the theorems of the previous sections to the process $\widetilde{X}$. Recall that, as in Theorem 2.5, for any $u > 1$, $x^*(u) < u$ denotes the positive solution to the equation $x = u(1 - e^{-x})$.

Theorem 5.3. *For the degree weighted continuous time process $\widetilde{X}$, we have the following asymptotic behaviour:*

(1) *If $\alpha < p\beta\log(1/p)$, then $\widetilde{X}$ is geometrically ergodic, and the stationary distribution has η-th moment finite for all $\eta < \eta^* := x^*(p\beta\log(1/p)/\alpha)/\log(1/p) < p\beta/\alpha$, and infinite for all $\eta > \eta^*$;*

(2) *If $\alpha = p\beta\log(1/p)$, then $\widetilde{X}$ is null recurrent;*

(3) *if $\alpha > p\beta\log(1/p)$, then $\widetilde{X}$ is transient.*

In case 1, $\mathbb{P}_i[\widetilde{X} = j] \to \tilde{p}_j(\infty)$ for each $j \geq 1$, where $\tilde{p}(\infty)$ is a probability distribution on $\mathbb{N}$. In view of Proposition 5.1, it follows that

$$e^{(1-2\alpha)t}\mathbb{P}_i[X_t = j] \;=\; (i/j)\mathbb{P}_i[\widetilde{X}_t = j] \;\to\; (i/j)\tilde{p}_j(\infty) \quad \text{as } t \to \infty,$$

so that the probabilities $\mathbb{P}_i[X_t = j]$ all decay exponentially in time, with the same rate. Furthermore, since the function $j \mapsto j^{-1}$ is bounded on $\mathbb{N}$,

$$e^{(1-2\alpha)t}\mathbb{P}_i[X_t \in A] \;=\; i\sum_{j\in A} j^{-1}\mathbb{P}_i[\widetilde{X}_t = j] \;\to\; i\sum_{j\in A} j^{-1}\tilde{p}_j(\infty), \tag{5.6}$$

and hence

$$\mathbb{P}_i[X_t = j \mid X_t \geq 1] \;=\; \frac{j^{-1}\mathbb{P}_i[\widetilde{X}_t = j]}{\sum_{l\geq 1} l^{-1}\mathbb{P}_i[\widetilde{X}_t = l]} \;\to\; \frac{j^{-1}\tilde{p}_j(\infty)}{\sum_{l\geq 1} l^{-1}\tilde{p}_l(\infty)}, \tag{5.7}$$

showing that then the conditional distribution of X_t, given $X_t \geq 1$, converges, and identifying the limit. In the basic model, with $q = 0$, (5.7) is Jordan (2018, Proposition 3.4). and the exponential tail in (5.6) is in Hermann and Pfaffelhuber (2020, Corollary 2.7(1)).

In the basic model, in which $\alpha = p(1-q) + q$ and $\beta = 1 - q$, for any choices of p and q such that $0 < p < 1$ and $0 \leq q < 1$, the behaviour given in Theorems 2.1 and 5.3 can be categorized according to the cases

$$\text{(A):} \;\; \alpha < p(1-q)\log(1/p), \qquad \text{(B):} \;\; p(1-q)\log(1/p) \leq \alpha < (1-q)\log(1/p)$$
$$\text{and} \;\; \text{(C):} \;\; \alpha > (1-q)\log(1/p);$$

see Figure 5.1. For fixed p, these cases can be represented by (A): $q < q_2(p)$, (B): $q_2(p) < q < q_1(p)$ and (C): $q > q_1(p)$, where q_1 is the inverse of the function p_* solving (2.19):

$$q_1(p) \;=\; \frac{\log(1/p) - p}{1 + \log(1/p) - p} \;>\; q_2(p) \;=\; \frac{p\log(1/p) - p}{1 + \log(1/p) - p},$$

and where $q_2(p) < 0$ if $p > e^{-1}$ and $q_1(p) < 0$ if $p > p_*(0) \approx 0.5671$.
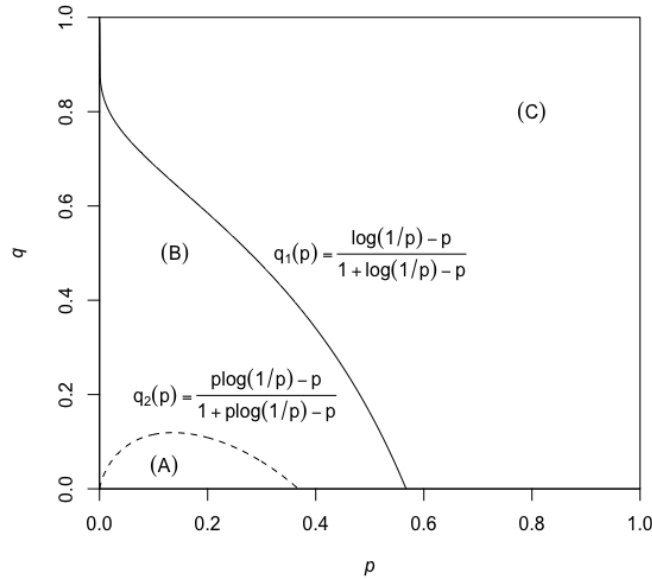

FIGURE 5.1. A $p - q$ plane. The labelled areas correspond to the cases above.

In particular, $q_2(0) = q_2(e^{-1}) = 0$, and the maximum of q_2 is given by $q_2(e^{-2}) = 1/(e^2 + 1) \approx 0.1192$. Case (A) holds for $0 < p < e^{-1}$ and $0 < q < q_2(p)$, and in this case (5.6) and (5.7) hold. In case (B), the process X is absorbed in zero with probability one, but the degree weighted process is transient. In case (C), the process X is transient and the central limit theorem of Theorem 4.1 can be used to approximate its (time-dependent) distribution.

When the process $\widetilde{X}$ is transient, Theorem 3.1 and 4.1 can be applied to it, in view of the inequalities (5.4) and of Lemma 5.2. However, because of the degree weighting, their conclusions no longer give information about the *typical* degree distribution.

6. Discrete time processes

6.1. *The DD model.* We now return to the analysis of the original probabilities $\mathfrak{p}_m^\top$, $m \in \mathbb{Z}_+$, that can be derived as the state probabilities of an inhomogeneous Markov chain $Y := (Y_m, m \in \mathbb{Z}_+)$, using (1.5). Note first that, given that the chain Y has a jump at step $(m + 1)$, the probability distribution of Y_{m+1}, given $Y_m = k$, depends only on the value k, and is the same for all m; the distribution is given by the transition probabilities out of the state k of the *jump chain* associated with the homogeneous process X in continuous time. Hence the Markov chain Y can be constructed from a realization of the jump chain of X, together with an independent realization of a sequence $(U_j, j \geq 1)$ of independent standard uniform random variables, which are used to determine the residence times in the successive states of the chain. In particular, properties such as recurrence and transience, which depend only on the sequence of states visited, can be deduced for Y from

the corresponding properties for X. However, the factors $(m+1)^{-1}$ in the definition of the jump probabilities at time m imply that the distribution of the residence times of Y in a given state k themselves depend on the time m at which k was reached, and that the residence times, as measured by the number of steps in k, become much longer as m increases. Instead, by compressing the time scale, a close analogue $Y^{(h)}$ of the homogeneous process X can be derived, as follows.

Let $(\widehat{X}_j, j \geq 0)$ be a realization of the jump chain of X. Then a version of X can be obtained from $\widehat{X}$ and an independent sequence of mutually independent standard uniform random variables $(U_j, j \geq 1)$ by setting

$$X_t = \widehat{X}_n \quad \text{if} \quad S_n := \sum_{j=1}^{n} V_j \leq t < S_{n+1}, \quad t \geq 0, \tag{6.1}$$

where, with $\hat{q}_k := -Q_{k,k}$, $V_j := \hat{q}_{\widehat{X}_{j-1}}^{-1} \{-\log(1-U_j)\}$ represents the j-th residence time of X, the residence time following its $(j-1)$-st jump, and has an exponential distribution with mean $\hat{q}_{\widehat{X}_{j-1}}^{-1}$. Note that, if $\widehat{X}_{j-1} = 0$, then $V_j = \infty$ a.s., because $q_0 = 0$. A similar construction can be used to define $Y^{(h)}$. Let N_j represent the step at which the j-th jump of Y occurs, with $N_0 = m_0$. For any $a > -1$, define the generalized harmonic numbers $h_a(j)$, $j \geq 1$, by $h_a(j) := \sum_{l=1}^{j} 1/(l+a)$, so that h_0 represents the harmonic numbers h defined in (3.6). Let $\widetilde{V}_j := h_0(N_j) - h_0(N_{j-1})$ represent the j-th residence time of Y, but in an almost logarithmically distorted time scale; again, if $\widehat{X}_{j-1} = 0$, then $\widetilde{V}_j = \infty$ a.s. Then set

$$Y_t^{(h)} := \widehat{X}_n \quad \text{if} \quad \widetilde{S}_n := \sum_{j=1}^{n} \widetilde{V}_j \leq t < \widetilde{S}_{n+1}, \quad t \geq 0. \tag{6.2}$$

From (6.1) and (6.2), it is immediate that the paths of X and $Y^{(h)}$ are close to one another if the partial sum processes S and $\widetilde{S}$ defined from the V_j and the $\widetilde{V}_j$ are close to one another. Indeed, defining

$$\Delta_0 := 0; \qquad \Delta_n := S_n - \widetilde{S}_n, \quad n \geq 1,$$

and then setting

$$\Delta(t) := \frac{\widetilde{S}_{n+1} - t}{\widetilde{S}_{n+1} - \widetilde{S}_n} \Delta_n + \frac{t - \widetilde{S}_n}{\widetilde{S}_{n+1} - \widetilde{S}_n} \Delta_{n+1} \quad \text{if} \quad \widetilde{S}_n \leq t < \widetilde{S}_{n+1}, \qquad t \geq 0, \tag{6.3}$$

with $\Delta(t) := \Delta_n$ if $t \geq \widetilde{S}_n$ and $\widetilde{S}_{n+1} = \infty$, we have

$$Y_t^{(h)} = X_{t+\Delta(t)}, \quad t \geq 0, \qquad \text{and also} \qquad Y_m = Y_{h_0(m)-h_0(m_0)}^{(h)}, \quad m \geq m_0 \in \mathbb{Z}_+. \tag{6.4}$$

In this way, the process $Y^{(h)}$ is represented as a random time shift of the process X, and the process Y can be deduced from $Y^{(h)}$ by a deterministic time change. We now show that the time shift connecting $Y^{(h)}$ and X is not substantial: in fact,

$$\Delta := \lim_{t \to \infty} \Delta(t) \quad \text{exists a.s.} \tag{6.5}$$

To establish (6.5), we need to examine the distribution of the random variables $\widetilde{V}_j$. From the definition of the process Y, conditional on the values of the whole of $\widehat{X}$ and on $\{N_l, 1 \leq l \leq j-1\}$, and on the event $\{(N_{j-1}, \widehat{X}_{j-1}) = (m, k)\}$, $\widetilde{V}_j$ has the distribution of $V_0(m, \hat{q}_k)$, where $V_a(m, b)$ denotes a random variable taking values in the discrete set $\{h_a(j) - h_a(m), j \geq m+1\}$ with

probabilities given by

$$\mathbb{P}[V_a(m,b) > h_a(r) - h_a(m)] \;=\; \prod_{l=m}^{r-1}(1 - b/(l+a+1)), \qquad r \geq m. \tag{6.6}$$

The next lemma shows that if E_b denotes a random variable having an exponential distribution with mean $1/b$, random variables with the distributions of $V_a(m,b)$ and E_b can be constructed on the same probability space, in such a way that they are close.

Lemma 6.1. *For $V_a(m,b)$ defined as above, $\mathbb{E}V_a(m,b) = 1/b$. Furthermore, if E_b, as defined above, is coupled to $V_a(m,b)$ using the quantile coupling, and if $b/(m+a+1) \leq \phi < 1$, then there exists a positive constant c_ϕ such that, for all $j \geq m+1$,*

$$-\frac{1}{j+a} \;\leq\; E_b - V_a(m,b) \;\leq\; c_\phi\,\frac{b(h_a(j) - h_a(m))}{m+a+1}, \tag{6.7}$$

almost surely on the event $\{V_a(m,b) = h_a(j) - h_a(m)\}$. As a consequence, there exists a positive constant k_ϕ such that

$$\mathbb{E}\{(E_b - V_a(m,b))^2\} \;\leq\; \frac{k_\phi}{(m+a+1)^2}\,.$$

Remark 6.2. The quantile coupling can be achieved by using a standard uniform random variable U: take $E_b := -b^{-1}\log(1-U)$ and

$$V_a(m,b) \;=\; h_a(r+1) - h_a(m) \quad\text{if}\quad \prod_{l=m}^{r}(1-b/(l+a+1)) \;\leq\; 1-U \;<\; \prod_{l=m}^{r-1}(1-b/(l+a+1)),$$

for $r \geq m$.

Proof: For any non-negative non-decreasing sequence $(f_j, j \geq 1)$ and any probability distribution $(p_j, j \geq 1)$ on $\mathbb{N}$, it is immediate that, if $\sum_{j\geq 1}p_j f_j < \infty$, then

$$\sum_{j\geq 1}p_j f_j \;=\; \sum_{j\geq 1}P_j(f_j - f_{j-1}),$$

where $P_j := \sum_{l\geq j}p_l$, and $f_0 := 0$. Hence, since $h_a(j+1) - h_a(j) = 1/(j+a+1)$, it follows that

$$\begin{aligned}
\mathbb{E}V_a(m,b) \;&=\; \sum_{j\geq m}\frac{1}{j+a+1}\prod_{l=m}^{j-1}(1-b/(l+a+1))\\
&=\; \frac{1}{b}\sum_{j\geq m}\Big\{\prod_{l=m}^{j-1}(1-b/(l+a+1)) - \prod_{l=m}^{j}(1-b/(l+a+1))\Big\}\\
&=\; \frac{1}{b}\,.
\end{aligned}$$

For the quantile coupling, first observe that

$$\prod_{l=m}^{j-1}(1-b/(l+a+1)) \;=\; \exp\Big\{\sum_{l=m}^{j-1}\log(1-b/(l+a+1))\Big\},$$

from which it follows that

$$e^{-b(h_a(j)-h_a(m))} \geq \mathbb{P}[V_a(m,b) > h_a(j) - h_a(m)] \tag{6.8}$$

$$= e^{-b(h_a(j)-h_a(m))} \exp\left\{\sum_{l=m}^{j-1}\left\{\log\left(1 - \frac{b}{l+a+1}\right) + \frac{b}{l+a+1}\right\}\right\}$$

$$= e^{-b(h_a(j)-h_a(m)+\psi_a(j,b))} = \mathbb{P}[E_b > h_a(j) - h_a(m) + \psi_a(j,b)], \tag{6.9}$$

say. Because $b \leq \phi(m+a+1)$ and $\log(1-x) > -x - c_\phi x^2$ in $0 < x < \phi$, for some constant $c_\phi > 0$, it is immediate that

$$\psi_a(j,b) = -\frac{1}{b}\sum_{l=m}^{j-1}\left\{\log\left(1 - \frac{b}{l+a+1}\right) + \frac{b}{l+a+1}\right\}$$

$$\leq \frac{c_\phi}{b}\sum_{l=m}^{j-1}\frac{b^2}{(l+a+1)^2} \leq c_\phi\frac{b(h_a(j)-h_a(m))}{m+a+1}.$$

Hence, using the quantile coupling, if $V_a(m,b) \leq h_a(j) - h_a(m)$, it follows that

$$E_b \leq h_a(j) - h_a(m) + \psi_a(j,b) \leq h_a(j) - h_a(m) + c_\phi\frac{b(h_a(j)-h_a(m))}{m+a+1},$$

and, in particular, on $\{V_a(m,b) \leq h_a(j) - h_a(m)\}$, the second inequality in (6.7) now follows. The first inequality in (6.7) follows similarly as a consequence of (6.8), applied with $j-1$ for j.

For the final statement, it follows from (6.7) and (6.8) that

$$\mathbb{E}\{(E_b - V_a(m,b))^2\}$$

$$\leq \sum_{j \geq m+1}\mathbb{P}[V_a(m,b) = h_a(j) - h_a(m)]\max\left\{\frac{1}{(m+a+1)^2}, \left(c_\phi\frac{b(h_a(j)-h_a(m))}{m+a+1}\right)^2\right\}$$

$$\leq \frac{1}{(m+a+1)^2} + \sum_{j \geq m+1}\frac{b}{j+a}e^{-b(h_a(j-1)-h_a(m))}\left(c_\phi\frac{b(h_a(j)-h_a(m))}{m+a+1}\right)^2$$

$$\leq \frac{1}{(m+a+1)^2} + \sum_{j \geq m}\frac{2eb}{j+a+1}e^{-b(h_a(j)-h_a(m))}\left(c_\phi\frac{b(h_a(j)-h_a(m))}{m+a+1}\right)^2.$$

Now, for a unimodal function $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ and for a sequence $0 = x_0 < x_1 < \cdots$, with $x_{j+1} - x_j \leq x_j - x_{j-1}$ for all $j \geq 1$, it follows by an elementary comparison that

$$\sum_{j \geq 0}(x_{j+1} - x_j)f(x_j) \leq \int_0^\infty f(x)\,dx + 2(x_1 - x_0)\sup_{x \geq 0}f(x).$$

Hence

$$\sum_{j \geq m}\frac{2eb}{j+a+1}e^{-b(h_a(j)-h_a(m))}\left(c_\phi\frac{b(h_a(j)-h_a(m))}{m+a+1}\right)^2$$

$$\leq \frac{2ec_\phi^2}{(m+a+1)^2}\left\{\int_0^\infty x^2e^{-x}\,dx + 8\phi e^{-2}\right\},$$

and the final part of the lemma follows. ∎

This enables the partial sum processes defined from the V_j and the $\widetilde{V}_j$ to be constructed so as to be close to one another.

Corollary 6.3. *Conditional on the whole of the jump chain $\widehat{X}$, the random variables $(\widetilde{V}_j,\, j \geq 1)$ and $(V_j,\, j \geq 1)$ can be coupled in such a way that*

$$\lim_{n\to\infty} \Delta_n \;=\; \lim_{t\to\infty} \Delta(t) \;=\; \Delta$$

exists a.s., with the convergence holding also in mean square.

Proof: In view of the constructions (6.1) and (6.2), it is enough to use the coupling in Remark 6.2 with $a = 0$ successively, with a sequence of independent standard uniform random variables $(U_j,\, j \geq 1)$, that are also independent of the jump chain $\widehat{X}$. The corollary then follows because $(\Delta_r,\, r \geq 0)$ is a square integrable martingale, as is implied by the variance bound given in Lemma 6.1, and because $\Delta(\cdot)$ interpolates the values of $(\Delta_r,\, r \geq 0)$. Note that, under the assumptions of Section 1, $\hat{q}_k \leq (m+1)\phi$ for all $k \leq m$, where $\phi := \frac{1}{2}(1+\alpha) < 1$, and that necessarily $\widehat{X}_j \leq N_j$ for all j, because upward jumps are of magnitude 1. ∎

As a result of Corollary 6.3, it is possible to transfer results about the long term behaviour of X into corresponding results about $Y^{(h)}$, and hence about Y. When X is absorbed in zero with probability one, as in Corollary 2.3, the same is true for $Y^{(h)}$ and Y, but this is already clear, because they have the same jump chain as X. If the process X is transient, then there is an almost sure growth theorem for Y that can be deduced from Theorem 3.1. To state it, let J_m denote the number of 'downward' jumps of Y up to time $m \geq m_0$, and let $J_t^{(h)}$ denote the number of downward jumps of $Y^{(h)}$ up to time t, where a downward jump is a jump that increases Z by one in the process (X, Z); since the jump chains of the three processes $X, Y^{(h)}$ and Y are identical, their downward jumps can also be defined to be the same. Note that $J_m = J_{h_0(m)-h_0(m_0)}^{(h)}$ for all $m \geq m_0$.

Theorem 6.4. *If $\alpha > \beta \log(1/p)$, let $W := \lim_{t\to\infty} e^{-\alpha t} p^{-Z_t} X_t$ be the almost sure limit given in Theorem 3.1. Then $m^{-\alpha} p^{-J_m} Y_m$ converges a.s. as $m \to \infty$ to a limit W' such that*

$$W' \;=_d\; e^{\alpha\{\Delta+(\log m_0-h_0(m_0)+\gamma)\}} W,$$

where $\Delta := \lim_{t\to\infty} \Delta(t)$, and γ is Euler's constant. Note that W and Δ are dependent random variables, and that $\lim_{m_0\to\infty}(\log m_0 - h_0(m_0) + \gamma) = 0$.

Proof: From the definition of W, it is immediate that

$$e^{-\alpha(t+\Delta(t))} p^{-Z_{t+\Delta(t)}} X_{t+\Delta(t)} \;\to\; W \quad \text{a.s. as } t \to \infty.$$

Then, because $Y^{(h)}$ and X can be constructed together in such a way that $Y_t^{(h)} = X_{t+\Delta(t)}$ for all $t \geq 0$, it follows that

$$e^{-\alpha(t+\Delta(t))} p^{-J_t^{(h)}} Y_t^{(h)} \;\to\; W \quad \text{a.s. as } t \to \infty.$$

Taking $t = h_0(m) - h_0(m_0)$ for $m \geq m_0$, $m \in \mathbb{N}$, and because $\Delta(t) \to \Delta$ a.s. as $t \to \infty$, this implies that

$$e^{-\alpha(h_0(m)-h_0(m_0))} p^{-J_m} Y_m \;\to\; e^{\alpha\Delta} W \quad \text{a.s. as } s \to \infty,$$

and the theorem follows, because $\lim_{m\to\infty}\{h_0(m) - \log m\} = \gamma$. ∎

The logarithm of the inhomogeneous discrete time process Y also satisfies a central limit theorem.

Theorem 6.5. *If $\nu := \alpha - \beta \log(1/p) > 0$, for any $x_0 \in \mathbb{N}$, the process Y satisfies*

$$\lim_{m\to\infty} \mathbb{P}[\{\log m\}^{-1/2}\{\log Y_m - \nu \log m\} \geq y \,|\, Y_{m_0} = x_0] \;=\; \mathbb{P}_{x_0}[\mathcal{X}](1 - \Phi(y/v)),$$

where $v^2 := \beta\{\log(1/p)\}^2$ and $\mathbb{P}_{x_0}[\mathcal{X}] := \mathbb{P}[\lim_{t\to\infty} X_t = \infty \,|\, X_0 = x_0]$.

Proof: We begin by proving the central limit theorem for $Y^{(h)}$, arguing using the coupling $Y_t^{(h)} = X_{t+\Delta(t)}$, $t \geq 0$, and deducing the theorem from Theorem 4.1.

Writing $\log X_t = \alpha t - Z_t \log(1/p) - \log W_t$, where $\log W_t \to \log W > -\infty$ a.s. on $\mathcal{X}$, and $\log W_t \to -\infty$ a.s. on $\mathcal{X}^c$, we have

$$|\log X_{t+\Delta(t)} - \log X_t| \leq \alpha|\Delta(t)| + \log(1/p)|Z_{t+\Delta(t)} - Z_t| + |\log W_{t+\Delta(t)} - \log W_t|. \qquad (6.10)$$

For $s \geq 1$ and $t \geq 2(s+1)$, define the events

$$A_s^{(1)} := \{|\Delta(u) - \Delta| \leq 1 \text{ for all } u \geq s\}; \qquad A_t^{(3)} := \{|Z_{t+t^{1/8}} - Z_{t-t^{1/8}}| \leq t^{1/4}\};$$

$$A_t^{(2)} := \{|\Delta| \leq t^{1/8}\}; \qquad A_s^{(4)} := \{|\log W_u - \log W_v| \leq 1 \text{ for all } u, v \geq s\}. \qquad (6.11)$$

The first three of these events are shown to have probabilities approaching 1 as $s, t \to \infty$, and the fourth event approaches the event $\mathcal{X}$ as $s \to \infty$. On $A_s^{(1)} \cap A_t^{(2)} \cap A_s^{(4)}$, with $s \geq 1$ and $t \geq 2(s+1)$, so that $t - t^{1/8} - 1 > s$, it follows that $|\log W_{t+\Delta(t)} - \log W_t| \leq 1$. Hence, if also $A_t^{(3)}$ holds, then, using (6.10),

$$|\log X_{t+\Delta(t)} - \log X_t| \leq \alpha(t^{1/8} + 1) + \log(1/p)t^{1/4} + 1 =: r(t);$$

here, we have used the fact that Z_u is a.s. non-decreasing in u. Note that $t^{-1/2}r(t) \to 0$ as $t \to \infty$, so that, on the event $\mathcal{E}_{s,t} := A_s^{(1)} \cap A_t^{(2)} \cap A_t^{(3)} \cap A_s^{(4)}$, the difference between $\log X_{t+\Delta(t)}$ and $\log X_t$ is negligible as far as the CLT is concerned. We exploit this as follows.

First, since $\log X_t \to -\infty$ a.s. as $t \to \infty$ on $\mathcal{X}^c$, it follows also that $\log X_{t+\Delta(t)} \to -\infty$ a.s. on $\mathcal{X}^c$. Hence, as in concluding the proof of Theorem 4.1,

$$\lim_{t\to\infty} \mathbb{P}_{x_0}[\{t^{-1/2}(\log X_{t+\Delta(t)} - \nu t) \geq y\} \cap \mathcal{X}^c] = 0. \qquad (6.12)$$

It thus remains to approximate the probability $\mathbb{P}_{x_0}[\{t^{-1/2}(\log X_{t+\Delta(t)} - \nu t) \geq y\} \cap \mathcal{X}]$. For this, we use the following sandwich:

$$\begin{aligned}
&\left(\{t^{-1/2}(\log X_t - \nu t) \geq y + t^{-1/2}r(t)\} \cap \mathcal{X}\right) \setminus (\mathcal{E}_{s,t}^c \cap \mathcal{X}) \\
&\subset \left(\{t^{-1/2}(\log X_{t+\Delta(t)} - \nu t) \geq y\} \cap \mathcal{X}\right) \qquad\qquad (6.13) \\
&\subset \left(\{t^{-1/2}(\log X_t - \nu t) \geq y - t^{-1/2}r(t)\} \cap \mathcal{X}\right) \cup (\mathcal{E}_{s,t}^c \cap \mathcal{X}).
\end{aligned}$$

The first step now is to show that the event

$$\mathcal{E}_{s,t}^c \cap \mathcal{X} \subset (A_s^{(1)})^c \cup (A_t^{(2)})^c \cup (A_t^{(3)})^c \cup (\mathcal{X} \setminus A_s^{(4)})$$

has small probability when s and t are large. It is immediate that

$$\lim_{s\to\infty} \mathbb{P}_{x_0}[(A_s^{(1)})^c] = \lim_{t\to\infty} \mathbb{P}_{x_0}[(A_t^{(2)})^c] = 0,$$

because $\Delta(t) \to \Delta$ a.s. and because Δ is a.s. finite. If $\beta^* := \sup_{k\geq 0} \beta_k$, a comparison between Z and a Poisson process with rate β^* shows that

$$\mathbb{P}_{x_0}[(A_t^{(3)})^c] \leq \text{Po}\left(\beta^*(t^{1/8} + 1)\right)\{(t^{1/4}, \infty)\} \to 0 \quad \text{as } t \to \infty.$$

For $\mathcal{X} \setminus A_s^{(4)}$, note that $A_s^{(4)}$ is an increasing sequence, and that $\mathcal{X} \subset \lim_{s\to\infty} A_s^{(4)}$, because $\mathcal{X} = \{\lim_{t\to\infty} \log W_t \in \mathbb{R}\}$. On the other hand, for all $s \geq 1$,

$$\mathcal{X}^c = \{\lim_{t\to\infty} \log W_t = -\infty\} \subset (A_s^{(4)})^c.$$

Hence $\mathcal{X} = \lim_{s\to\infty} A_s^{(4)}$ a.s., and so $\lim_{s\to\infty} \mathbb{P}_{x_0}[\mathcal{X} \setminus A_s^{(4)}] = 0$. Thus it follows that

$$\lim_{s\to\infty} \lim_{t\to\infty} \mathbb{P}_{x_0}[\mathcal{E}_{s,t}^c \cap \mathcal{X}] = 0. \qquad (6.14)$$

Returning to (6.13), we note that

$$\lim_{t\to\infty} \mathbb{P}_{x_0}[\{t^{-1/2}(\log X_t - \nu t) \geq y + t^{-1/2}r(t)\} \cap \mathcal{X}]$$
$$= \lim_{t\to\infty} \mathbb{P}_{x_0}[\{t^{-1/2}(\log X_t - \nu t) \geq y - t^{-1/2}r(t)\} \cap \mathcal{X}] = \mathbb{P}_{x_0}[\mathcal{X}](1 - \Phi(y/v)),$$

by (4.29) and because $t^{-1/2}r(t) \to 0$ as $t \to \infty$. It thus follows that

$$\limsup_{t\to\infty} |\mathbb{P}_{x_0}[\{t^{-1/2}(\log X_{t+\Delta(t)} - \nu t) \geq y\} \cap \mathcal{X}] - \mathbb{P}_{x_0}[\mathcal{X}](1 - \Phi(y/v))| \leq \limsup_{t\to\infty} \mathbb{P}_{x_0}[\mathcal{E}_{s,t}^c \cap \mathcal{X}]$$

for any $s \geq 1$, and hence, from (6.12) and (6.14), that

$$\lim_{t\to\infty} \mathbb{P}_{x_0}[t^{-1/2}(\log Y_t^{(h)} - \nu t) \geq y] = \mathbb{P}_{x_0}[\mathcal{X}](1 - \Phi(y/v))$$

for all $y \in \mathbb{R}$. The conclusion of the theorem now follows by replacing t by $h_0(m) - h_0(m_0)$ for $m \geq m_0$, $m \in \mathbb{N}$, and noting that $h_0(m) - h_0(m_0) \sim \log(m/m_0)$ as $m \to \infty$. ∎

6.2. *The discrete degree weighted process.* There is also a degree weighted process $\widetilde{Y}$ in discrete time. Multiplying equation (1.4) for $\mathfrak{p}_{m+1,k}$ by k, it follows that, with $u_{m,k} := k\mathfrak{p}_{m,k}$,

$$u_{m+1,k} = u_{m,k} + \frac{1}{m+1}\left\{-u_{m,k}(1+\alpha k) + \alpha k u_{m,k-1} + q_k u_{m,k} + \sum_{j\geq k} p_j(1-q_j)u_{m,j}\tilde{\pi}_{jk}\right\},$$

where $\tilde{\pi}_{jk} := k\pi_{jk}/\{jp_j\}$, so that $\widetilde{\Pi}_j$ is the size–biased transformation of Π_j. This can be re-written in vector form as

$$u_{m+1}^\top = u_m^\top\left\{I + \frac{1}{m+1}\{[\widetilde{Q}]_{m+1} + (2\alpha - 1)I\}\right\},$$

where $\widetilde{Q}$ is the Q-matrix defined in (5.1) with $\beta_j := 1 - q_j$; recall that, for each $j \geq 1$, $0 < q_j + p_j(1-q_j) = \alpha < 1$ is constant. Hence, defining $v_m := \prod_{s=m_0}^{m-1}\{1 + (2\alpha-1)/(s+1)\}^{-1}u_m$, this implies that, in parallel to (1.5),

$$v_{m+1}^\top = v_m^\top\left\{I + \frac{1}{m+2\alpha}[\widetilde{Q}]_{m+1}\right\}, \quad \text{and so} \quad v_m^\top = v_{m_0}^\top \prod_{j=m_0+1}^{m}\left\{I + \frac{1}{j+2\alpha-1}[\widetilde{Q}]_j\right\}.$$

These equations show that, if the initial configuration at time m_0 has j_0 vertices, then the family of vectors $j_0^{-1}(v_m, m \geq m_0)$ can be interpreted as the sequence of probability distributions $(\tilde{\mathfrak{p}}_m^{(j_0,m_0)}, m \geq m_0)$ of a time inhomogeneous Markov chain $\widetilde{Y}$ starting in state j_0 at time m_0, since $v_{m_0,j_0} = u_{m_0,j_0} = j_0$ if $\mathfrak{p}_{m_0}$ is the distribution with point mass at j_0, and since $I + (j+2\alpha-1)^{-1}[\widetilde{Q}]_j$ is a stochastic matrix, for each j. This in turn implies that, if $\mathfrak{p}_m^{(j_0,m_0)}$ denotes the distribution of Y, starting in state j_0 at time m_0, then

$$\mathfrak{p}_{m,k}^{(j_0,m_0)} = j_0 k^{-1} \prod_{s=m_0}^{m-1}\{1 + (2\alpha-1)/(s+1)\}\tilde{\mathfrak{p}}_{m,k}^{(j_0,m_0)},$$

so that the fractions $\mathfrak{p}_m^{(j_0,m_0)}$ can be deduced from the probabilities $\tilde{\mathfrak{p}}_t^{(j_0,m_0)}$. This is the discrete analogue of Proposition 5.1; note that, for large m_0 and $m \geq m_0$,

$$\prod_{s=m_0}^{m-1}\{1 + (2\alpha-1)/(s+1)\} \approx e^{-(1-2\alpha)(m-m_0)},$$

making the comparison clearer.

A time changed version $\widetilde{Y}^{(h)}$ of the discrete time inhomogeneous degree weighted process $\widetilde{Y}$ can be closely coupled to the continuous time degree weighted process, using arguments exactly like those above. The processes $\widetilde{Y}$ and $\widetilde{X}$ have the same jump chains, and can be constructed in parallel, using

the same realizations of the jump chain, and the same sequence of independent uniform random variables. The time scale for the discrete process that matches that of the continuous time process is $(h_a(m) - h_a(m_0), m \geq m_0)$, where $a := 2\alpha - 1$, and the waiting time until the next jump in the time scaled discrete process $\widetilde{Y}^{(h)}$, when starting in state k at time $h_a(m) - h_a(m_0)$ (or at step m of the process $\widetilde{Y}$), has the distribution of $V_a(m, -\widetilde{Q}_{k,k})$. The subsequent arguments, showing that, with this coupling, $\widetilde{Y}_t^{(h)} = \widetilde{X}_{t+\widetilde{\Delta}(t)}$ for all $t \geq 0$, for a difference $\widetilde{\Delta}(t)$ defined analogously to the definition (6.3) of $\Delta(t)$, and that $\widetilde{\Delta}(t)$ converges a.s. as $t \to \infty$, are substantially the same.

The next lemma can be applied to show that the limiting probabilities for the degree weighted inhomogeneous chain $\widetilde{Y}$ are the same as those for $\widetilde{X}$, if the process $\widetilde{X}$ is positive recurrent. The lemma can also be applied to the discrete versions of the unweighted chain, in models, such as the DD model with 'random re-wiring' in Section 7.2, in which the original chain is positive recurrent. For the basic DD model, with 0 an absorbing state, it is not needed.

Lemma 6.6. *Suppose that $(X_t, t \geq 0)$ is an irreducible positive recurrent pure jump Markov process on $\mathbb{Z}$, with some initial distribution λ and with stationary distribution π. Let $(\Delta(t), t \geq 0)$ be a càdlàg stochastic process such that $\Delta := \lim_{t \to \infty} \Delta(t)$ exists and is finite a.s. Then, for each $j \in \mathbb{Z}$,*

$$\lim_{t \to \infty} \mathbb{P}[X_{t+\Delta(t)} = j] \ = \ \pi_j.$$

Proof: Suppose first that Δ is integrable. Write $\widehat{\Delta}(t) := \mathbb{E}\{\Delta \mid \mathcal{F}_t\}$, where $\mathcal{F}_t := \sigma\{X_s, 0 \leq s \leq t\}$. Define the events

$$A_s(\varepsilon) \ := \ \{|\Delta(t) - \Delta| \leq \varepsilon \text{ for all } t \geq s\}; \quad \widehat{A}_s(\varepsilon) \ := \ \{|\widehat{\Delta}(t) - \Delta| \leq \varepsilon \text{ for all } t \geq s\},$$

noting that, for each $\varepsilon > 0$,

$$\lim_{s \to \infty} \mathbb{P}[A_s(\varepsilon)] \ = \ \lim_{s \to \infty} \mathbb{P}[\widehat{A}_s(\varepsilon)] \ = \ 1.$$

Now, for any $j \in \mathbb{Z}$, $s, t > 0$ and $\varepsilon > 0$,

$$\{X_{t+\Delta(t)} = j\} \ \supset \ \{X_{t+\widehat{\Delta}(t)+u} = j, \ -\varepsilon \leq u \leq \varepsilon\} \cap A_s(\varepsilon/2) \cap \widehat{A}_s(\varepsilon/2)$$
$$\supset \ \{X_{t+\widehat{\Delta}(s)+u} = j, \ -\varepsilon \leq u \leq \varepsilon\} \cap A_s(\varepsilon/2) \cap \widehat{A}_s(\varepsilon/4),$$

from which it follows that

$$\mathbb{P}[X_{t+\Delta(t)} = j] \ \geq \ \mathbb{P}[X_{t+\widehat{\Delta}(s)+u} = j, \ -\varepsilon \leq u \leq \varepsilon] - \mathbb{P}[A_s^c(\varepsilon/2)] - \mathbb{P}[\widehat{A}_s^c(\varepsilon/4)]. \qquad (6.15)$$

Then, by the Markov property,

$$\mathbb{P}[X_{t+\widehat{\Delta}(s)+u} = j, \ -\varepsilon \leq u \leq \varepsilon \mid \mathcal{F}_s] \, I[t + \widehat{\Delta}(s) - \varepsilon \geq s]$$
$$= \ e^{-2q_j\varepsilon} \, \mathbb{P}[X_{t+\widehat{\Delta}(s)-\varepsilon} = j \mid \mathcal{F}_s] \, I[t + \widehat{\Delta}(s) - \varepsilon \geq s]$$
$$= \ e^{-2q_j\varepsilon} \, P_{X_s,j}(t + \widehat{\Delta}(s) - \varepsilon - s) \, I[t + \widehat{\Delta}(s) - \varepsilon \geq s],$$

where q_j denotes the jump rate from the state j, and $P_{i,j}(u) := \mathbb{P}[X_u = j \mid X_0 = i]$. Hence, and because $\lim_{t \to \infty} I[t + \widehat{\Delta}(s) - \varepsilon \geq s] = 1$ a.s., we have

$$\lim_{t \to \infty} \mathbb{P}[X_{t+\widehat{\Delta}(s)+u} = j, \ -\varepsilon \leq u \leq \varepsilon \mid \mathcal{F}_s]$$
$$= \ \lim_{t \to \infty} \{\mathbb{P}[X_{t+\widehat{\Delta}(s)+u} = j, \ -\varepsilon \leq u \leq \varepsilon \mid \mathcal{F}_s] \, I[t + \widehat{\Delta}(s) - \varepsilon \geq s]\}$$
$$= \ e^{-2q_j\varepsilon} \, \pi_j \quad \text{a.s.},$$

and thus, by dominated convergence,

$$\lim_{t \to \infty} \mathbb{P}[X_{t+\widehat{\Delta}(s)+u} = j, \ -\varepsilon \leq u \leq \varepsilon] \ = \ e^{-2q_j\varepsilon} \, \pi_j.$$

It now follows from (6.15) that

$$\liminf_{t\to\infty} \mathbb{P}[X_{t+\Delta(t)} = j] \ \geq \ e^{-2q_j\varepsilon}\,\pi_j - \mathbb{P}[A_s^c(\varepsilon/2)] - \mathbb{P}[\widehat{A}_s^c(\varepsilon/4)],$$

for any $s > 0$, $j \in \mathbb{Z}$ and $\varepsilon > 0$. Letting $s \to \infty$ and then $\varepsilon \to 0$ thus implies that

$$\liminf_{t\to\infty} \mathbb{P}[X_{t+\Delta(t)} = j] \ \geq \ \pi_j \tag{6.16}$$

for all $j \in \mathbb{Z}$. The inequality (6.16), used in the second and final inequalities below, now implies that, for any $j_0 \in \mathbb{Z}$ and any *finite* set $J \subset \mathbb{Z}$ such that $j_0 \in J$,

$$1 - \sum_{j\in J}\pi_j \ \geq \ \limsup_{t\to\infty} \mathbb{P}[X_{t+\Delta(t)} \in J] - \sum_{j\in J}\pi_j \tag{6.17}$$

$$\geq \ \limsup_{t\to\infty}\left(\mathbb{P}[X_{t+\Delta(t)} = j_0] - \pi_{j_0}\right) \ \geq \ \liminf_{t\to\infty}\left(\mathbb{P}[X_{t+\Delta(t)} = j_0] - \pi_{j_0}\right) \ \geq \ 0.$$

Letting the set J increase towards the whole of $\mathbb{Z}$, the sum $\sum_{j\in J}\pi_j$ can be made arbitrarily close to 1, and so the left hand side of (6.17) can be made arbitrarily small, implying that

$$\lim_{t\to\infty} \mathbb{P}[X_{t+\Delta(t)} = j_0] \ = \ \pi_{j_0}$$

for any $j_0 \in \mathbb{Z}$, as required.

If Δ is not integrable, for any $M > 0$, define $\Delta_M := [\Delta]_{-M}^M := \max\{-M, \min(\Delta, M)\}$ to be Δ truncated within the range $[-M, M]$, and let $\Delta_M(t) := [\Delta(t)]_{-M}^M$. Then, because $[\cdot]_{-M}^M$ is a continuous function, $\lim_{t\to\infty}\Delta_M(t) = \Delta_M$ a.s., and hence, by the previous result,

$$\lim_{t\to\infty} \mathbb{P}[X_{t+\Delta_M(t)} = j] \ = \ \pi_j, \qquad j \in \mathbb{Z},$$

since Δ_M is integrable. But

$$|\mathbb{P}[X_{t+\Delta_M(t)} = j] - \mathbb{P}[X_{t+\Delta(t)} = j]| \ \leq \ \mathbb{P}[\Delta_M(t) \neq \Delta(t)] \ = \ \mathbb{P}[|\Delta(t)| > M],$$

and, since $\Delta(t) \to \Delta$ a.s., $\limsup_{t\to\infty}\mathbb{P}[|\Delta(t)| \geq M] \ \leq \ \mathbb{P}[|\Delta| \geq M]$. Hence, for all $M > 0$,

$$\pi_j - \mathbb{P}[|\Delta| \geq M] \ \leq \ \liminf_{t\to\infty}\mathbb{P}[X_{t+\Delta(t)} = j] \ \leq \ \limsup_{t\to\infty}\mathbb{P}[X_{t+\Delta(t)} = j] \ \leq \ \pi_j + \mathbb{P}[|\Delta| \geq M],$$

implying that $\lim_{t\to\infty}\mathbb{P}[X_{t+\Delta(t)} = j] = \pi_j$, and completing the proof. ∎

Applying Lemma 6.6 to the process $\widetilde{Y}^{(h)}$ shows that, if $\widetilde{X}$ is positive recurrent, with stationary distribution $\tilde{p}(\infty)$, then, in view of (6.4),

$$\lim_{t\to\infty}\mathbb{P}[\widetilde{Y}_t^{(h)} = j] \ = \ \tilde{p}_j(\infty), \qquad j \geq 1.$$

It thus also follows that

$$\lim_{m\to\infty}\mathbb{P}[\widetilde{Y}_m = j] \ = \ \tilde{p}_j(\infty), \qquad j \geq 1,$$

and hence that the inhomogeneous process $\widetilde{Y}$ has the same limiting distribution as $\widetilde{X}$, since $\widetilde{Y}_m = \widetilde{Y}_{h_0(m)-h_0(m_0)}^{(h)}$, and $\lim_{m\to\infty}\{h_0(m) - h_0(m_0)\} = \infty$.

7. Variants of the basic process

7.1. *The process with removal of edges.* A process, in which edges are also allowed to be removed, has been discussed in Hermann and Pfaffelhuber (2021). The analogue in our formulation is obtained by modifying the process X of Section 2 by allowing deaths at rate $k\delta_k$, $k \geq 1$, with $|\delta_k - \delta| \leq c_5 k^{-\gamma}$ for all k. This gives a Q-matrix

$$Q_{k,k+1}^{(d)} \ = \ k\alpha_k; \qquad Q_{k,k-1}^{(d)} \ = \ k\delta_k + \beta_k\pi_{k,k-1}; \qquad Q_{k,k}^{(d)} \ = \ -\{k(\alpha_k + \delta_k) + \beta_k(1 - \pi_{kk})\};$$

$$Q_{k,j}^{(d)} \ = \ \beta_k\pi_{kj}, \quad 0 \leq j \leq k - 2; \qquad Q_{k,j}^{(d)} \ = \ 0, \quad j \geq k + 2, \tag{7.1}$$

for the modified process $X^{(d)}$, whose properties are much as for the process without deaths. First, the asymptotic growth rate for the process without catastrophes becomes $\alpha^{(d)} := \alpha - \delta$, rather that α. This affects the criterion for transience and recurrence in the obvious way; Theorem 2.3 holds, but with α replaced by $\alpha^{(d)}$, and the proofs are otherwise unchanged. The net growth rate for the process with catastrophes now becomes $\nu^{(d)} := \alpha^{(d)} - \beta \log(1/p)$, and Theorem 3.1 holds with $\alpha^{(d)}$ for α and with $\nu^{(d)}$ for ν. Here, some small modifications to the proofs of two of the lemmas are needed, largely because the process is no longer increasing between catastrophes.

In the discussion that follows, we drop the superscript $^{(d)}$, where possible. First, in the proof of Lemma 3.2, the martingale $(M_l, l \geq 0)$ in (3.8) has to be modified. Here, $\alpha_{Y_u^{(i)}}$ is to be replaced by $\tilde{\alpha}_{Y_u^{(i)}}$, where, in order to obtain a martingale, we define $\tilde{\alpha}_j := \alpha_j - j\delta_j/(j-1)$ in $j \geq 2$; the quantity $\tilde{\alpha}_j$ is not quite the same as $\alpha_j - \delta_j$, but is close to it when j is large, which is the case in the arguments of Lemma 3.2. A stopped version M^σ is used in the proofs, where $M_l^\sigma := M_{l \wedge \sigma}$, and

$$\sigma := \sigma^{(i)} := \inf\{l \geq 1 : Y_{T_{[i-1]}+l}^{(i)} \leq \tfrac{1}{2}\{X_{[i-1]} + l\rho(\tfrac{1}{2}X_{[i-1]})\}\},$$

where $\rho(j) := \inf_{k \geq j}\{(\alpha_k - \delta_k)/(\alpha_k + \delta_k)\}$, a lower bound for the drift of the asymmetric random walk described by the jump chain of $Y^{(i)}$ in $\mathbb{Z} \cap [j, \infty)$. Provided that $X_{[i-1]} \geq 2$, there is thus no need to define $\tilde{\alpha}_j$ for $j \leq 1$. It is then possible to apply Doob's inequality to control the fluctuations of M^σ, and this can be translated into bounds on the probability of the events A_i, again with α replaced by $\alpha^{(d)}$. Note that the variance of M^σ and the means and variances of $\Delta_l^\sigma := \Delta_{l \wedge \sigma}$, conditional on $\mathcal{F}_{T_{i-1}}$, can be bounded much as in (3.9) and (3.11), but with different constants. For instance, temporarily defining

$$\hat{p}_j := 1 - \hat{q}_j := \frac{\alpha_j}{\alpha_j + \delta_j},$$

we have

$$\mathbb{E}\{(M_l - M_{l-1})^2 \,|\, \mathcal{F}_{S_{l-1}} \cap \{\sigma^{(i)} > l-1\} \cap \{Y_{S_{l-1}}^{(i)} = j\}\}$$
$$= \hat{p}_j\hat{q}_j\Big\{\frac{1}{j} + \frac{1}{j-1}\Big\}^2 + \Big\{\frac{\hat{p}_j}{j} - \frac{\hat{q}_j}{j-1}\Big\}^2 \leq 4j^{-2}, \tag{7.2}$$

in $j \geq 2$, where the first term in (7.2) comes from the variability in $h(Y_{S_l}^{(i)})$ and the second from the variability in $\tilde{\alpha}(S_l - S_{l-1})$. Hence, for any $l \geq 1$

$$\mathrm{Var}\,\{M_l^\sigma \,|\, \mathcal{F}_{T_{i-1}}\} \leq 4 \sum_{k \geq 1} \mathbb{E}\{(Y_{S_{k-1}}^{(i)})^{-2} I[\sigma^{(i)} > k-1] \,|\, \mathcal{F}_{T_{i-1}}\}. \tag{7.3}$$

But since, for any $r > 1$,

$$\sum_{l \geq 0} (Y_{S_l}^{(i)})^{-r} I[\sigma^{(i)} > l] \leq \sum_{l \geq 0} \big\{\tfrac{1}{2}\{X_{[i-1]} + l\rho(\tfrac{1}{2}X_{[i-1]})\}\big\}^{-r}$$
$$\leq \frac{2^r}{X_{[i-1]}^{r-1}}\Big\{\frac{1}{X_{[i-1]}} + \frac{1}{(r-1)\rho(\tfrac{1}{2}X_{[i-1]})}\Big\},$$

the sum in (7.3) is bounded by $16X_{[i-1]}^{-1}(1 + 1/\rho(j_0))$, uniformly in $X_{[i-1]} \geq 2j_0$, for any j_0 such that $\rho(j_0) > 0$. This is the analogue of (3.9). The bounds analogous to (3.9) and (3.11) in turn imply analogues of the probability bounds (3.10) and (3.12) for the deviations of M^σ and Δ^σ, again with different constants. By using a Wald martingale for the simple random walk with drift $\rho(\tfrac{1}{2}X_{[i-1]})$, the probability $\mathbb{P}[\sigma^{(i)} < \infty \,|\, \mathcal{F}_{T_{i-1}}]$ is bounded by $e^{-cX_{[i-1]}}$, for a suitable constant $c = c(\rho(\tfrac{1}{2}X_{[i-1]}))$. This is enough to establish Lemma 3.2, under the extended model.

In the proof of Lemma 3.3, the fact that the process $Y^{(i)}$ is non-decreasing is used, but only for convenience. The proof can easily be modified for the process in which deaths are allowed to occur. It is enough, instead, to know that $Y^{(i)}$ does not fall below the value $\tfrac{1}{2}X_{[i-1]}$, and to define η so

that $(2/\eta)^{4/\gamma} > 2j'$ in the analogue of Lemma 3.3. Then, in the induction argument used to prove Lemma 3.3, note that on A_i, for any $t \in [T_{i-1}, T_i)$,

$$\log X_t \geq \log X_{[i-1]} - X_{[i-1]}^{-\gamma/4} \geq \log(X_{[i-1]}/2),$$

as desired, if $\eta < 2\log 2$, by the induction hypothesis, so that then $\beta_{X_t} \leq \beta'$ for all $T_{i-1} \leq t < T_i$. The remainder of the proof is as before. In the analogue of Lemma 3.5, η should also now satisfy $(2/\eta)^{4/\gamma} > 2j'$.

Theorem 4.1 also holds, with $\nu^{(d)}$ for ν. As for Theorem 3.1, some small modifications in the proofs are needed, again because the process is no longer increasing between catastrophes.

The biological motivation for allowing edges to be removed is that some interactions may decline in importance over time, being replaced by more advantageous interactions. However, these removals would not occur only at the occasions on which a vertex is duplicated, so that the original discrete process is not well suited to such a modification. Hermann and Pfaffelhuber (2021) introduce edge deletions in a process very similar to the continuous time model of Jordan (2018), in which edges can easily be modelled as having independent exponential lifetimes.

In Hermann and Pfaffelhuber (2020), extra terms are introduced into the birth–death–catastrophe process, allowing for upward jumps of sizes greater than one, those corresponding to a Markov branching process. Here, we can again allow such jumps, now allowing the jump rates to vary a little with the state k. Writing $\mathbb{N}_{-1} := \{-1\} \cup \mathbb{N}$, we can modify the matrix Q further by setting

$$Q_{k,k+j}^{(b)} = ka_{k,j}, \quad j \in \mathbb{N}; \qquad Q_{k,k-1}^{(b)} = ka_{k,-1} + \beta_k \pi_{k,k-1};$$

$$Q_{k,k}^{(b)} = -\left\{ k \sum_{j \in \mathbb{N}_{-1}} a_{k,j} + \beta_k(1 - \pi_{kk}) \right\}; \qquad Q_{k,j}^{(b)} = \beta_k \pi_{kj}, \quad 0 \leq j \leq k-2. \tag{7.4}$$

Here, we assume that $\inf_{k \geq 1} \sum_{j \in \mathbb{N}_{-1}} a_{k,j} > 0$, and that

$$|a_{k,j} - a_{*,j}| \leq c_{*,j} k^{-\gamma}, \quad j \in \mathbb{N}_{-1}, \quad k \geq 1,$$

where

$$\sum_{j \in \mathbb{N}_{-1}} j a_{*,j} =: \alpha^{(b)} < \infty; \qquad \sum_{j \in \mathbb{N}_{-1}} j c_{*,j} < \infty,$$

and $\sum_{j \in \mathbb{N}_{-1}} j^2 a_{k,j}$ is uniformly bounded in k. It then follows that the asymptotic growth rate for the process without catastrophes is $\alpha^{(b)}$, rather that α, which is consistent with $\alpha^{(d)}$ when $a_{*,j} = 0$ for all $j \geq 2$, and with α in Section 2, when also $a_{*,-1} = 0$. Largely using the proof of Theorem 2.1, we can show that the process $X^{(*,b)}$, that has an additional transition from state 0 to state 1, satisfies the following theorem.

Theorem 7.1. *For the process $X^{(*,b)}$, we have the following behaviour:*

$$
\begin{aligned}
(1): &\quad \text{if } \alpha^{(b)} < \beta \log(1/p), \text{ then } X^{(*,b)} \text{ is geometrically ergodic;} \\
(2): &\quad \text{if } \alpha^{(b)} = \beta \log(1/p), \text{ then } X^{(*,b)} \text{ is null recurrent;} \\
(3): &\quad \text{if } \alpha^{(b)} > \beta \log(1/p), \text{ then } X^{(*,b)} \text{ is transient.}
\end{aligned}
\tag{7.5}
$$

Proof: The difference from the proof of Theorem 2.1 comes from needing to find the leading terms in $k \sum_{j \in \mathbb{N}_{-1}} a_{k,j}(f(k+j) - f(k))$, rather than just in $k(f(k+1) - f(k))$, for the same four choices of f. For the two recurrence arguments, when f is concave, we immediately have

$$\sum_{j \in \mathbb{N}} a_{k,j}(f(k+j) - f(k)) \leq \sum_{j \in \mathbb{N}} j a_{k,j} f'(k),$$

and the rest of the proof is straightforward, under the assumptions on the $a_{k,j}$. For the remaining two arguments, we use (2.13) with $x = k$ and $y = k + j$ in proving null recurrence, and (2.16) in proving transience.   ∎

The analogue of Theorem 2.5 is also true, in case (1), with $\alpha^{(b)}$ in place of α.

In case (3), the analogue of Theorem 3.1, with $\alpha^{(b)}$ in place of α, is also true. As discussed above for $X^{(d)}$, there are modifications to the proofs of Lemmas 3.2, 3.3 and 3.5, because of the deaths. In Lemma 3.2, the martingale M in (3.8) again has to be modified, now with $\alpha_{Y_u^{(i)}}$ replaced by $\tilde{\alpha}_{Y_u^{(i)}}^{(b)}$, where we define

$$\tilde{\alpha}_j^{(b)} := j \sum_{k \in \mathbb{N}_{-1}} a_{j,k}(h(j+k-1) - h(j-1)) \approx \alpha^{(b)}$$

in $j \geq 2$. The lower bound $\rho(j)$ on the drift of the jump chain, used in defining the stopping time $\sigma^{(i)}$, also needs to be modified to

$$\rho(j) := \inf_{k \geq j}\left\{ \frac{\sum_{l \in \mathbb{N}_{-1}} l a_{k,l}}{\sum_{l \in \mathbb{N}_{-1}} a_{k,l}} \right\}.$$

As for $X^{(d)}$, it is necessary to be able to bound the variance of M^σ and the means and variances of $\Delta_l^\sigma := \Delta_{l \wedge \sigma}$, conditional on $\mathcal{F}_{T_{l-1}}$, much as in (3.9) and (3.11), but with different constants. Under our assumptions on the values of $a_{k,j}$, the calculations are routine. Finally, again with small modifications to the proof, Theorem 4.1 also holds for $X^{(b)}$, with $\nu^{(b)}$ for ν.

7.2. *The process with random re-wiring.* In Pastor-Satorras et al. (2003), the basic DD model is modified to accommodate random connections, added to the duplicate vertex after copying and thinning. At time $(m+1)$, each vertex, other than the vertex being copied and its neighbours, may be connected to the duplicate, independently with probability r/m.

In this case, the equations (1.4) for probabilities in the basic DD model, for $k \geq 1$, are modified to

$$\begin{aligned}
\mathfrak{p}_{m+1,k}^{(r)} = \mathfrak{p}_{m,k}^{(r)} + \frac{1}{m+1}\Big\{ & -\{1 + \alpha k + r(1 - (k+1)/m)\}\mathfrak{p}_{m,k}^{(r)} \\
& + \{(k-1)\alpha + r(1 - k/m)\}\mathfrak{p}_{m,k-1}^{(r)} \\
& + q\sum_{j \leq k} \mathfrak{p}_{m,j}^{(r)}\hat{\pi}_{jk}^{(r,m)} + \sum_{j \geq 0}(1-q)\mathfrak{p}_{m,j}^{(r)}\pi_{jk}^{(r,m)} \Big\},
\end{aligned} \tag{7.6}$$

where, as before, $\alpha := q + p(1-q)$; the probability distribution $\widehat{\Pi}_k^{(r,m)}$ is the binomial distribution $\mathrm{Bi}\,(m-k, r/m)$, having point probabilities $\hat{\pi}_{kj}^{(r,m)}$, and the probability distribution $\Pi_k^{(r,m)}$, with point probabilities $\pi_{kj}^{(r,m)}$, is the convolution of Π_k and $\widehat{\Pi}_k^{(r,m)}$, and has mean $kp + r - rk/m$. In vector notation, the equations can be written as

$$(\mathfrak{p}_{m+1}^{(r)})^\top = (\mathfrak{p}_m^{(r)})^\top\{I + (m+1)^{-1}[Q^{(r,m+1)}]_{m+1}\}, \tag{7.7}$$

and so

$$(\mathfrak{p}_m^{(r)})^\top = (\mathfrak{p}_{m_0}^{(r)})^\top \prod_{j=m_0+1}^{m} \{I + j^{-1}[Q^{(r,j)}]_j\}, \tag{7.8}$$

where

$$\begin{aligned}
Q_{k,k+1}^{(r,m)} &= \alpha k + r(1 - (k+1)/m) + (1-q)\pi_{k.k+1}^{(r,m)} + q\hat{\pi}_{k.k+1}^{(r,m)}; \\
Q_{k,k}^{(r,m)} &= -\{\alpha k + r(1 - (k+1)/m) + (1-q)(1 - \pi_{kk}^{(r,m)}) + q(1 - \hat{\pi}_{kk})^{(r,m)}\}; \\
Q_{k,j}^{(r,m)} &= (1-q)\pi_{kj}^{(r,m)}, \quad 0 \leq j \leq k-1; \quad Q_{k,j}^{(r,m)} = (1-q)\pi_{kj}^{(r,m)} + q\hat{\pi}_{kj}^{(r,m)}, \quad j \geq k+2,
\end{aligned} \tag{7.9}$$

for each $k \geq 1$, and with

$$Q_{0,1}^{(r,m)} = r(1 - 1/m) + \hat{\pi}_{01}^{(r,m)}; \qquad Q_{0,j}^{(r,m)} = \hat{\pi}_{0j}^{(r,m)}, \quad j \geq 2;$$
$$Q_{0,0}^{(r,m)} = -\{r(1 - 1/m) + 1 - \hat{\pi}_{00}^{(r,m)}\}. \tag{7.10}$$

Note that, if $r = 0$, $Q^{(r,m)}$ reduces to the Q-matrix for the basic DD model. If m is large, the Q-matrices $Q^{(r,m)}$ are very close to the matrix $Q^{(r)}$, in which, for all $k \geq 0$, the fractions $(k+1)/m$ are replaced by zero, $\widehat{\Pi}_k^{(r,m)}$ is replaced by $\widehat{\Pi}^{(r)}$, the Poisson distribution with mean r, and $\Pi_k^{(r,m)}$ is replaced by $\Pi_k^{(r)}$, the convolution of Π_k and $\widehat{\Pi}^{(r)}$.

Now the process $X^{(r)}$ with Q-matrix $Q^{(r)}$ is almost of the same form as that for the irreducible process X^* in Section 2, with $\alpha_0 := r$, and with $\alpha_k := \alpha + r/k$ and $\beta_k := (1 - q)$ for all $k \geq 1$. The difference is that it also has extra elements in the rates, because of the (limited) re-wiring.

First, instead of having rate αk of upward jumps in state k, with α constant, we have a rate $ka_{k,1}^{(r)}$, where

$$a_{k,1}^{(r)} = \alpha + \{r + (1 - q)\pi_{k.k+1}^{(r)} + q\hat{\pi}_{k,k+1}^{(r)}\}/k = \alpha + O(k^{-1}), \quad k \geq 1, \tag{7.11}$$

and there are also jump rates $ka_{k,j}^{(r)}$ for $j \geq 2$ as in (7.4), with

$$a_{k,j}^{(r)} = \{(1 - q)\pi_{k.k+j}^{(r)} + q\hat{\pi}_{k,k+j}^{(r)}\}/k = O(k^{-1}), \quad k \geq 1.$$

Thus the framework is technically that of the process $X^{(b)}$, as given by (7.4), though the quantities $a_{*,j}$ are all zero, except for $j = 1$. Then, instead of the copy distribution from a vertex of degree k being a mixture of the point mass on k with probability q and the distribution Π_k with probability $1 - q$, the distribution is convolved with the Poisson distribution $\mathrm{Po}(r)$. This in part contributes to the extra terms $a_{k,j}^{(r)}$, but also has the consequence that the distribution $\Pi_k^{(r)}$, when restricted to the set $\{0, 1, \ldots, k\}$, has probability mass less than 1. In order to fit with a process of the form $X^{(b)}$, the factor $1 - q$ should be replaced by a parameter $\beta_k^{(r)}$, where $\beta_k^{(r)} := (1 - q)\Pi_k^{(r)}\{[0, k]\}$, and the downward jump distribution should be modified to $\widetilde{\Pi}_k^{(r)} := (1 - q)\Pi_k^{(r)}/\beta_k^{(r)}$. Under the assumptions on Π_k, $\beta_k^{(r)} = (1 - q) + O(k^{-\gamma})$, the mean of $\widetilde{\Pi}_k^{(r)}$ is asymptotically $kp(1 + O(k^{-\gamma}))$, and its variance is of order $O(k^{2-\gamma})$.

Thus the process $X^{(r)}$ can be analyzed in the same way as for $X^{(b)}$. For instance, Theorem 7.1 shows that $X^{(r)}$ is positive recurrent if $\alpha < (1 - q)\log(1/p)$, null recurrent if $\alpha = (1 - q)\log(1/p)$ and transient if $\alpha > (1 - q)\log(1/p)$. In the case of positive recurrence, Theorem 2.5 indicates which moments of the stationary distribution exist. Note that the critical exponent η^* is smaller than $(1 - q)/\alpha = (1 - q)/(q + p(1 - q)) \leq 1/p$, so that higher moments can only be finite when p is small enough. In the case of transience, as for the process $X^{(b)}$, the conclusions of Theorems 3.1 and 4.1 hold.

We now show that the behaviour of the inhomogeneous Markov chain $\widehat{Y}^{(r)}$, whose probabilities are governed by the equations (7.8), can indeed be deduced from that of the process $X^{(r)}$. To do this, we couple $\widehat{Y}^{(r)}$ to a process $Y^{(r)}$ whose probabilities are given by (7.8), but with $Q^{(r,m)}$ replaced by $Q^{(r)}$. If the two processes are started together in the state (m_1, j_1), construct them jointly so as to make identical transitions as far as possible. If, at any step m, the two processes are in the same state k, the jump probabilities differ in two respects. First, the probability of jumping up by 1 because of a re-wiring from a copied vertex that is not a neighbour is an amount $r(k+1)/\{m(m+1)\}$ larger in $Y^{(r)}$ than in $\widehat{Y}^{(r)}$. Secondly, the copy distributions differ, in that one of them is Π_k convolved with $\mathrm{Bi}(m - (k+1), r/m)$ and the other with $\mathrm{Po}(r)$. As in Prokhorov (1953), the total variation distance between these two distributions is at most cm^{-1}, where c can be taken to be r. Hence the jump distributions of $Y^{(r)}$ and $\widehat{Y}^{(r)}$ when leaving state k at time $m + 1$

differ in total variation by at most $r(k+2)/\{m(m+1)\}$. This suggests that the two processes can be coupled so as to have identical paths with high probability, if m_1 is large enough.

More precisely, the probability that the two processes fail to remain identically coupled is at most

$$E_{j_1,m_1} = r \sum_{m > m_1} \frac{\mathbb{E}\{Y_m^{(r)} + 2 \,|\, Y_{m_1}^{(r)} = j_1\}}{m(m+1)}.$$

Now, from the one-step probabilities for $Y^{(r)}$, we have

$$\mathbb{E}\{Y_{m+1}^{(r)} \,|\, Y_m^{(r)} = j\} = j + \frac{\alpha j + 2r + (1-q)(p-1)j}{m+1} = j\left\{1 + \frac{2\alpha - 1}{m+1}\right\} + \frac{2r}{m+1},$$

and hence, for $m \geq m_1$,

$$\mathbb{E}\{Y_{m+1}^{(r)} \,|\, Y_{m_1}^{(r)} = j_1\} = \left\{1 + \frac{2\alpha - 1}{m+1}\right\}\mathbb{E}\{Y_m^{(r)} \,|\, Y_{m_1}^{(r)} = j_1\} + \frac{2r}{m+1}.$$

Iterating, we deduce that

$$\mathbb{E}\{Y_m^{(r)} \,|\, Y_{m_1}^{(r)} = j_1\} \leq \begin{cases} \left(j_1 + \frac{2r}{2\alpha - 1}\right)\prod_{l=m_1+1}^{m}\left\{1 + \frac{2\alpha-1}{l}\right\}, & \text{if } 2\alpha > 1; \\ j_1 + \sum_{l=m_1+1}^{m}\frac{2r}{l}, & \text{if } 2\alpha = 1; \\ j_1 \prod_{l=m_1+1}^{m}\left\{1 + \frac{2\alpha-1}{l}\right\} + \frac{2r}{1-2\alpha}, & \text{if } 2\alpha < 1, \end{cases}$$

for $m > m_1$, and hence that

$$\mathbb{E}\{Y_m^{(r)} \,|\, Y_{m_1}^{(r)} = j_1\} \leq \begin{cases} C(j_1 + 2r/(2\alpha - 1))(m/m_1)^{2\alpha-1}, & \text{if } 2\alpha > 1; \\ j_1 + 2r\{\log(m/m_1) + \gamma\}, & \text{if } 2\alpha = 1; \quad\quad (7.12) \\ Cj_1(m_1/m)^{1-2\alpha} + 2r/(1-2\alpha), & \text{if } 2\alpha < 1: \end{cases}$$

here, γ denotes Euler's constant. In all cases, this implies that $E_{j_1,m_1} \leq C'j_1 m_1^{-1}$, for a finite constant C', implying that the probability that the processes $Y^{(r)}$ and $\widehat{Y}^{(r)}$ ever differ, if started at time m_1 in state j_1, is of order $O(j_1 m_1^{-1})$.

Now, for the process $\widehat{Y}^{(r)}$ started at any time m_0 in any state j_0, couple it to a process $Y^{(r)}$ starting at time m_1 in the state $\widehat{Y}_{m_1}^{(r)}$. For $2\alpha > 1$, the probability that the two processes ever differ in $m \geq m_1$ is at most

$$\frac{C'}{m_1}\mathbb{E}\{\widehat{Y}_{m_1}^{(r)} \,|\, \widehat{Y}_{m_0}^{(r)} = j_0\} \leq \frac{C'}{m_1}C(j_0 + 2r/(2\alpha - 1))\left(\frac{m_1}{m_0}\right)^{2\alpha-1},$$

where, for the last bound, we have used (7.12), because the mean of $Y^{(r)}$ grows faster than that of $\widehat{Y}^{(r)}$. Hence the probability that the two processes ever differ in $m \geq m_1$ is at most

$$\frac{C''j_0}{m_0^{2\alpha-1}}\frac{1}{m_1^{2(1-\alpha)}},$$

and can be chosen to be as small as desired, for fixed initial state (j_0, m_0), by choosing m_1 large enough. For $2\alpha < 1$, the comparable bound is of order $O(j_0 m_1^{-1})$, and, for $2\alpha = 1$, it is of order $O(j_0 \log(m_1)m_1^{-1})$. Hence, after some finite random time, the paths of $\widehat{Y}^{(r)}$ are exactly those of a process $Y^{(r)}$. In consequence, the conclusions of Theorems 7.1 and 2.5 hold for $\widehat{Y}^{(r)}$, if $\alpha < (1-q)\log(1/p)$. If $\alpha > (1-q)\log(1/p)$, as in Theorem 6.4, $m^{-\alpha}p^{-J_m}Y_m^{(r)}$ converges a.s., and, as in Theorem 6.5,

$$\lim_{m \to \infty}\mathbb{P}[\{\log m\}^{-1/2}\{\log Y_m^{(r)} - \nu \log m\} \geq y \,|\, Y_{m_0} = x_0] = 1 - \Phi(y/v),$$

where $v^2 = (1-q)\{\log(1/p)\}^2$.

Note that the argument relating $\widehat{Y}^{(r)}$ to $Y^{(r)}$ does not greatly involve the detail of the re-wiring mechanism. For instance, at step $m+1$, suppose that each new vertex adds extra connections to a

random number of other vertices, chosen without replacement from all those other than the vertex being copied, and that double edges are then merged, as proposed in Bebek et al. (2006). Let the distribution of the number of extra connections be denoted by $\widehat{\Pi}^{(m)}$, having point probabilities $\hat{\pi}^{(m)}$ and mean r_m. The distribution $\Pi_k^{(m)}$, with point probabilities $\pi_k^{(m)}$, that represents the number of edges in the copy of a vertex of degree k at step $m+1$ is not quite the convolution of Π_k and $\widehat{\Pi}^{(m)}$, because some of the extra connections may duplicate edges to neighbours of the copied vertex, but it is easy to check that the difference in total variation between the two distributions is at most kr_m/m. The equations corresponding to (7.9), defining the analogue $Q^{(m)}$ of $Q^{(r,m)}$, have α replaced by $\alpha_m := \alpha + (1-\alpha)r_m/m$, $\hat{\pi}^{(r,m)}$ replaced by $\hat{\pi}^{(m)}$ and $\pi_k^{(r,m)}$ replaced by $\pi_k^{(m)}$. Then, if there is a distribution $\widehat{\Pi}^{(\infty)}$ with mean $r^{(\infty)}$ such that $d_{TV}(\widehat{\Pi}^{(m)}, \widehat{\Pi}^{(\infty)}) = O(m^{-1})$, and if $|r_m - r^{(\infty)}| + O(m^{-1})$, the previous arguments can be used to show that the process is asymptotically equivalent to one associated with the Q-matrix $Q^{(\infty)}$, obtained by formally replacing m by ∞ in the definitions of $Q^{(m)}$, to which the asymptotic theorems of the paper can once more be applied.

Acknowledgements

References

Bansaye, V., Pardo Millan, J. C., and Smadi, C. On the extinction of continuous state branching processes with catastrophes. *Electron. J. Probab.*, **18**, no. 106, 31 (2013). MR3151726.

Barbour, A. D., Heesterbeek, J. A. P., and Luchsinger, C. J. Thresholds and initial growth rates in a model of parasitic infection. *Ann. Appl. Probab.*, **6** (4), 1045–1074 (1996). MR1422977.

Bartoszyński, R., Bühler, W. J., Chan, W., and Pearl, D. K. Population processes under the influence of disasters occurring independently of population size. *J. Math. Biol.*, **27** (2), 167–178 (1989). MR991048.

Bebek, G., Berenbrink, P., Cooper, C., Friedetzky, T., Nadeau, J., and Sahinalp, S. C. The degree distribution of the generalized duplication model. *Theoret. Comput. Sci.*, **369** (1-3), 239–249 (2006). MR2277572.

Bhan, A., Galas, D. J., and Dewey, T. G. A duplication growth model of gene expression networks. *Bioinformatics*, **18**, 1486–1493 (2002). DOI: 10.1093/bioinformatics/18.11.1486.

Brockwell, P. J. The extinction time of a general birth and death process with catastrophes. *J. Appl. Probab.*, **23** (4), 851–858 (1986). MR867182.

Bühler, W. J. and Puri, P. S. The linear birth and death process under the influence of independently occurring disasters. *Probab. Theory Related Fields*, **83** (1-2), 59–66 (1989). MR1012494.

Chung, F., Lu, L., Dewey, T. G., and Galas, D. J. Duplication models for biological networks. *Journal of Computational Biology*, **10**, 677–687 (2003). DOI: 10.1089/106652703322539024.

Frieze, A., Turowski, K., and Szpankowski, W. Degree distribution for duplication-divergence graphs: large deviations. In *Graph-theoretic concepts in computer science*, volume 12301 of *Lecture Notes in Comput. Sci.*, pp. 226–237. Springer, Cham (2020). MR4184219.

Gibson, T. A. and Goldberg, D. S. Improving evolutionary models of protein interaction networks. *Bioinformatics*, **27**, 376–382 (2011). DOI: 10.1093/bioinformatics/btq623.

Hamza, K. and Klebaner, F. C. Conditions for integrability of Markov chains. *J. Appl. Probab.*, **32** (2), 541–547 (1995). MR1334906.

Hermann, F. and Pfaffelhuber, P. Large-scale behavior of the partial duplication random graph. *ALEA Lat. Am. J. Probab. Math. Stat.*, **13** (2), 687–710 (2016). MR3531396.

Hermann, F. and Pfaffelhuber, P. Markov branching processes with disasters: extinction, survival and duality to p-jump processes. *Stochastic Process. Appl.*, **130** (4), 2488–2518 (2020). MR4074709.

Hermann, F. and Pfaffelhuber, P. The partial duplication random graph with edge deletion. *ALEA Lat. Am. J. Probab. Math. Stat.*, **18** (1), 325–347 (2021). MR4213861.

Jacquet, P., Turowski, K., and Szpankowski, W. Power-law degree distribution in the connected component of a duplication graph. In *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 159 of *LIPIcs. Leibniz Int. Proc. Inform.*, pp. 16:1–16:14. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern (2020). DOI: 10.4230/LIPIcs.AofA.2020.16.

Jordan, J. The connected component of the partial duplication graph. *ALEA Lat. Am. J. Probab. Math. Stat.*, **15** (2), 1431–1445 (2018). MR3882189.

Kaplan, N., Sudbury, A., and Nilsen, T. S. A branching process with disasters. *J. Appl. Probability*, **12**, 47–59 (1975). MR362531.

Meyn, S. P. and Tweedie, R. L. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, **25** (3), 518–548 (1993). MR1234295.

Middendorf, M., Ziv, E., and Wiggins, C. H. Inferring network mechanisms: the Drosophila melanogaster protein interaction network. *Proceedings of the National Academy of Sciences*, **102**, 3192–3197 (2005). DOI: 10.1073/pnas.0409515102.

Pakes, A. G. The Markov branching-catastrophe process. *Stochastic Process. Appl.*, **23** (1), 1–33 (1986). MR866285.

Pastor-Satorras, R., Smith, E., and Solé, R. V. Evolving protein interaction networks through gene duplication. *J. Theoret. Biol.*, **222** (2), 199–210 (2003). MR2070214.

Pollett, P. K. Reversibility, invariance and μ-invariance. *Adv. in Appl. Probab.*, **20** (3), 600–621 (1988). MR955506.

Prokhorov, Y. V. Asymptotic behavior of the binomial distribution. *Uspehi Matem. Nauk (N.S.)*, **8** (3(55)), 135–142 (1953). MR0056861.

Tweedie, R. L. Sufficient conditions for regularity, recurrence and ergodicity of Markov processes. *Math. Proc. Cambridge Philos. Soc.*, **78**, 125–136 (1975). MR391269.

Tweedie, R. L. Criteria for classifying general Markov chains. *Advances in Appl. Probability*, **8** (4), 737–771 (1976). MR451409.

Vázquez, A., Flammini, A., Maritan, A., and Vespignani, A. Modeling of protein interaction networks. *Complexus*, **1**, 38–44 (2003). DOI: 10.1159/000067642.