Electronic Journal of Probability

Electron. J. Probab. **23** (2018), no. 123, 1–37.
ISSN: 1083-6489 https://doi.org/10.1214/18-EJP248

A central limit theorem for the gossip process

A. D. Barbour* Adrian Röllin†

Abstract

The Aldous gossip process represents the dissemination of information in geographical space as a process of locally deterministic spread, augmented by random long range transmissions. Starting from a single initially informed individual, the proportion of individuals informed follows an almost deterministic path, but for a random time shift, caused by the stochastic behaviour in the very early stages of development. In this paper, it is shown that, even with the extra information available after a substantial development time, this broad description remains accurate to first order. However, the precision of the prediction is now much greater, and the random time shift is shown to have an approximately normal distribution, with mean and variance that can be computed from the current state of the process.

Keywords: gossip process; deterministic approximation; branching processes; central limit theorem.
AMS MSC 2010: 92H30; 60K35; 60J85.
Submitted to EJP on June 26, 2017, final version accepted on November 19, 2018.

1 Introduction

A model for the dissemination of information in space, in which random long-range contacts facilitate spread, was introduced in Aldous [1]. In an idealized version, proposed by Chatterjee & Durrett [6], individuals are represented as a continuum, evenly distributed over a two-dimensional torus of large area L. Information spreads locally at constant rate from individuals to their neighbours, so that a disc of informed individuals, centred on an initial informant, grows steadily in the torus. However, information is also spread by long range transmissions to other, randomly chosen points of the torus, according to a Poisson process, whose rate is proportional to the area of currently informed individuals. Any such transmission initiates a new disc of informed individuals.

*Institut für Mathematik, Universität Zürich, Winterthurertrasse 190, CH-8057 ZÜRICH.
E-mail: a.d.barbour@math.uzh.ch Work begun while ADB was Saw Swee Hock Professor of Statistics at the National University of Singapore, and supported in part by Australian Research Council Grants Nos DP120102728, DP120102398, DP150101459 and DP150103588.

†Department of Statistics and Applied Probability, National University of Singapore, 6 Science Drive 2, 117546 Singapore. E-mail: adrian.roellin@nus.edu.sg AR was supported in part by NUS Research Grant R-155-000-167-112 and Australian Research Council Grant No. DP150101459.

The process can also be interpreted as a model of the spread of an SI disease, in which local infection is supplemented by occasional long-range contacts.

With L_t denoting the area of informed individuals by time t, Chatterjee & Durrett [6] showed that, after some randomness in the initial stages of the process, the proportion of the torus L_t/L that has been informed by time t closely follows a particular, deterministic path. The times at which L_t/L increases from almost zero to almost one is relatively short, and occurs around a time t_L, which is a fixed multiple of $\log L$. In what follows, we therefore concentrate on times relative to t_L. Roughly speaking, Chatterjee & Durrett [6] showed that, for large L, we have

$$\frac{L_{t_L+u/\lambda}}{L} \approx \ell(u+U) \qquad \text{for any } u \in \mathbb{R}$$

for some function ℓ, where λ is a scaling factor related to the speed of spread of information, and where U is a random variable. The path ℓ is the same for all realizations of the process, but the position on the path at a particular time varies from realization to realization because of the random time shift U. This result was generalized to gossip processes on rather general homogeneous Riemannian manifolds by Barbour & Reinert [5], hereafter referred to as [BR], as well as to related 'small world' processes; they also derived a uniform bound on the approximation error. In addition, the equation describing the deterministic development was interpreted in terms of the Laplace transform of the limiting random variable corresponding to an associated Crump–Mode–Jagers (CMJ) branching process (see Jagers [9]).

By analogy with the theory of Markov population processes (Kurtz [11, 12]), one might expect that the fluctuations around the deterministic path of the proportions informed would be approximately Gaussian, with standard deviation $O(L^{-1/2})$, at least while the proportion informed is not too small or too close to 1. Here, however, the random quantity of most interest — the difference between the actual course of the process and a prediction of the course based on information available early in its development — involves the fluctuations of the process while the proportion informed is rather small, and the standard analogy does not apply. Instead, in view of the approximation already established, it seems reasonable at times $v \ll t_L$ to predict the value of $L_{t_L+u/\lambda}/L$ by $\ell(u+\widehat{U}(v))$, where $\widehat{U}(v)$ is the expected value of U, given the information at time v, and to augment the point prediction with a confidence interval around $\ell(u+\widehat{U}(v))$, derived from the (approximate) conditional distribution of $L_{t_L+u/\lambda}/L$, given the current information.

The validity of the procedure is justified in detail in Section 3. The broad argument is to exploit the fact that $L_{t_L+u/\lambda}/L$ is the probability that a point K, chosen independently and uniformly at random in $\mathcal{C}$, belongs to the informed set $\mathcal{L}_{t_L+u/\lambda}$:

$$L_{t_L+u/\lambda}/L \;=\; \mathbb{P}[K \in \mathcal{L}_{t_L+u/\lambda} \,|\, \mathcal{L}_{t_L+u/\lambda}].$$

As it stands, this changes nothing. However, it indicates that a good approximation might be obtained by replacing $\mathbb{P}[K \in \mathcal{L}_{t_L+u/\lambda} \,|\, \mathcal{L}_{t_L+u/\lambda}]$ by $\mathbb{P}[K \in \mathcal{L}_{t_L+u/\lambda} \,|\, \mathcal{L}_s]$, or, equivalently, replacing $L_{t_L+u/\lambda}/L$ by $\mathbb{E}\{L_{t_L+u/\lambda}/L \,|\, \mathcal{L}_s\}$, for $s < t_L + u/\lambda$ chosen so that s is close enough to $t_L + u/\lambda$. In particular, for prediction from v, we need to choose $s \in (v, t_L + u/\lambda)$ so that

$$\mathbb{E}_v\big|(L_{t_L+u/\lambda}/L) - \mathbb{E}\{L_{t_L+u/\lambda}/L \,|\, \mathcal{L}_s\}\big| \;\ll\; \mathrm{SD}_v(L_{t_L+u/\lambda}/L), \tag{1.1}$$

where $\mathbb{E}_v$ and SD_v denote expectation and standard deviation given the information at time v.

The advantage of using $\mathbb{E}\{L_{t_L+u/\lambda}/L \,|\, \mathcal{L}_s\}$ is that $\mathbb{P}[K \in \mathcal{L}_{t_L+u/\lambda} \,|\, \mathcal{L}_s]$ can be approximated as the probability of at least one of many small balls, with centres chosen independently and at random in $\mathcal{C}$, intersecting $\mathcal{L}_s$. These balls are the islands in an independent 'backwards' gossip process, run for a length of time $t_L+(u/\lambda)-s$ from K. There are many such balls if $t_L+(u/\lambda)-s$ is not too small, and the intersection probability can be approximated by a Poisson probability, using the Stein–Chen method; see Lemma 3.4. The mean of the Poisson distribution can, with considerable effort, be shown to be close to $\ell(\log[CW(s,v)]+u)$, where $W(s,v)$ is a quantity that can be simply expressed in terms of a carefully chosen branching process, and C is a constant. Now, given the information available at time v, the quantity $W(v,v)$ (which loosely corresponds to $\exp\{\widehat{U}(v)\}$) is known, and the conditional distribution of the difference $W(s,v)-W(v,v)$ is approximately normal, as is shown in Theorem 2.8 in Section 2. This, in turn, leads to a normal approximation for the difference between $\ell(\log[CW(s,v)]+u)$ and its prediction $\ell(\log[CW(v,v)]+u)$ at time v. This implies the main result of the paper, that

$$\sigma^{-1}\big(L_{t_L+u/\lambda}/L-\ell(\log[CW(v,v)]+u)\big)\ \approx_d\ \mathcal{N}(0,1), \tag{1.2}$$

for suitable choice of the standard deviation σ depending on u and $W(v,v)$; a precise statement is given in Theorem 1.1. The error in the normal approximation is shown to be small if the *number* of individuals informed at time v is large, even if their proportion in the whole population may be very small. For practical purposes, in an epidemic, the very earliest development may well pass almost unnoticed — the origins are often obscure — but prediction on the basis of the information gained from the first few hundred cases is an important public health goal, in which case using the normal approximation is reasonable.

1.1 Detailed formulation

We now describe the problem in more detail. We consider the gossip process $(\mathcal{L}_t,\ t\geq 0)$ evolving on a smooth closed homogeneous Riemannian manifold $\mathcal{C}$ of dimension d, such as a sphere or a torus, having large finite volume $|\mathcal{C}| =: L$ with respect to its intrinsic metric. An individual at point $P\in\mathcal{C}$ informed at time 0 gives rise to deterministic local spread that informs the set $\mathcal{K}(P,s)$ by time $s>0$; in addition, random 'long range transmissions' to independent and uniformly distributed points of $\mathcal{C}$ occur at rate ρ times the intrinsic volume of the set currently informed. Thus the process can be constructed from knowledge of the points $0=\tau_0<\tau_1<\cdots$ of a point process Π on $\mathbb{R}_+$ (characterized immediately below), together with an independent sequence of independent points $P_1,P_2,\ldots$, uniformly distributed in $\mathcal{C}$, and an initial point $P=P_0$. The informed set and its volume are denoted by

$$\mathcal{L}_t\ :=\ \bigcup_{j:\ \tau_j\leq t}\mathcal{K}(P_j,t-\tau_j)\quad\text{and}\quad L_t\ :=\ |\mathcal{L}_t|. \tag{1.3}$$

The point process Π is simple, and has conditional intensity ρL_t at time t with respect to the filtration $(\mathcal{F}_t,\ t\geq 0)$, where $\mathcal{F}_t:=\sigma((\tau_j,P_j),\ j\geq 0,\tau_j\leq t)$.

The sets $\mathcal{K}(P,s)$ are assumed to be closed balls, centred at P and of radius s, with respect to a metric that makes $\mathcal{C}$ a geodesic space: $P'\in\mathcal{K}(P,2t)$ exactly when $\mathcal{K}(P,t)\cap\mathcal{K}(P',t)\neq\emptyset$. Since $\mathcal{C}$ is assumed to be homogeneous, the volume of $\mathcal{K}(P,s)$ is independent of P, and we will therefore denote it by $\nu_s=\nu_s(\mathcal{K})$. The sets $\mathcal{K}(P,s)$ are also assumed to be locally almost Euclidean in the sense that $\nu_s\approx s^d\nu$ for some constant $\nu=\nu(\mathcal{K})>0$. More precisely, we will assume that, for constants $c_g,\gamma_g>0$,

$$\left|\frac{\nu_s}{s^d\,\nu}-1\right|\ \leq\ c_g\left(\frac{s^d\,\nu}{L}\right)^{\gamma_g/d},\quad s>0. \tag{1.4}$$

The quantity $\nu>0$ has physical dimensions $(\text{length}/\text{time})^d$, so that $\nu^{1/d}$ can be interpreted as a local velocity of spread of information in any particular direction. Assumption (1.4) is satisfied, for instance, for balls with respect to geodesic distance on the surface of a $(d+1)$-dimensional sphere of large radius R, when $L = c_d R^d$ and

$$\frac{\nu_s}{s^d\,\nu} - 1 \;=\; \frac{dR^d}{s^d}\int_0^{s/R} (\sin t)^{d-1}\,dt - 1 \;=\; O\big((s/R)^2\big),$$

(Li, 2011), in which case we can take $\gamma_g = 2$ in all dimensions $d \geq 2$.

Using (1.4), the probability of there being no long range transmission before time u is given by

$$\exp\Big\{-\int_0^u \rho\nu_s\,ds\Big\} \;\approx\; \exp\Big\{-\int_0^u \rho s^d \nu\,ds\Big\} \;=\; \exp\big\{-\rho\nu u^{d+1}/(d+1)\big\},$$

so that the mean time to the first long range transmission is approximately

$$\int_0^\infty \exp\big\{-\rho\nu u^{d+1}/(d+1)\big\}\,du \;=\; (\rho\nu)^{-1/(d+1)}\int_0^\infty e^{-w^{d+1}/(d+1)}\,dw.$$

Thus

$$\lambda \;:=\; (\rho d!\nu)^{1/(d+1)}, \tag{1.5}$$

having physical dimensions $(1/\text{time})$, is such that $1/\lambda$ represents the time scale for the first long range transmission, and then $\lambda^{-d}\nu$ reflects the size of the initial neighbourhood when the first long range transmission occurs; the exact specification of λ is to make it equal to the growth rate of the associated CMJ process ([BR], p.986). For our approximations to be good, the size of the initial neighbourhood when the first long range transmission occurs should be small compared to L, so that, defining

$$\Lambda \;:=\; L\lambda^d/\nu, \tag{1.6}$$

a quantity without physical dimension, we shall take Λ to be large. Note that, if this is so, the approximations made above have small error, in view of (1.4).

To start with, the points of Π closely match the birth events of a CMJ process $\overline{X}$, whose birth intensity as a function of age s is given by $\rho\nu_s$. In fact, the approximation $\overline{\mathcal{L}}_t$ of $\mathcal{L}_t$, constructed by using the CMJ process $\overline{X}$ to approximate Π and with the same sequence of points $(P_j,\, j \geq 1)$, is excellent for times $t \leq \alpha\lambda^{-1}\log\Lambda$ if $\alpha < 1/2$ ([BR], §2.2), and still gives an approximation to the volume L_t of $\mathcal{L}_t$ at time t that is accurate to the first order if $\alpha < 1$ ([BR], Theorem 3.2 and (2.23)). This CMJ approximation takes the form

$$L_t/L \;\sim\; K\Lambda^{-1}e^{\lambda t + \log W}, \quad t \to \infty, \tag{1.7}$$

for a constant K, where W is a limiting random variable associated with the CMJ process $\overline{X}$. Taking

$$t \;=\; t_\Lambda(u) \;:=\; \lambda^{-1}(\log\Lambda + u), \tag{1.8}$$

with $u \leq (\alpha-1)\log\Lambda$ *large and negative* in the range in which this approximation holds, this implies that $L_{t_\Lambda(u-\log W)}/L$ closely follows the curve

$$u \;\mapsto\; \ell_0(u), \tag{1.9}$$

where $\ell_0(u) := Ke^u$.

In [BR], Theorem 3.2, an analogous approximation

$$L_{t_\Lambda(u-\log W)}/L \;\approx\; \ell(u + \log\hat{c}_d)$$

is established, with uniformly small error, for *all* values of u, with $\hat{c}_d$ defined before (1.11), and with the time shift U given by $\lambda^{-1}\log W + c$, for a suitably chosen constant c.

Clearly, to be compatible with (1.9), $\ell(u) \sim Ke^u$ as $u \to -\infty$, as follows from ([BR], following (2.23)).

For any fixed u, the distribution of $L_{t_\Lambda(u)}/L$ is close to that of $\ell(u + \log W + \log \hat{c}_d)$, and is a bounded random variable. Hence it can only be approximately normally distributed, after appropriate centring and normalization, in circumstances in which the distribution of $\log W$ is concentrated close to some fixed value. This is not true of the distribution of W at time 0. However, when predicting from a time $v = \alpha\lambda^{-1} \log \Lambda$ for any fixed α, $0 < \alpha < 1$, the conditional distribution of W, given the information up to time v, is concentrated close to an approximation $W(v, v)$ provided only that $\alpha > 0$, even though the size of the informed set is still relatively small when compared to L for any $\alpha < 1$. The aim is now to show that the difference $\Delta(v) := W(v, v) - W$, suitably normalized, is approximately normally distributed.

It turns out to be easier to work with a 'flattened' CMJ process $\widehat{X}$, rather than with the original CMJ process $\overline{X}$. The process $\widehat{X}$ has birth rate at age s given by $\rho s^d \nu$, and is thus the same process for all L, whereas $\overline{X}$ depends implicitly on L through the function ν_s. The quantity λ then turns out to be the Malthusian parameter of $\widehat{X}$. In a CMJ process with Malthusian parameter μ, at large times, a randomly sampled individual has average age approximately $1/\mu$. For $\widehat{X}$, $\mu = \lambda$, and replacing s by $1/\lambda$ in (1.4) confirms that the two CMJ processes $\overline{X}$ and $\widehat{X}$ have birth rates that are close to each other if Λ is large. The essentials of the proof of the normal approximation to $\Delta(v)$ are carried out in Section 2. The argument hinges on examining a collection of (complex valued) martingales $(W_j(\cdot), 0 \le j \le d)$ associated with $\widehat{X}$, that are defined in (2.13) below. In particular, $W(t, v) := W_0(t)$, $t \ge v$, is non-negative and square integrable, having limit $W_0(\infty) =: W$. It is then shown that $W_0(v) - W$, suitably normalized, is close enough to the integral of a function $f(W_0(v), u)$ with respect to an independent standard Brownian motion $B(u)$, giving the normal approximation.

The arguments in Section 3, as outlined before (1.2), rely heavily on comparisons between birth and growth processes. The actual process $(\mathcal{L}_t, t \ge 0)$ is compared with the branching approximation $\overline{X}$, and $\overline{X}$ is compared to its flattened version $\widehat{X}$. Further (flattened) CMJ processes $\widehat{X}^+$ and $\widehat{X}^-$ are then introduced, to act as upper and lower bounds for $\overline{X}$; the comparison is formalized in Lemma 3.1. All the detailed computations in Section 3 are made using these processes, including the reduction of the intersection probability in Lemma 3.4 to a tractable form in Lemma 3.7.

To state our theorem, we take

$$\widehat{W}(v) \;:=\; e^{-\lambda v} \sum_{l=0}^{d} \sum_{j \in \widehat{J}_v} \frac{\{\lambda(v - \tau_j)\}^l}{l!} \tag{1.10}$$

as an approximation to W, where the set $\widehat{J}_v$ indexes the set of all *non-intersecting* neighbourhoods of $\mathcal{L}_v$. For each of these, the radii $(v - \tau_j)$ can be determined, and so $\widehat{W}(v)$ can be derived from $\mathcal{L}_v$. Then let $\hat{c}_d := d!/(d+1)$, and

$$\zeta(d) \;:=\; \begin{cases} 1/2 & \text{if } d \le 6, \\ 1 - \cos(2\pi/d) & \text{if } d \ge 7, \end{cases} \tag{1.11}$$

and define

$$\ell(u) \;:=\; 1 - \phi_\infty(e^u), \quad \text{where} \quad \phi_\infty(\theta) \;:=\; \mathbb{E}\{e^{-\theta W}\}, \tag{1.12}$$

where W is as above; see also (2.13) and (2.18). Let d_{BW} denote the bounded Wasserstein distance between probability measures on $\mathbb{R}$:

$$d_{\mathrm{BW}}(P, Q) \;:=\; \sup_{f \in F_{\mathrm{BW}}} \Big\{ \Big| \int f\, dP - \int f\, dQ \Big| \Big\},$$

where F_{BW} consists of all Lipschitz functions $f\colon \mathbb{R} \to [-1,1]$ whose Lipschitz constant is at most 1. The theorem is as follows.

Theorem 1.1. *With the above definitions, suppose that $v = \alpha\lambda^{-1}\log\Lambda$ for $0 < \alpha < 2\min\{\gamma_g/d, \zeta(d)/(1+\zeta(d))\}$, where γ_g is as in* (1.4)*. Then, for any $u_1 < u_0 \in \mathbb{R}$, there exists a $\gamma > 0$ and an event $E^*(v) \in \sigma(\mathcal{L}_v)$ with $\mathbb{P}[E^*(v)^c] = \mathrm{O}(\Lambda^{-\gamma})$ such that*

$$d_{\mathrm{BW}}\big(\mathcal{L}\big\{e^{\lambda v/2}\{L_{t_\Lambda(u)}/L - \ell(u+\log[\hat{c}_d\widehat{W}(v)])\}\,\big|\,\mathcal{F}_v \cap E^*(v)\big\}, \mathcal{N}(0, \sigma^2(u, \widehat{W}(v)))\big) \\ = \mathrm{O}(\Lambda^{-\gamma}),$$

uniformly in $u_1 \le u \le u_0$, where $t_\Lambda(u) = \lambda^{-1}(\log\Lambda + u)$ as in (1.8) *and*

$$\sigma^2(u,w) \;:=\; \frac{\{D\ell(u+\log[\hat{c}_d w])\}^2}{(d+1)w}.$$

So, for instance, for spherical neighbourhoods in $d \le 6$, it is possible to take any α strictly between 0 and $2/3$ in Theorem 1.1. The order statements can be replaced by inequalities, valid for all Λ sufficiently large, in which the constants depend only on d, u_1 and u_0; however, the lower bound on the value of Λ then also involves α and the constants c_g and γ_g from (1.4).

In fact, the proof shows a little more: that we could realize the normal random variables $\mathcal{N}(0, \sigma^2(u, W(v,v)))$, for different values of u, as $\sigma(u, W(v,v))N$ for the *same* standard normal random variable N. The interpretation of this is that the fluctuations in $L_{t_\Lambda(u)}/L$ are essentially those of $\ell(u+\log[\hat{c}_d W])$, and that the remaining randomness after time v is overwhelmingly that of the difference $W - W(v,v)$, a single random variable. This, at first sight surprising, result reflects the phenomenon common to branching processes, that the randomness determining the growth of a super-critical branching process occurs at the very beginning of its development.

2 The branching process

In this section, we investigate the limit W, as $t \to \infty$, of a martingale $W(t)$ associated with a particular CMJ branching process. We show that $(W(t) - W)$ is approximately normally distributed, and give an explicit bound on the accuracy of the approximation. Although, for a (multitype) Galton–Watson process, a central limit theorem of this sort is not difficult to establish (Asmussen & Hering [2, Theorem 7.1]), the corresponding theorems for general CMJ processes seem not to be available. Here, we are able to exploit the particular structure of our CMJ process to prove what we need.

We start by identifying the branching process that we work with, which can be expressed as a Markov process in a $(d+1)$-dimensional space. The properties of the coordinate processes $(H_j(t),\, 0 \le j \le d)$, and of some equivalent (complex valued) martingales $(W_j(t),\, 0 \le j \le d)$ are established in Lemma 2.1. The component W_0 is a non-negative real valued martingale, and W is its limit as $t \to \infty$. Using Kolmogorov's inequality, the fluctuations of the sample paths of the processes W_j are controlled in Lemma 2.2, and this in turn gives control over the processes H_j.

The martingale difference $W_0(v+t) - W_0(v)$ is written in (2.23) as an integral of an explicit function of the process $H_{d+1}(u) := \lambda\int_0^h H_d(w)\,dw$ with respect to a standard compensated Poisson process. Using the control that we have over the H_j, we determine successively simpler approximations to this process, in (2.29) and (2.31), at each stage making sure that the error incurred is sufficiently small (Lemma 2.4 and Corollary 2.6). Finally, in (2.35), an expression is obtained in which integration with respect to the compensated Poisson process has been replaced by integration with respect to standard Brownian motion, and this can be used with an error controlled in Lemma 2.7. The results of these steps are collected as a functional approximation in Theorem 2.8. The version that is used to prove Theorem 3.10 in Section 3 is given as Corollary 2.10.

2.1 Properties of the flattened process

The first step is to determine a suitable W. We do so by way of a 'flattened' version $\widehat{X}$ of the CMJ branching process $\overline{X}$. The process $\widehat{X}$ is the counting process associated with a point process $(\hat{\tau}_j,\, j \geq 0)$ on $\mathbb{R}_+$, with $\hat{\tau}_0 = 0$ a.s., whose compensator is given by $\widehat{A}(t) := \int_0^t \hat{a}(u)\,du$, where $\hat{a}(u) := \rho\nu \sum_{j:\,\hat{\tau}_j \leq u}(u - \hat{\tau}_j)^d$, and where ρ, as before, denotes the intensity per unit volume. At time t, $\widehat{X}(t)$ can be thought of as consisting of $M_0(t) := 1 + \max\{r \geq 0\colon\ \hat{\tau}_r \leq t\}$ neighbourhoods, whose volumes at time t are given by $(t - \hat{\tau}_r)^d\nu$, asymptotically close to, but not the same as the volume $\nu_{t-\hat{\tau}_r}$. The intensity $\hat{a}$ is then precisely that of a CMJ process, in which neighbourhoods play the part of individuals, and the point process ξ of an individual's offspring is an inhomogeneous Poisson process with rate $\rho\nu s^d$ at age s. The mean number of offspring of an individual is thus infinite, but the Malthusian parameter λ, chosen so that the equation

$$\int_0^\infty e^{-\lambda s}\rho\nu s^d\,ds \;=\; 1$$

is satisfied, is finite, and is given by $\lambda := (d!\rho\nu)^{1/(d+1)}$. Note that

$$(\xi(t),\, t \geq 0) \;=_d\; (\xi^1(\lambda t),\, t \geq 0), \tag{2.1}$$

where ξ^1 is the inhomogeneous Poisson process with rate $s^d/d!$ at age s.

We can immediately deduce some useful general properties of the process $\widehat{X}$. To start with, because the variance of the discounted offspring number $\int_0^\infty e^{-\lambda s}\xi(ds)$ is finite, being given by $\int_0^\infty e^{-2\lambda s}\rho\nu s^d\,ds$, it follows from Ganuza & Durham [7, Theorem 1] that there exist finite constants c_1 and c_2 such that, for all $u > 0$,

$$e^{-\lambda u}\mathbb{E}M_0(u) \;\leq\; c_1; \qquad e^{-2\lambda u}\mathbb{E}\{M_0^2(u)\} \;\leq\; c_2; \tag{2.2}$$

in view of (2.1), c_1 and c_2 depend only on d. Then the intensity $\hat{a}(u)$ can be expressed as $\rho\nu M_d(u)$, where

$$M_d(u) \;=\; u^d + \int_{(0,u]}(u-v)^d M_0(dv) \;=\; d\int_0^u (u-v)^{d-1}M_0(v)\,dv. \tag{2.3}$$

This in turn implies from (2.2) that

$$e^{-\lambda u}\mathbb{E}M_d(u) \;\leq\; c_1 d!\lambda^{-d}; \qquad e^{-2\lambda u}\mathbb{E}\{M_d^2(u)\} \;\leq\; c_2\{d!\lambda^{-d}\}^2, \qquad u > 0, \tag{2.4}$$

using Cauchy–Schwarz for the second inequality.

However, $\widehat{X}$ also has special structure that will prove useful in what follows, relating to the sums

$$M_l(t) \;=\; \sum_{j=1}^{M_0(t)}(t - \tau_{j-1})^l, \quad l \geq 1, \tag{2.5}$$

of the l-th powers of the ages of the neighbourhoods. Note that $M_d(t)$ is as defined previously, and that

$$\begin{aligned} \frac{d}{dt}M_1(t) &= M_0(t) \quad \text{for a.e. } t; \\ \frac{d}{dt}M_i(t) &= iM_{i-1}(t), \quad i \geq 2. \end{aligned} \tag{2.6}$$

Since M_0 has intensity $\hat{a} = \rho\nu M_d$, letting Z denote a unit rate Poisson process, we can write

$$M_0(t) \;=\; M_0(0) + Z\left(\rho\nu \int_0^t M_d(u)\,du\right). \tag{2.7}$$

Defining $H_i(t) := M_i(t)\lambda^i/i!$, for any $\lambda > 0$, the equations (2.6) reduce to

$$\begin{aligned} \frac{d}{dt}H_1(t) &\;=\; \lambda H_0(t) \quad \text{for a.e. } t; \\ \frac{d}{dt}H_i(t) &\;=\; \lambda H_{i-1}(t), \quad i \geq 2; \end{aligned} \tag{2.8}$$

with the particular choice $\lambda := (d!\rho\nu)^{1/(d+1)}$, equation (2.7) becomes

$$H_0(t) \;=\; M_0(0) + Z\left(\rho\nu \int_0^t d!\lambda^{-d} H_d(u)\,du\right) \;=\; H_0(0) + Z(H_{d+1}(t)), \tag{2.9}$$

so that $\widehat{A}(t) = H_{d+1}(t)$. In particular, from (2.8) and (2.9), it follows that the process $\widetilde{H}$ defined by

$$\widetilde{H}(t) \;:=\; (H_0(t), H_1(t), \ldots, H_d(t)) \tag{2.10}$$

is a Markov process. It also follows directly from (2.8) and (2.9), or as a consequence of (2.1), that

$$\{\widetilde{H}(t),\, t \geq 0\} \;=_d\; \{\widetilde{H}^1(\lambda t),\, t \geq 0\}, \tag{2.11}$$

where $\widetilde{H}^1$ denotes the process with $\lambda = 1$. Note that ρ may depend on L, as also may λ.

In order to describe the properties of the process $\widehat{X}$ in more detail, we introduce the (complex valued) processes

$$W_j(t) \;=\; 1 + \int_{(0,t]} e^{-\lambda x_j u}\{M_0(du) - \widehat{A}(du)\}, \tag{2.12}$$

where $x_j := \exp\{2\pi \imath j/(d+1)\} \in \mathbb{C}$, $j \in \{0, 1, \ldots, d\}$, which are martingales with respect to the natural filtration $(\widehat{\mathcal{F}}_t,\, t \geq 0)$ of $\widehat{X}$. In particular, for $j = 0$, we have $x_j = 1$, and

$$W(t) \;:=\; W_0(t) \;=\; 1 + \int_{(0,t]} e^{-\lambda u}\{M_0(du) - \widehat{A}(du)\} \tag{2.13}$$

is a real valued, càdlàg martingale, and plays a key part our arguments. It is shown in the next lemma that it is also non-negative, and the rest of the section is then devoted to proving a normal approximation to $e^{\lambda t/2}(W(t) - W(\infty))$, which is the basis for the central limit theorem for the gossip process itself. Note that the distribution of $W(\cdot)$ can be derived from the corresponding martingale $W^1(\cdot)$ for the process with $\lambda = 1$, since, from (2.11),

$$\{W(t),\, t \geq 0\} \;=_d\; \{W^1(\lambda t),\, t \geq 0\}; \tag{2.14}$$

from this, it also follows that the distribution of $W(\infty)$ is the same for all λ. The remaining martingales W_j are useful, because they enable the quantities $H_j(\cdot)$ to be expressed in a tractable form, as in the next lemma.

Lemma 2.1. *With notation as above, we have*

$$W(t) \;=\; \sum_{r=0}^{d} e^{-\lambda t} H_r(t) \;>\; 0, \quad t \geq 0,$$

and

$$e^{-\lambda t} H_j(t) \;=\; \frac{1}{d+1} \sum_{l=0}^{d} x_j^l \, e^{-\lambda(1-x_l)t}\, W_l(t).$$

Proof. It follows from (2.8) that, for any $x \in \mathbb{C}$,

$$\frac{d}{dt}\{e^{-\lambda xt}x^r H_r(t)\} \;=\; \lambda x e^{-\lambda xt}\{-x^r H_r(t) + x^{r-1}H_{r-1}(t)\}, \quad r \geq 1,$$

and, by partial integration, that

$$\int_{[0,t]} e^{-\lambda xu} H_0(du) \;=\; e^{-\lambda xt}H_0(t) + \lambda x \int_0^t e^{-\lambda xu} H_0(u)\,du.$$

Hence

$$\frac{d}{dt}\sum_{r=1}^{d}\{e^{-\lambda xt}x^r H_r(t)\} \;=\; \lambda x e^{-\lambda xt}\{-x^d H_d(t) + H_0(t)\},$$

and thus

$$\sum_{r=0}^{d}\{e^{-\lambda xt}x^r H_r(t)\} \;=\; \int_{[0,t]} e^{-\lambda xu}\{H_0(du) - \lambda x^{d+1} H_d(u)\,du\}. \tag{2.15}$$

Taking $x = x_j$ for any $j \in \{0,1,\ldots,d\}$, we have $x^{d+1} = 1$, making the right hand side equal to $W_j(t)$, because $\lambda H_d(u)\,du = H_{d+1}(du) = \widehat{A}(du)$, by (2.8) and (2.9); hence

$$W_j(t) \;=\; \sum_{r=0}^{d}\{e^{-\lambda x_j t}x_j^r H_r(t)\}. \tag{2.16}$$

The first statement of the lemma follows by taking $j = 0$, and the second by using the orthogonality relation $\sum_{l=0}^{d} x_j^l x_l^r = (d+1)\delta_{jr}$. ∎

Now, writing $r_j := \Re x_j$ and noting that $\hat{a}(u) = \lambda H_d(u) \leq \lambda e^{\lambda u} W(u)$, it follows from (2.12) that, for $0 \leq j \leq d$ and for $v < t < w$,

$$\begin{aligned}\mathbb{E}\{|W_j(w) - W_j(t)|^2 \,|\, \widehat{\mathcal{F}}_v\} \;&=\; \int_{(t,w]} e^{-2\lambda r_j u}\,\mathbb{E}\{\hat{a}(u)\,|\,\widehat{\mathcal{F}}_v\}\,du \\ &\leq\; W(v)\int_{(t,w]} \lambda e^{-\lambda(2r_j - 1)u}\,du.\end{aligned} \tag{2.17}$$

Using this bound with $v = 0$, we see that the variances of the terms with $1 \leq l \leq d$ in the sum in Lemma 2.1 converge to zero as $t \to \infty$. However, the term with $l = 0$ remains significant as $t \to \infty$, since, by (2.17) with $v = 0$ and $j = 0$, it follows that $W(\cdot)$ is square integrable, and that

$$W \;:=\; W(\infty) \;:=\; \lim_{t\to\infty} W(t) \quad \text{exists a.s.;} \quad \text{and } \mathbb{E}W \;=\; 1, \quad \operatorname{Var} W \leq 1. \tag{2.18}$$

Note that the distribution of W, through its Laplace transform ϕ_∞ as in (1.12), already appears in the statement of Theorem 1.1, and is the same for all λ, as remarked following (2.14). Thus each of the H_j satisfies

$$e^{-\lambda t}H_j(t) \;\to_P\; W/(d+1) \quad \text{as } t \to \infty. \tag{2.19}$$

We shall exploit more detailed versions of these asymptotics in Section 3.

In order to use Lemma 2.1 to describe further the behaviour of the $H_j(t)$, we need good control of the fluctuations of the processes $(W_l,\, 0 \leq l \leq d)$. As indicated by (2.17), their asymptotic behaviour depends substantially on whether or not $r_l > 1/2$. Note, for future reference, that $\min\{(1 - r_1), 1/2\} = \zeta(d)$, where $\zeta(d)$ is as in (1.11).

Lemma 2.2. *For any* $1 \le l \le d$ *and* $0 < \eta < \min\{(1-r_l), 1/2\}$*, and for any* $K > 0$*, define the events*

$$E_{1l}^{\eta}(v;K) \quad := \quad \Big\{\sup_{t\ge v}\{e^{-\lambda t(1-r_l-\eta)}|W_l(t)-W_l(v)|\} \le K\Big\};$$

similarly, for $0 < \eta < 1/2$*, define*

$$E_{10}^{\eta}(v;K) \;:=\; \Big\{\sup_{t\ge v}\{e^{\lambda\eta t}|W(t)-W(\infty)|\} \le K\Big\}.$$

Then there exist constants $C(l,\eta)$*,* $0 \le l \le d$*, such that, for all* $K > 0$*,*

$$\mathbb{P}[\{E_{1l}^{\eta}(v;K)\}^c \,|\, \widetilde{H}(v)] \quad \le \quad C(l,\eta)K^{-2}W(v)e^{-\lambda(1-2\eta)v}.$$

Proof. Combining (2.16) with (2.10), it follows that $\mathcal{L}\big((W_0(s),\ldots,W_d(s)),\ s \ge v \,|\, \widehat{\mathcal{F}}_v\big)$ depends on $\widehat{\mathcal{F}}_v$ only through the value of $\widetilde{H}(v)$. Then, noting that, for $r+\eta \le 1$, $1 \le l \le d$ and for any $w > t \ge v$,

$$\sup_{t\le s\le w}\{e^{-\lambda s(1-r-\eta)}|W_l(s)-W_l(v)|\} \;\le\; e^{-\lambda t(1-r-\eta)}\sup_{t\le s\le w}|W_l(s)-W_l(v)|,$$

and using Kolmogorov's inequality on the real and imaginary parts of W_l, it follows that

$$\begin{aligned}\mathbb{P}\Big[&\sup_{t\le s\le w}\{e^{-\lambda s(1-r_l-\eta)}|W_l(s)-W_l(v)|\} \ge K \,\Big|\, \widetilde{H}(v)\Big]\\ &\le\; 4K^{-2}e^{-2\lambda t(1-r_l-\eta)}\mathbb{E}\{|W_l(w)-W_l(v)|^2 \,|\, \widetilde{H}(v)\}\,.\end{aligned}$$

For $r_l > 1/2$, taking $w = \infty$, it follows from (2.17) that

$$\begin{aligned}\mathbb{P}\Big[&\sup_{s\ge v}\{e^{-\lambda s(1-r_l-\eta)}|W_l(s)-W_l(v)|\} \ge K \,\Big|\, \widetilde{H}(v)\Big]\\ &\le\; 4K^{-2}e^{-2\lambda v(1-r_l-\eta)}W(v)e^{-\lambda v(2r_l-1)}/(2r_l-1) \;=\; 4K^{-2}W(v)e^{-\lambda v(1-2\eta)}/(2r_l-1).\end{aligned}$$

For $r_l = 1/2$, taking $t = v + j\lambda^{-1}$ and $w = v + (j+1)\lambda^{-1}$, it follows from (2.17) that

$$\mathbb{P}\Big[\sup_{t\le s\le w}\{e^{-\lambda s(1-r_l-\eta)}|W_l(s)-W_l(v)|\} \ge K \,\Big|\, \widetilde{H}(v)\Big] \;\le\; 4K^{-2}W(v)e^{-(\lambda v+j)(1-2\eta)}(j+1),$$

and adding over $j \in \mathbb{Z}_+$ gives

$$\mathbb{P}\Big[\sup_{s\ge v}\{e^{-\lambda s(1-r_l-\eta)}|W_l(s)-W_l(v)|\} \ge K \,\Big|\, \widetilde{H}(v)\Big] \;\le\; \frac{4W(v)e^{-\lambda v(1-2\eta)}}{K^2(1-e^{-(1-2\eta)})^2}\,.$$

For $r_l < 1/2$, taking $t = v + j\lambda^{-1}$ and $w = v + (j+1)\lambda^{-1}$, it follows from (2.17) that

$$\mathbb{P}\Big[\sup_{t\le s\le w}\{e^{-\lambda s(1-r_l-\eta)}|W_l(s)-W_l(v)|\} \ge K \,\Big|\, \widetilde{H}(v)\Big] \;\le\; \frac{4W(v)e^{-(\lambda v+j)(1-2\eta)}\,e^{1-2r_l}}{K^2(1-2r_l)}\,,$$

and adding over $j \in \mathbb{Z}_+$ gives

$$\mathbb{P}\Big[\sup_{s\ge v}\{e^{-\lambda s(1-r_l-\eta)}|W_l(s)-W_l(v)|\} \ge K \,\Big|\, \widetilde{H}(v)\Big] \;\le\; \frac{4eW(v)e^{-\lambda v(1-2\eta)}}{K^2(1-e^{-(1-2\eta)})(1-2r_l)}.$$

For $l = 0$, the result is proved in analogous fashion, starting from

$$\sup_{t\le s\le t+\lambda^{-1}}\{e^{\lambda\eta s}|W(s)-W(\infty)|\} \;\le\; 2e^{\eta(\lambda t+1)}\sup_{s\ge t}|W(s)-W(t)|,$$

and observing that, from (2.17),

$$\mathbb{P}\big[\sup_{s\geq t}|W(s)-W(t)|>a\,|\,\widetilde{H}(v)\big] \;\leq\; a^{-2}\mathbb{E}\{W(t)e^{-\lambda t}\,|\,\widetilde{H}(v)\} \;=\; a^{-2}e^{-\lambda t}W(v). \qquad \blacksquare$$

As a result of this lemma, we can sharpen (2.19) by giving an explicit bound on the error made when approximating $e^{-\lambda t}H_j(t)$ by $W(v)/(d+1)$ for any $t \geq v$. To state the bound, we define

$$Q(v) \;:=\; d+2+\sum_{l=1}^{d} e^{-\lambda(1-r_l-\eta)v}\,|W_l(v)|; \qquad E_1^{\eta}(v) \;:=\; \bigcap_{l=0}^{d} E_{1l}^{\eta}(v;1), \tag{2.20}$$

noting that, on $E_1^{\eta}(v)$, $Q(t) \leq Q(v)+d$ for all $t \geq v$. Then for all $t \geq v$ and $0 \leq j \leq d$, and if $\eta < \zeta(d)$, we have

$$\begin{aligned}
&\left|e^{-\lambda t}H_j(t)-\frac{W(v)}{(d+1)}\right| \\
&\quad\leq\; \frac{1}{d+1}\left\{|W(t)-W(v)|+\sum_{l=1}^{d} e^{-\lambda(1-r_l)t}\,\{|W_l(v)|+|W_l(t)-W_l(v)|\}\right\} \\
&\quad\leq\; \frac{1}{d+1}\left\{\sum_{l=1}^{d} e^{-\lambda(1-r_l)t}\,|W_l(v)|+(d+2)e^{-\lambda\eta t}\right\} \;\leq\; \frac{e^{-\lambda\eta v}Q(v)}{d+1}\,,
\end{aligned} \tag{2.21}$$

on $E_1^{\eta}(v)$. Furthermore, from Lemma 2.2,

$$\mathbb{P}[\{E_1^{\eta}(v)\}^c\,|\,\widetilde{H}(v)] \;\leq\; \theta_1(v) \;:=\; W(v)e^{-\lambda(1-2\eta)v}\Big(C(0,\eta)+\sum_{l=1}^{d} C(l,\eta)\Big). \tag{2.22}$$

2.2 Approximating an integral representation of $W(v+t)-W(v)$

The aim of this section is to prove an approximation theorem, when v is large, for the process $X_v^{(0)}(t) := W(v+t)-W(v)$ in $t \geq 0$. We recall (2.7) and (2.9), and use the representation (2.12), writing

$$\begin{aligned}
X_v^{(0)}(t) \;&=\; \int_v^{v+t} e^{-\lambda u}\{M_0(du)-H_{d+1}(du)\} \\
&=\; \int_{H_{d+1}(v)}^{H_{d+1}(v+t)} e^{-\lambda H_{d+1}^{-1}(w)}\{Z^{(1)}(dw)-dw\}\,,
\end{aligned} \tag{2.23}$$

where $Z^{(1)}$ is a unit rate Poisson process, with increments independent of $\widehat{\mathcal{F}}_v$, starting with $Z^{(1)}(H_{d+1}(v)) = M_0(v) = H_0(v)$, and where $H_l(u)$, $l \geq 0$, are constructed in $u \geq v$ from the Poisson process $Z^{(1)}$, using (2.8) and (2.9), with initial values $H_l(v)$, $0 \leq l \leq d$. Once again, the process $X_v^{(0)}$ depends on its past $\widehat{\mathcal{F}}_v$ only through $\widetilde{H}(v)$. Since the expression (2.23) is too complicated to use directly, we simplify it in a series of stages.

We start by approximating $H_{d+1}^{-1}(w)$ in $w \geq H_{d+1}(v)$. In view of (2.21), we have $H_{d+1}(t) \approx e^{\lambda t}W(v)/(d+1)$, or $w \approx e^{\lambda H_{d+1}^{-1}(w)}W(v)/(d+1)$; the precise result is as follows. Note that, for our purposes, $\gamma^{\eta}(v)$ can be thought of as small.

Lemma 2.3. *Fix any $\eta < \zeta(d)$. Then, on the event $E_1^{\eta}(v)$, we have*

$$\frac{W(v)(1-\gamma^{\eta}(v))}{w(d+1)} \;\leq\; e^{-\lambda H_{d+1}^{-1}(w+H^*(v))} \;\leq\; \frac{W(v)(1+\gamma^{\eta}(v))}{w(d+1)},$$

for all $w \geq \{W(v)/(d+1)\}e^{\lambda v}$, where $\gamma^{\eta}(v) := (d+1)\{Q(v)/W(v)\}e^{-\lambda\eta v}$, $H^*(v) := H_{d+1}(v)-e^{\lambda v}W(v)/(d+1)$, and $Q(v)$ is as defined in (2.20).

Proof. We begin by noting that $H_{d+1}(u) = \int_0^u \lambda H_d(t)\,dt$, so that, from (2.21), for $u \geq v$,

$$\begin{aligned}\Big| H_{d+1}(u) - H_{d+1}(v) - (e^{\lambda(u-v)} - 1)e^{\lambda v}\frac{W(v)}{d+1}\Big| \\ \leq \int_v^u \lambda e^{\lambda t}\left\{\sum_{l=1}^d |W_l(v)|e^{-\lambda(1-r_l)t} + (d+2)e^{-\lambda\eta v}\right\} dt \\ \leq Q(v)e^{\lambda(u-v)}e^{\lambda(1-\eta)v}\,. \end{aligned} \tag{2.24}$$

So, defining

$$t_v(s) := \lambda^{-1}\log\left\{1 + \frac{s(d+1)}{e^{\lambda v}W(v)}\right\} \quad \text{and} \quad t_v^{-1}(u) := \frac{e^{\lambda v}W(v)}{d+1}(e^{\lambda u} - 1), \tag{2.25}$$

it follows that, on $E_1^\eta(v)$,

$$\begin{aligned}|\{H_{d+1}(t_v(s) + v) - H_{d+1}(v)\} - s| \\ \leq Q(v)e^{\lambda(1-\eta)v}\left\{1 + \frac{s(d+1)}{e^{\lambda v}W(v)}\right\} =: h_v(s). \end{aligned} \tag{2.26}$$

Now substitute $s = t_v^{-1}(u)$ into (2.26) for $u \geq 0$, giving

$$\frac{W(v)}{d+1}e^{\lambda(u+v)}(1 - \gamma^\eta(v)) + H^*(v) \leq H_{d+1}(u+v) \leq \frac{W(v)}{d+1}e^{\lambda(u+v)}(1 + \gamma^\eta(v)) + H^*(v).$$

Writing $w = H_{d+1}(u+v)$ and inverting, it then follows immediately that

$$\lambda^{-1}\log\left\{\frac{(w - H^*(v))(d+1)}{W(v)(1+\gamma^\eta(v))}\right\} \leq H_{d+1}^{-1}(w) \leq \lambda^{-1}\log\left\{\frac{(w - H^*(v))(d+1)}{W(v)(1-\gamma^\eta(v))}\right\},$$

establishing the lemma. ∎

This now allows (2.23) to be rewritten in the form

$$X_v^{(0)}(t) = \int_0^{H_{d+1}(v+t) - H_{d+1}(v)} e^{-\lambda H_{d+1}^{-1}(w + H_{d+1}(v))}(Z^{(2)}(dw) - dw), \tag{2.27}$$

where $Z^{(2)}$ is a unit rate Poisson process, with respect to which both upper limit and integrand are predictable, the latter being decreasing in w and bounded between

$$\frac{W(v)(1 - \gamma^\eta(v))}{w(d+1) + W(v)e^{\lambda v}} \quad \text{and} \quad \frac{W(v)(1 + \gamma^\eta(v))}{w(d+1) + W(v)e^{\lambda v}}, \tag{2.28}$$

for all $w \geq 0$, on the event $E_1^\eta(v)$. In order to show that we can replace both the integrand and the upper limit of integration in (2.27) with simpler expressions, without making too great an error, we use Lemma 4.1 from the Appendix.

We first replace the integrand in (2.27), showing that $X_v^{(0)}$ is close to $X_v^{(1)}$, defined by

$$X_v^{(1)}(t) := \int_0^{H_{d+1}(v+t) - H_{d+1}(v)} \frac{W(v)}{w(d+1) + W(v)e^{\lambda v}}(Z^{(2)}(dw) - dw), \tag{2.29}$$

using (2.28). We set

$$v_-(\eta) := \max\big\{0, [\lambda(1-\eta)]^{-1}\log\{e^{-2}(d+1)\}\big\}.$$

Lemma 2.4. *With the above definitions, for any $\eta < \zeta(d)$ and any $v \geq v_-(\eta)$, we have*

$$\mathbb{P}\Big[e^{\lambda v/2} \sup_{t\geq 0} |X_v^{(0)}(t) - X_v^{(1)}(t)| > \{W(v)Q(v)\gamma^\eta(v)\}^{1/2} \,|\, \widetilde{H}(v)\Big] \\ \leq \; \theta_2(v) \; := \; \theta_1(v) + \tilde{\theta}_2(v),$$

where $\theta_1(v)$ is as in (2.22)*, and $\tilde{\theta}_2(v) := 2e^{-W(v)e^{\lambda\eta v}/\{2e\}}$.*

Proof. It follows from (2.27) that $X(t) = X_v^{(0)}(t) - X_v^{(1)}(t)$ is an integral of the form considered in Lemma 4.1, albeit with a random upper limit, and its corresponding function F satisfies

$$|F(u)| \;\leq\; G(u) \;:=\; \frac{\gamma^\eta(v)W(v)}{u(d+1) + W(v)e^{\lambda v}}, \quad \text{for all } u \geq 0, \tag{2.30}$$

on $E_1^\eta(v)$, in view of (2.28). We can thus apply Lemma 4.1 to the process $\widetilde{X}$ with $\widetilde{F}(t) := F(t)\mathbf{1}\{|F(u)| \leq G(u),\, 0 \leq u < t\}$ and with $\widetilde{G}(u) := G(u)$ as in (2.30), noting that then, recalling (2.22),

$$\mathbb{P}[X(t) = \widetilde{X}(t) \text{ for all } t \geq 0 \,|\, \widehat{\mathcal{F}}_v] \;\geq\; \mathbb{P}[E_1^\eta(v) \,|\, \widetilde{H}(v)] \;\geq\; 1 - \theta_1(v).$$

Now, from (2.30), we have $\widetilde{G}_2(0,\infty) = \{\gamma^\eta(v)\}^2\{W(v)/(d+1)\}e^{-\lambda v}$. We can then choose $a := e^{-\lambda v/2}\{W(v)Q(v)\gamma^\eta(v)\}^{1/2}$ in Lemma 4.1, because

$$a \;\leq\; e\widetilde{G}_2(0,\infty)/\widetilde{G}^*(0,\infty) \;=\; e\gamma^\eta(v)\{W(v)/(d+1)\}$$

if $v \geq v_-(\eta)$, and the result follows. ∎

The next step is to simplify the upper limit in (2.29), using Lemma 4.1 to show that, with $t_v(s)$ as defined in (2.25), $(X_v^{(1)}(t_v(s)),\, s \geq 0)$ is close to the process $(X_v^{(2)}(s),\, s \geq 0)$ given by

$$X_v^{(2)}(s) \;:=\; \int_0^s \frac{W(v)}{w(d+1) + W(v)e^{\lambda v}}\,(Z^{(2)}(dw) - dw). \tag{2.31}$$

For this, we need to control $\sup_{s\geq 0,\,|z|<h_v(s)} |X_v^{(2)}(s+z) - X_v^{(2)}(s)|$, for $h_v(s)$ defined in (2.26).

Lemma 2.5. *With the definitions given in* (2.26), (2.29) *and* (2.31)*, and for any $\eta < \zeta(d)$, we have*

$$\mathbb{P}\Big[e^{\lambda v/2} \sup_{s\geq 0,\,|z|<h_v(s)} |X_v^{(2)}(s+z) - X_v^{(2)}(s)| > 4\varepsilon^\eta(v) \,\Big|\, \widetilde{H}(v)\Big] I[E_{21}^\eta(v)] \\ \leq \; \theta_3(v) \; := \; \left\{2\left[1 + \frac{W(v)}{g(v)}\right] + \frac{8e\, e^{2\lambda\eta v/3}}{Q(v)(d+1)^2}\right\} e^{-W(v)e^{\lambda\eta v/3}/\{2e(d+2)\}},$$

where $\varepsilon^\eta(v) := \{W(v)Q(v)\}^{1/2}e^{-\lambda\eta v/3}$, $g(v) := Q(v)(d+2)e^{-\lambda\eta v}$ and

$$E_{21}^\eta(v) \;:=\; \{W(v) \leq e^2(d+2)^2Q(v)e^{\lambda v/3}\} \cap \{Q(v) \leq 2e(d+1)^{-2}e^{2\lambda\eta v/3}\} \;\in\; \sigma(\widetilde{H}(v)). \tag{2.32}$$

Proof. We consider the ranges $0 \leq s \leq W(v)e^{\lambda v}$ and $s > W(v)e^{\lambda v}$ separately. In the first range of s, define $s_j := je^{\lambda v}g(v)$ for $0 \leq j \leq M := \lfloor W(v)/g(v) \rfloor$, and set $s_{M+1} := W(v)e^{\lambda v}$: then $s_{j+1} - s_j \geq h_v(s_j)$ for each j. By Lemma 4.1, with $G(u)$ the constant $e^{-\lambda v}$ and $a := e^{-\lambda v/2}\varepsilon^\eta(v)$, we have

$$\mathbb{P}\Big[\sup_{s_j\leq s\leq s_{j+1}} e^{\lambda v/2}|X_v^{(2)}(s) - X_v^{(2)}(s_j)| > \varepsilon^\eta(v) \,|\, \widetilde{H}(v)\Big] I[E_{21}^\eta(v)] \;\leq\; 2\exp\{-\varepsilon^\eta(v)^2/(2eg(v))\},$$

for $0 \le j \le M$, since $a \le eg(v) = eG(s_j)(s_{j+1} - s_j)$ on $E_{21}^{\eta}(v)$. Hence, by a standard argument,

$$\mathbb{P}\Big[\sup_{0\le s\le W(v)e^{\lambda v},\,|z|<h_v(s)} e^{\lambda v/2}|X_v^{(2)}(s+z) - X_v^{(2)}(s)| > 3\varepsilon^{\eta}(v) \,\Big|\, \widetilde{H}(v)\Big] I[E_{21}^{\eta}(v)] \\ \le\; 2\{1 + W(v)/g(v)\} \exp\{-W(v)e^{\lambda\eta v/3}/\{2e(d+2)\}\}. \tag{2.33}$$

In the second range of s, we define

$$s_j \;:=\; W(v)e^{\lambda v}(1+\tilde{g}(v))^j, \quad \text{where} \quad \tilde{g}(v) \;:=\; g(v)(d+1)/W(v),$$

noting that $s_{j+1} - s_j = s_j\tilde{g}(v) \ge h_v(s_j)$. By Lemma 4.1 with $G(u) := s_j^{-1}\{W(v)/(d+1)\}$, we have

$$\mathbb{P}\Big[\sup_{s_j\le s\le s_{j+1}} e^{\lambda v/2}|X_v^{(2)}(s) - X_v^{(2)}(s_j)| > \varepsilon^{\eta}(v) \,\Big|\, \widetilde{H}(v)\Big] I[E_{21}^{\eta}(v)] \\ \le\; 2\exp\{-\{\varepsilon^{\eta}(v)\}^2(d+1)^2(1+\tilde{g}(v))^j/(2eW(v)\tilde{g}(v))\}, \quad j \ge 0,$$

since $a := e^{-\lambda v/2}\varepsilon^{\eta}(v) \le eg(v) = e\{W(v)/(d+1)\}\tilde{g}(v) = eG(s_j)(s_{j+1} - s_j)$ on $E_{21}^{\eta}(v)$, and hence

$$\mathbb{P}\Big[\sup_{s\ge W(v)e^{\lambda v},\,|z|<h_v(s)} e^{\lambda v/2}|X_v^{(2)}(s+z) - X_v^{(2)}(s)| > 4\varepsilon^{\eta}(v) \,\Big|\, \widetilde{H}(v)\Big] I[E_{21}^{\eta}(v)] \\ \le\; 2\exp\{-W(v)(d+1)e^{\lambda\eta v/3}/\{2e(d+2)\}\} \sum_{j\ge 0} \exp\{-j\{\varepsilon^{\eta}(v)\}^2(d+1)^2/(2eW(v))\} \\ \le\; \frac{8e\, e^{2\lambda\eta v/3}}{Q(v)(d+1)^2} \exp\{-W(v)(d+1)e^{\lambda\eta v/3}/\{2e(d+2)\}\}, \tag{2.34}$$

since also $\{\varepsilon^{\eta}(v)^2\}(d+1)^2/(2eW(v)) \le 1$ on $E_{21}^{\eta}(v)$. We need $4\varepsilon^{\eta}(v)$ here as the bound on the supremum difference, rather than the usual $3\varepsilon^{\eta}(v)$, because it is possible to have $s(1-\tilde{g}(v)) < s_{j-1}$ for some $s_j < s < s_{j+1}$; however, it then has to be the case that, for such s, $s(1-\tilde{g}(v)) \ge s_{j-2}$ if $\tilde{g}(v) \le 1/2$, which is the case on $E_{21}^{\eta}(v)$. ∎

In view of Lemma 2.5 and (2.26), we immediately have the following corollary.

Corollary 2.6. *With the definitions of Lemma 2.5,*

$$\mathbb{P}\Big[e^{\lambda v/2}\sup_{s\ge 0}|X_v^{(1)}(t_v(s)) - X_v^{(2)}(s)| > 4\varepsilon^{\eta}(v) \,\Big|\, \widetilde{H}(v)\Big] I[E_{21}^{\eta}(v)] \;\le\; \theta_1(v) + \theta_3(v).$$

We now show that $X_v^{(2)}$ is close in distribution to the process $X_v^{(3)}$ defined by

$$X_v^{(3)}(s) \;:=\; \int_0^s \frac{W(v)}{w(d+1) + W(v)e^{\lambda v}}\, B(dw), \tag{2.35}$$

where, for the integrator, the compensated Poisson process $Z^{(2)}(w) - w$ from $X_v^{(2)}$ has been replaced by a standard Brownian motion $B(w)$. Note that $e^{\lambda v/2}X_v^{(3)}$ is itself just a time-changed Brownian motion:

$$\Big(\{(d+1)/W(v)\}^{1/2}\, X_v^{(3)}\big(\{W(v)/(d+1)\}e^{\lambda v}s\big),\, s\ge 0\Big) \;=_d\; \big(B(s/(s+1)),\, s\ge 0\big), \tag{2.36}$$

and so, conditional on $W(v)$, $X_v^{(3)}(\infty) \sim \mathcal{N}(0, W(v)/(d+1))$.

Lemma 2.7. *Fix $r \ge 1$. Then there are constants c_{r1} and c_{r2}, depending only on d, with the following properties. For all v such that $\lambda v \ge c_{r1}$, it is possible to construct $X_v^{(2)}$ and $X_v^{(3)}$ on the same probability space, in such a way that*

$$\mathbb{P}\Big[e^{\lambda v/2}\sup_{s\ge 0}|X_v^{(3)}(s) - X_v^{(2)}(s)| \ge c_{r2}(1+W(v))\lambda v e^{-\lambda v/2}\Big] \;\le\; \theta_4(v) \;:=\; e^{-3r\lambda v}.$$

Proof. For any $r \geq 1$, there are constants C_r, K_r with the property that, for any $n \geq 1$, a standard Poisson process Z and a standard Brownian motion B can be constructed on the same probability space in such a way that $\mathbb{P}[A_r^c(n)] \leq K_r n^{-(r+1)}$, where

$$A_r(n) \;:=\; \left\{ \sup_{0\leq s\leq n} \frac{|Z(s)-s-B(s)|}{\log n} \leq C_r \right\}.$$

This follows from Komlós, Major & Tusnády [10, Theorem 1 (ii)], together with elementary exponential bounds for the fluctuations of the standard Poisson process and Brownian motion over the time interval $[0,1]$. Fix r, and take $n := e^{3\lambda v}$ for $v \geq v_1$, where v_1 is chosen so that $e^{3\lambda v_1} \geq 2K_r$, implying that $\mathbb{P}[A_r^c(n)] \leq \frac{1}{2}e^{-3r\lambda v}$. Then use the corresponding choices of Z and B to realize $X_v^{(2)}$ and $X_v^{(3)}$, which we express, by partial integration, in the form

$$\begin{aligned} X_v^{(2)}(s) &= W(v)\left\{ \frac{Z(s)-s}{s(d+1)+W(v)e^{\lambda v}} + \int_0^s \frac{Z(u)-u}{(u(d+1)+W(v)e^{\lambda v})^2}\,du \right\}, \\ X_v^{(3)}(s) &= W(v)\left\{ \frac{B(s)}{s(d+1)+W(v)e^{\lambda v}} + \int_0^s \frac{B(u)}{(u(d+1)+W(v)e^{\lambda v})^2}\,du \right\}. \end{aligned} \quad (2.37)$$

Taking the difference, it is immediate that, for $0 \leq s \leq e^{3\lambda v}$ and on $A_r(e^{3\lambda v})$,

$$\frac{|Z(s)-s-B(s)|}{s(d+1)+W(v)e^{\lambda v}} \;\leq\; C_r\,\frac{3\lambda v}{W(v)e^{\lambda v}}$$

and that

$$\int_0^s \frac{|Z(u)-u-B(u)|}{(u(d+1)+W(v)e^{\lambda v})^2}\,du \;\leq\; C_r\,\frac{3\lambda v}{W(v)(d+1)e^{\lambda v}}\,.$$

This shows that, on $A_r(e^{3\lambda v})$,

$$e^{\lambda v/2}|X_v^{(3)}(s) - X_v^{(2)}(s)| \;\leq\; 6C_r\lambda v e^{-\lambda v/2} \quad \text{for} \quad 0 \leq s \leq e^{3\lambda v}.$$

Then, taking $F(u) = W(v)/\{u(d+1)+W(v)e^{\lambda v}\}$, $a = eC_r\{W(v)/(d+1)\}\lambda v e^{-\lambda v}$, $t_1 = e^{3\lambda v}$ and $t_2 = \infty$ in Lemma 4.1, with the choice of a permissible for all $v \geq v_2$, where $v_2 \geq \lambda^{-1}$ is chosen such that $\lambda v_2 e^{-\lambda v_2} \leq 1/C_r$, we have

$$\mathbb{P}\Big[\sup_{e^{3\lambda v}\leq s<\infty} |X_v^{(2)}(s) - X_v^{(2)}(e^{3\lambda v})| > eC_r\{W(v)/(d+1)\}\lambda v e^{-\lambda v}\Big] \leq 2\exp\{-(e/2)(C_r\lambda v)^2 e^{\lambda v}\}.$$

The same bound is satisfied also for $\sup_{e^{3\lambda v}\leq s<\infty}|X_v^{(3)}(s) - X_v^{(3)}(e^{3\lambda v})|$, as can be deduced from the representation (2.36). Now choose $v_3 \geq \lambda^{-1}$ so that $8\exp\{-(e/2)(C_r\lambda v_3)^2 e^{\lambda v_3}\} \leq e^{-3r\lambda v}$, and set $v_0 := \max\{v_1, v_2, v_3\}$. ∎

Summarizing the conclusions Lemmas 2.4 and 2.7 and of Corollary 2.6, we have the following theorem. In the error terms, $\theta_1(v)$ is defined in (2.22), $\theta_2(v)$ in Lemma 2.4, $\theta_3(v)$ in Lemma 2.5 and $\theta_4(v)$ in Lemma 2.7.

Theorem 2.8. *With the definitions* (2.12), (2.25) *and* (2.35)*, fixing any $\eta < \zeta(d)$, we can construct W and a time changed Brownian motion $X_v^{(3)}$ on the same probability space, in such a way that, for all $v \geq \lambda^{-1}c_{1*}$,*

$$\mathbb{P}\Big[e^{\lambda v/2}\sup_{u\geq 0}|\{W(u+v)-W(v)\} - X_v^{(3)}(t_v^{-1}(u))| > K(v)e^{-\lambda\eta v/3}\,\Big|\,\widetilde{H}(v)\Big] I[E_{21}^{\eta}(v)] \;\leq\; \sum_{i=1}^{4}\theta_i(v),$$

where $K(v) := 4\{W(v)Q(v)\}^{1/2} + Q(v)\sqrt{d+1} + c_{2*}(1 + W(v)e^{-\lambda v/3})$, $E_{21}^{\eta}(v) \in \sigma(\widetilde{H}(v))$ is as defined in (2.32)*, and the constants c_{1*} and c_{2*}, which depend only on d, can be deduced from Lemma 2.7 with $r = 1$.*

2.3 Consequences for the gossip process

Theorem 2.8 is not yet in a form easily applied to the gossip process. To start with, the statement of the theorem involves the $\sigma(\widetilde{H}(v))$-measurable random variables $W(v)$, $Q(v)$, $K(v)$ and $\theta_i(v)$, $1 \le i \le 4$, and it is useful to have some idea of their magnitude. It is also useful to specify how big the probability $\mathbb{P}[E^\eta_{21}(v)]$ may be. To derive appropriate statements, we begin with the random elements $W(v)$ and $W_l(v)$, $1 \le l \le d$.

Lemma 2.9. *For any* $0 < \eta < \zeta(d)$*, we have*

$$\mathbb{P}[e^{-\lambda(1-r_l-\eta)v}\,|W_l(v)| > 2] \;\le\; \begin{cases} e^{-2\lambda v(1-r_l-\eta)}(2r_l-1)^{-1} & \text{if } r_l > 1/2; \\ \lambda v\, e^{-\lambda v(1-2\eta)} & \text{if } r_l = 1/2; \\ e^{-\lambda v(1-2\eta)}(1-2r_l)^{-1} & \text{if } r_l < 1/2, \end{cases} \tag{2.38}$$

for $1 \le l \le d$*. Furthermore, for any* $s > 0$*,*

$$\mathbb{P}[W(v) \ge 1+s] \;\le\; s^{-2} \quad\text{and}\quad \mathbb{P}[W(v) \le s] \;\le\; \exp\left\{-\frac{\{\log_+(w_0/s)\}^{d+1}}{2\,(d+1)!}\right\}, \tag{2.39}$$

for a suitably chosen $w_0 > 0$*.*

Proof. The first part follows from (2.17) and Chebyshev's inequality, and, for $W(v)$, the bound on the upper tail holds because $\mathrm{Var}\, W(v) \le \mathrm{Var}\, W(\infty) \le 1$ and $\mathbb{E}W(v) = 1$. For the lower tail, note that $W(\infty) > 0$ a.s., so that, because $W(\cdot)$ is càdlàg and positive on $\mathbb{R}_+$, we have $W_* := \inf_{t>0} W(t) > 0$ a.s. also. Suppose that $w_0 > 0$ is chosen so that $\mathbb{P}[W_* \ge w_0] \ge 1/2$. Then, for $0 < x \le w_0$, $W(t) > x$ if any of the offspring of the initial individual that are born before time t_x generate families with $W_* > w_0$, where $e^{-\lambda t_x} = x/w_0$. The probability that there are no such offspring is just $\exp\{-\rho\nu t_x^{d+1}/\{2(d+1)\}\}$. Hence, for $t \ge t_x$ and $x \le w_0$,

$$\mathbb{P}[W(t) \le x] \;\le\; \exp\left\{-\frac{\rho\nu\{\log(w_0/x)\}^{d+1}}{2\lambda^{(d+1)}(d+1)}\right\} \;=\; \exp\left\{-\frac{\{\log(w_0/x)\}^{d+1}}{2\,(d+1)!}\right\}.$$

∎

In view of (2.20), if $0 < \eta < \zeta(d)$, then $Q(v) \le 3(d+1)$ on the event

$$E^\eta_{22}(v) \;:=\; \bigcap_{l=1}^{d}\{e^{-\lambda(1-r_l-\eta)v}\,|W_l(v)| \le 2\}, \tag{2.40}$$

and the first part (2.38) of Lemma 2.9 directly implies that

$$\mathbb{P}[\{E^\eta_{22}(v)\}^c] \;\le\; c(d)(1+\lambda v\mathbf{1}_{\{d=6\}})\,e^{-2\lambda v(\zeta(d)-\eta)}, \tag{2.41}$$

for a suitable constant $c(d)$; of course, by definition, $Q(v) \ge d+2$. The second part of Lemma 2.9 implies that $E_{23}(v) := \{W(v) \le 1 + e^{\lambda\eta v/3}\}$ is such that $\mathbb{P}[\{E_{23}(v)\}^c] \le e^{-2\lambda\eta v/3}$. From these observations and (2.32), it follows that

$$E^\eta_{22}(v) \cap E_{23}(v) \;\subset\; E^\eta_{21}(v),$$

if v is such that $e^{2\lambda\eta v/3} \ge (d+1)^3$, and hence, for such v,

$$\mathbb{P}[\{E^\eta_{21}(v)\}^c] \;\le\; c(d)(1+\lambda v)e^{-2\lambda v(\zeta(d)-\eta)} + e^{-2\lambda\eta v/3}; \tag{2.42}$$

in addition,

$$K(v)e^{-\lambda\eta v/3} \;\le\; \sqrt{d+1}\,\{4\sqrt{2}+2(d+1)+3c_{2*}\}\,e^{-\lambda\eta v/6}$$

on $E^\eta_{22}(v) \cap E_{23}(v)$ also.

For the quantities θ_i, $1 \le i \le 4$, note that, from (2.22),

$$\theta_1(v) \ \le\ C(d,\eta)e^{-\lambda v(\zeta(d)-\eta)} \quad \text{on the event } E^\eta_{24}(v) \ :=\ \{W(v) \le 1 + e^{\lambda v(\zeta(d)-\eta)}\}, \quad (2.43)$$

and that $\mathbb{P}[\{E^\eta_{24}(v)\}^c] \le e^{-2\lambda v(\zeta(d)-\eta)}$. Then, as in Lemma 2.7, $\theta_4(v) = e^{-3\lambda v}$ if we take $r = 1$. From Lemma 2.4, $\theta_2(v) = \theta_1(v) + \tilde\theta_2(v)$, and both $\tilde\theta_2(v)$ and $\theta_3(v)$, defined in Lemma 2.5, are super-exponentially small in $\lambda\eta v$ on the event $E^\eta_{25}(v) := \{W(v) \ge e^{-\lambda\eta v/6}\}$. Finally, by the last inequality in Lemma 2.9,

$$\mathbb{P}[\{E^\eta_{25}(v)\}^c] \ \le\ \exp\left\{-(1/2)\{\log(w_0) + \lambda\eta v/6\}^{d+1}/(d+1)!\right\},$$

which is also super-exponentially small in $\lambda\eta v$. Hence, taking

$$E^\eta(v) \ :=\ E^\eta_{22}(v) \cap E_{23}(v) \cap E^\eta_{24}(v) \cap E^\eta_{25}(v), \quad (2.44)$$

for which $\mathbb{P}[\{E^\eta(v)\}^c] \le C(d)(\lambda v e^{-2\lambda v(\zeta(d)-\eta)} + e^{-2\lambda\eta v/3})$, and assuming that v is such that $e^{2\lambda\eta v/3} \ge (d+1)^3$, we have the following consequence of Theorem 2.8. To state it, and for future use, we define

$$t_{\max}(\Lambda) \ :=\ (3/2)\lambda^{-1}\log\Lambda, \quad (2.45)$$

an upper bound for the times to be considered in proving the central limit theorem.

Corollary 2.10. *For any $0 < \eta < \zeta(d)$ and $v \le t_{\max}(\Lambda)$ such that $e^{2\lambda\eta v/3} \ge (d+1)^3$ and $\lambda v \ge c_{1*}$, there are constants $C = C(d,\eta)$ and $C' = C'(d)$ and an event $E^\eta(v) \in \sigma(\widetilde H(v))$, with $\mathbb{P}[\{E^\eta(v)\}^c] \le C'\lambda v e^{-2\lambda v(\zeta(d)-\eta)} + e^{-2\lambda\eta v/3})$, such that, for any $u \ge 0$ such that $t_\Lambda(u) \le t_{\max}(\Lambda)$,*

$$\begin{aligned} &|\mathbb{E}\{f(e^{\lambda v/2}\{W(u+v) - W(v)\}) \,|\, \widehat{\mathcal F}_v\} - \mathbb{E}\{f(e^{\lambda v/2}X^{(3)}_v(t_v^{-1}(u))) \,|\, \widehat{\mathcal F}_v\}| I[E^\eta(v)] \\ &\quad \le\ C\{e^{-\lambda\eta v/6} + e^{-\lambda v(\zeta(d)-\eta)}\}, \end{aligned} \quad (2.46)$$

uniformly for all $f \in F_{\rm BW}$.

Taking any $c_0,\ldots,c_d \in \mathbb{R}_+$ and setting $C(x) := \sum_{l=0}^d c_l x^l$, we also observe from Lemma 2.1 that

$$\Big|\sum_{l=0}^d c_l e^{-\lambda s}H_l(s) - (d+1)^{-1}C(1)W(s)\Big| \ \le\ \frac{C(1)}{d+1}\sum_{l=1}^d e^{-\lambda(1-r_l)s}|W_l(s)| \ \le\ C(1)e^{-\lambda\eta s} \quad (2.47)$$

on $E^\eta_{22}(s)$, the probability of whose complement is bounded in (2.41).

3 The central limit theorem

In this section, the central limit theorem is proved much as outlined in the introduction. With $\sigma^2_L(v,u) := \mathrm{Var}\,\{L_{t_\Lambda(u)}/L \,|\, \mathcal F_v\}$, we show in Lemma 3.2 that

$$\mathbb{E}\left\{\big|(L_{t_\Lambda(u)}/L) - \mathbb{E}\{L_{t_\Lambda(u)}/L \,|\, \mathcal F_s\}\big| \,\Big|\, \mathcal F_v\right\} \ \ll\ \sigma_L(v,u), \quad (3.1)$$

if s is chosen to be sufficiently long after v. The approximation of $\mathbb{E}\{L_{t_\Lambda(u)}/L \,|\, \mathcal F_s\}$ as a Poisson probability is then accomplished in Lemma 3.4, with an error that is small if $t_\Lambda(u) - s$ is sufficiently large. Lemmas 3.5–3.7 approximate the mean of the Poisson distribution by successively simpler quantities, and bound the errors involved in the approximations. The combined result of these steps is summarized in Corollary 3.8, showing that, given $\mathcal F_v$, the distribution of $L_{t_\Lambda(u)}/L$ is close to that of $\ell(\log[\hat c_d W(s,v)] + u)$.

Now the normalized difference $e^{\lambda v/2}(W(s,v) - W(v,v))$ can be shown, using Corollary 2.10, to have a normal approximation. Because of the normalization, it is important at this point to check that the approximation errors in the previous steps are all much smaller than $e^{-\lambda v/2}$; this places some restrictions on how large v may be. The linearization of the difference $\ell(\log[\hat{c}_d W(s,v)] + u) - \ell(\log[\hat{c}_d W(v,v)] + u)$, needed to show that it is itself approximately normally distributed, is accomplished in Lemma 3.9, and the final result is given in Theorem 3.10.

3.1 Comparisons of processes

The detailed calculations make heavy use of comparisons between a number of processes, that we justify in Lemma 3.1 by realizing them on the same probability spaces. The process $\mathcal{L}$ itself can be realized by starting with the times $(\bar{\tau}_j,\, j \geq 0)$ of the branching process $\overline{X}$, paired with a sequence of independent uniform points $(\overline{P}_j,\, j \geq 0)$ of $\mathcal{C}$. This yields a process

$$Y(t) \;:=\; \{(\bar{\tau}_j, \overline{P}_j),\, j \in \overline{J}_t\}, \quad t \geq 0, \tag{3.2}$$

in terms of which we define

$$\overline{J}_t \;:=\; \{j \geq 0 \colon \bar{\tau}_j \leq t\}; \quad \overline{N}_t \;:=\; |\overline{J}_t|; \quad \overline{M}_t \;:=\; \sum_{j \in \overline{J}_t} (t - \bar{\tau}_j)^d. \tag{3.3}$$

We can then define the set valued process

$$\overline{\mathcal{L}}(t) \;:=\; \bigcup_{j \in \overline{J}_t} \mathcal{K}(\overline{P}_j, t - \bar{\tau}_j), \tag{3.4}$$

obtained by taking the unions of the neighbourhoods generated by $Y(t)$. The process Y can be augmented to a process $\widetilde{Y}$ of quadruples, by including a set of pairs $(K(j), \overline{Q}_j)$, $j \geq 0$, where $0 \leq K(j) < j$ and $\overline{Q}_j \in \mathcal{C}$, denoting the subsets from which the long range contacts were made and the positions of the individuals within them: given $Y(\bar{\tau}_j-)$,

$$\mathbb{P}[K(j) = l] \;=\; \frac{\nu_{\bar{\tau}_j - \bar{\tau}_l}}{\sum_{l'=0}^{j} \nu_{\bar{\tau}_j - \bar{\tau}_{l'}}}, \quad 0 \leq l < j,$$

and $\overline{Q}_j$ is then chosen uniformly from the set $\mathcal{K}(\overline{P}_{K(j)}, \bar{\tau}_j - \bar{\tau}_{K(j)})$. The process $\mathcal{L}$ is derived from $\widetilde{Y}$ sequentially, by thinning. The pair $(\bar{\tau}_j, \overline{P}_j)$ is not included in $\mathcal{L}$ unless $K(j) = \min\{l \geq 0 \colon \overline{Q}_j \in \mathcal{K}(\overline{P}_l, \bar{\tau}_j - \bar{\tau}_l)\}$. This thinning process ensures that, when neighbourhoods overlap in $\mathcal{C}$, only contacts from the neighbourhood that was informed earliest are allowed, ensuring that the rate of long range transmissions from $\mathcal{L}_t$ remains equal to ρL_t. Note that, if $\overline{P}_j \in \mathcal{L}_{\bar{\tau}_j-}$, the pair $(\bar{\tau}_j, \overline{P}_j)$ is included in defining $\mathcal{L}$; however, it is redundant in (1.3), the newly informed individual having previously been informed, and it never contributes to further transmission, because of the definition of the thinning step. The resulting set of times and positions we denote by $((\tau_j, P_j),\, j \geq 0)$, with

$$J_s \;:=\; \{j \geq 0 \colon \tau_j \leq s\}; \quad N_s \;:=\; |J_s|; \quad M_s \;:=\; \sum_{j \in J_s} (s - \tau_j)^d, \tag{3.5}$$

and $\mathcal{L}$ is as given by (1.3); it satisfies $\mathcal{L}_t \subset \overline{\mathcal{L}}_t$, with strict inclusion for all large enough times.

The process $\overline{\mathcal{L}}$ acts as a tractable upper bound for $\mathcal{L}$, and it is useful also to have tractable lower bounds. In particular, when calculating the probability that a neighbourhood $\mathcal{K}(P,s)$ intersects $\mathcal{L}_t$, where s is fixed and P is a uniform random point of $\mathcal{C}$, the way in which the neighbourhoods of $\mathcal{L}_t$ intersect one another enters in a complicated way. However, if $\mathcal{L}_t$ happened to consist of a union of *non-intersecting* neighbourhoods, which were also separated from one another by distance at least $2s$, then the probability could be deduced by simply adding the intersection probabilities for the individual neighbourhoods. Then, because the neighbourhoods $\mathcal{K}$ are balls in a geodesic metric space, the probability of two neighbourhoods $\mathcal{K}(P,s)$ and $\mathcal{K}(Q,t)$ intersecting, if one or both of P and Q are chosen uniformly and independently in $\mathcal{C}$, is given by

$$q_L(s,t) \;=\; L^{-1}\nu_{s+t}, \tag{3.6}$$

where ν_{s+t} can be estimated in terms of $\nu(s+t)^d$, in view of (1.4). Of course, as t grows, intersections occur in $\mathcal{L}_t$, but, at least for a while, their effect may not be too large. So the next step is to construct subsets of $\mathcal{L}_t$ with the necessary separation properties, and which are amenable to analysis.

Fix any $s,t > 0$, and thin the process $\widetilde{Y}$ to obtain a set valued process $\mathcal{L}^{s,t}$ as follows. Start with $\tau_0^{s,t} = 0$ and $P_0^{s,t} = P_0$, defining

$$\mathcal{L}_u^{s,t} \;:=\; \mathcal{K}(P_0,u) \quad \text{for } 0 \le u < \bar\tau_1;$$

let $R_0^{s,t} := \emptyset$ denote the initial set of indices of censored points of $\widetilde{Y}$. Then proceed sequentially. Suppose that the quadruples $((\bar\tau_l, \overline{P}_l, K(l), \overline{Q}_l),\ 0 \le l \le j-1) \subset \widetilde{Y}$ have already been considered. If $K(j) \in R_{j-1}^{s,t}$, set $R_j^{s,t} := R_{j-1}^{s,t} \cup \{j\}$ and proceed to the next quadruple; descendants of censored points are also censored. If not, thin much as in the construction of $\mathcal{L}$, except that a point $\overline{P}_j$ is also thinned if it belongs to $N_{2s+t-\bar\tau_j}(\mathcal{L}_{\bar\tau_j-}^{s,t})$, where, for $V \subset \mathcal{C}$ and $u > 0$,

$$N_u(V) \;:=\; \bigcup_{y\in V} \mathcal{K}(y,u); \tag{3.7}$$

set

$$\mathcal{L}_u^{s,t} \;:=\; \bigcup_{l=0}^{j} \mathbf{1}_{\{l \notin R_j^{s,t}\}} \mathcal{K}(\overline{P}_l, u-\bar\tau_l), \quad \bar\tau_j \le u < \bar\tau_{j+1}. \tag{3.8}$$

The extra thinning in (3.7) ensures that the neighbourhoods in $\mathcal{L}_t^{s,t}$ are at distance at least $2s$ from one another. If $J_u^{s,t}$ denotes the set of indices of the points of $\widetilde{Y}$ that enter $\mathcal{L}^{s,t}$ up to time u, then $\mathcal{L}_u^{s,t}$ consists of the collection of disjoint neighbourhoods $(\mathcal{K}(\overline{P}_j, u-\bar\tau_j),\ j \in J_u^{s,t})$, and new points are generated at rate $\rho\sum_{j\in J_u^{s,t}} \nu_{u-\bar\tau_j}(1-\pi_u^{s,t})$, where the censoring probability $\pi_u^{s,t}$ is given by

$$\pi_u^{s,t} \;:=\; L^{-1} \sum_{j\in J_u^{s,t}} \nu_{2s+(t-\bar\tau_j)+(t-u)}. \tag{3.9}$$

In our applications, we can find suitably small bounds for $\pi_u^{s,t}$, so that the growth of the numbers of neighbourhoods in $\mathcal{L}^{s,t}$ is still reasonably close to that of the CMJ process $\overline{X}$. In view of the 'hard core' censoring, the points $(\overline{P}_j,\ j \in J_u^{s,t})$ are no longer independent of one another, but their marginal distribution is still uniform on $\mathcal{C}$ if P_0 is chosen at random. Note also that $\mathcal{L}_u^{s,t} \subset \mathcal{L}_u$ for each $s,t \ge 0$ and $0 < u \le t$.

We shall also use comparisons between the CMJ process $\overline{X}$ and 'flattened' versions $\widehat{X}_-$, $\widehat{X}_0$ and $\widehat{X}_+$ that are of the form discussed in the previous section. We start by noting that, from the inequality (1.4),

$$\nu s^d\{1-\eta_\Lambda\} \;\le\; \nu_s \;\le\; \nu s^d\{1+\eta_\Lambda\}, \quad 0 < s \le t_{\max}(\Lambda), \tag{3.10}$$

where $t_{\max}(\Lambda) := \frac{3}{2\lambda}\log\Lambda$ is as in (2.45), and

$$\eta_\Lambda \;:=\; c_g\left(\frac{3\log\Lambda}{2\Lambda^{1/d}}\right)^{\gamma_g}. \tag{3.11}$$

Hence, up to time $t_{\max}(\Lambda)$, the process $\overline{X}$ is stochastically dominated by the flattened process $\widehat{X}_+$, defined as in the previous section, having intensity $\rho_+ := \rho(1+\eta_\Lambda)$ per unit volume, and hence growth rate $\lambda_+ := \lambda\{1+\eta_\Lambda\}^{1/d}$; similarly, it stochastically dominates the flattened process $\widehat{X}_-$ with $\rho_- := \rho(1-\eta_\Lambda)$ and $\lambda_- := \lambda\{1-\eta_\Lambda\}^{1/d}$. We also define the flattened process $\widehat{X}_0$ with intensity ρ per unit volume, and with growth rate λ. The quantities M_j^+, M_j^0 and M_j^-, and their standardized versions H_j^+, H_j^0 and H_j^-, correspond to these processes. We make the relationships between the processes precise with the following construction.

Lemma 3.1. *Let the successive birth times in the branching processes $\overline{X}$, $\widehat{X}_-$, $\widehat{X}_0$ and $\widehat{X}_+$ be denoted by $(\bar{\tau}_j, \hat{\tau}_j^-, \hat{\tau}_j^0, \hat{\tau}_j^+, j \ge 0)$, respectively, and let $(T_t, T_t^-, T_t^0, T_t^+)$ denote the sets of birth times up to time t in each of the processes. If, for some $0 \le s < t_{\max}(\Lambda)$, $T_s^- \subset T_s \subset T_s^+$ and $T_s^- \subset T_s^0 \subset T_s^+$, then the processes $\overline{X}$, $\widehat{X}_-$, $\widehat{X}_0$ and $\widehat{X}_+$ can be defined on the same probability space, in such a way that, for all $s \le t \le t_{\max}(\Lambda)$,*

$$T_t^- \subset T_t \subset T_t^+ \quad \textit{and} \quad T_t^- \subset T_t^0 \subset T_t^+ \quad \textit{a.s.}$$

Proof. The birth rate of $\overline{X}$ at time t is given by

$$r(\overline{X}, t) \;:=\; \rho \sum_{j\,:\,\tau_j \in T_t} \nu_{t-\bar{\tau}_j},$$

and of $\widehat{X}_0$ by

$$r(\widehat{X}_0, t) \;:=\; \lambda H_d^0(t) \;=\; \lambda^{d+1} \sum_{j\,:\,\hat{\tau}_j^0 \in T_t^0} (t-\hat{\tau}_j^0)^d/d! \;=\; \rho\nu \sum_{j\,:\,\hat{\tau}_j^0 \in T_t^0} (t-\hat{\tau}_j^0)^d,$$

with analogous representations for $r(\widehat{X}_-, t)$ and $r(\widehat{X}_+, t)$. Thus, for any time t such that

$$T_t^- \subset T_t \subset T_t^+ \quad \text{and} \quad T_t^- \subset T_t^0 \subset T_t^+, \tag{3.12}$$

we have $r(\widehat{X}_-, t) \le r(\overline{X}, t) \le r(\widehat{X}_+, t)$ and $r(\widehat{X}_-, t) \le r(\widehat{X}_0, t) \le r(\widehat{X}_+, t)$. Hence, for s as given, we can construct all four processes on the same probability space, for $s \le t \le t_{\max}(\Lambda)$, by realizing $\widehat{X}_+$ on $[s, t_{\max}(\Lambda)]$ together with an independent sequence of independent random variables $(U_j,\, j \ge 1)$ uniformly distributed on $[0,1]$, and then thinning in the following way. At each successive point $\hat{\tau}_j^+ > s$, include it as a point of $\overline{X}$ if $U_j r(\widehat{X}_+, t) \le r(\overline{X}, t)$; similarly, if $U_j r(\widehat{X}_+, t) \le r(\widehat{X}_-, t)$, include $\hat{\tau}_j^+$ as a point of $\widehat{X}_-$, and if $U_j r(\widehat{X}_+, t) \le r(\widehat{X}_0, t)$, include $\hat{\tau}_j^+$ as a point of $\widehat{X}_0$. This construction preserves the inclusions (3.12) for all times up to $t_{\max}(\Lambda)$, and, because independently thinned Poisson processes are again Poisson processes, also yields the right distributions for the processes $\overline{X}$, $\widehat{X}_0$ and $\widehat{X}_-$. ∎

In what follows, we shall use $\mathcal{F}_t^{++}$ to denote the filtration for the combined construction in Lemma 3.1. We shall henceforth only consider times in $[0, t_{\max}(\Lambda)]$, and will take Λ large enough that

$$\exp\{3\eta_\Lambda t_{\max}(\Lambda)\} \;\le\; 2 \quad \text{and} \quad \eta_\Lambda \;\le\; 1. \tag{3.13}$$

3.2 Relating the proportion informed to the function ℓ

The first step in our detailed calculations is to replace L_t/L with $\mathbb{E}\{L_t/L \,|\, \widetilde{\mathcal{F}}_s\}$, where $\widetilde{\mathcal{F}}_s := \sigma(\widetilde{Y}_u,\, 0 \le u \le s)$, for suitable $s < t$; this conditional expectation is easier to handle. We start by bounding the conditional variance $\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_s\}$, for suitable values of $s < t$.

The basis for our argument is given by the observations that

$$\mathbb{E}\{1 - L_t/L \,|\, \widetilde{\mathcal{F}}_s\} \;=\; \mathbb{P}[K \notin \mathcal{L}_t \,|\, \widetilde{\mathcal{F}}_s] \quad \text{and} \quad \mathbb{E}\{(1-L_t/L)^2 \,|\, \widetilde{\mathcal{F}}_s\} \;=\; \mathbb{P}[K, K' \notin \mathcal{L}_t \,|\, \widetilde{\mathcal{F}}_s], \tag{3.14}$$

where K and K' are chosen independently and uniformly in $\mathcal{C}$, implying that

$$\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_s\} \;=\; \mathbb{P}[K, K' \notin \mathcal{L}_t \,|\, \widetilde{\mathcal{F}}_s] - \{\mathbb{P}[K \notin \mathcal{L}_t \,|\, \widetilde{\mathcal{F}}_s]\}^2. \tag{3.15}$$

On the other hand,

$$\{K \notin \mathcal{L}_t\} \;=\; \{\widetilde{\mathcal{L}}^K_{t,s} \cap \mathcal{L}_s = \emptyset\}, \tag{3.16}$$

where $\widetilde{\mathcal{L}}^K_{t,s}$ denotes the set of all points at time s that, if informed, would inform K by time t. Now, for the gossip process, $\widetilde{\mathcal{L}}^K_{t,s}$ is *independent* of $\widetilde{\mathcal{F}}_s$, and has the same distribution as $\mathcal{L}_{t-s}$. In view of (3.16), we thus have

$$\mathbb{P}[K \notin \mathcal{L}_t \,|\, \widetilde{\mathcal{F}}_s] \;=\; \mathbb{P}[\widetilde{\mathcal{L}}^K_{t,s} \cap \mathcal{L}_s = \emptyset \,|\, \widetilde{\mathcal{F}}_s], \tag{3.17}$$

where $\mathcal{L}_s$ is $\widetilde{\mathcal{F}}_s$-measurable and $\widetilde{\mathcal{L}}^K_{t,s}$ is independent of $\widetilde{\mathcal{F}}_s$, and

$$\mathbb{P}[K, K' \notin \mathcal{L}_t \,|\, \widetilde{\mathcal{F}}_s] \;=\; \mathbb{P}[\{\widetilde{\mathcal{L}}^K_{t,s} \cap \mathcal{L}_s = \emptyset\} \cap \{\widetilde{\mathcal{L}}^{K'}_{t,s} \cap \mathcal{L}_s = \emptyset\} \,|\, \widetilde{\mathcal{F}}_s], \tag{3.18}$$

with $\widetilde{\mathcal{L}}^K_{t,s}$ and $\widetilde{\mathcal{L}}^{K'}_{t,s}$ independent of $\widetilde{\mathcal{F}}_s$, but not of each other. Indeed, in view of (3.15), it is the extent of their dependence that measures $\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_s\}$.

Writing $t_s := t - s$, our argument now involves bounding the differences

$$\mathbb{P}[\widetilde{\mathcal{L}}^K_{t,s} \cap \mathcal{L}_s = \emptyset \,|\, \widetilde{\mathcal{F}}_s] - \mathbb{P}[\overline{\mathcal{L}}^K(t_s) \cap \mathcal{L}_s = \emptyset \,|\, \widetilde{\mathcal{F}}_s] \quad \text{and} \tag{3.19}$$

$$\mathbb{P}[\{\widetilde{\mathcal{L}}^K_{t,s} \cap \mathcal{L}_s = \emptyset\} \cap \{\widetilde{\mathcal{L}}^{K'}_{t,s} \cap \mathcal{L}_s = \emptyset\} \,|\, \widetilde{\mathcal{F}}_s] - \mathbb{P}[\{\overline{\mathcal{L}}^K(t_s) \cap \mathcal{L}_s = \emptyset\} \cap \{\overline{\mathcal{L}}^{K'}(t_s) \cap \mathcal{L}_s = \emptyset\} \,|\, \widetilde{\mathcal{F}}_s] \tag{3.20}$$

between the probabilities (3.17) and (3.18) and the smaller ones obtained by replacing $\widetilde{\mathcal{L}}^K_{t,s}$ and $\widetilde{\mathcal{L}}^{K'}_{t,s}$ by their related (independent) branching and growth processes $\overline{\mathcal{L}}^K$ and $\overline{\mathcal{L}}^{K'}$. These, as observed in the joint construction at the beginning of the section, give rise to stochastically larger sets than $\widetilde{\mathcal{L}}^K_{t,s}$ and $\widetilde{\mathcal{L}}^{K'}_{t,s}$. If both of the differences (3.19) and (3.20) are smaller than some ε, then the independence of $\overline{\mathcal{L}}^K$ and $\overline{\mathcal{L}}^{K'}$ immediately implies that $\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_s\} \le 4\varepsilon$. Using this strategy, we prove the following lemma.

Lemma 3.2. *Under the above assumptions, there is a constant $C_{3.2} = C_{3.2}(d)$ such that*

$$\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_s\} \;\le\; C_{3.2}\Lambda^{-2}(1 + (\lambda s)^d)e^{2\lambda(t-s)}(\lambda^d M_s + N_s).$$

Proof. To control the differences (3.19) and (3.20), we begin by running a process $\widetilde{Y}^K$, defined following (3.2), until time t_s, and thin to obtain $\widetilde{\mathcal{L}}^K_{t,s}$. As in (3.3), define $\overline{J}^K_u := \{j \ge 0 \colon \bar\tau^K_j \le u\}$, and set $\overline{N}^K_u := |\overline{J}^K_u|$ and $\overline{M}^K_u := \sum_{j \in \overline{J}^K_u} (u - \bar\tau^K_j)^d$. We then thin $\widetilde{Y}^K$ further to construct the process $(\mathcal{L}^{0,t_s,K}(u),\, 0 \le u \le t_s)$, by the method used to construct $\mathcal{L}^{s,t}$ in (3.8).

We now consider the difference

$$\Delta_{s,t} \;:=\; \mathbb{P}[\mathcal{L}^{0,t_s,K}(t_s) \cap \mathcal{L}_s = \emptyset \,|\, \widetilde{\mathcal{F}}_s] - \mathbb{P}[\overline{\mathcal{L}}^K_{t_s} \cap \mathcal{L}_s = \emptyset \,|\, \widetilde{\mathcal{F}}_s],$$

which is an upper bound for the real quantity (3.19) of interest to us. The quantity $\Delta_{s,t}$ is no larger than the conditional expectation given $\widetilde{\mathcal{F}}_s$ of the number $Z^K_{t,s}$ of intersections between *censored* islands of $\overline{\mathcal{L}}^K_{t_s}$ and the islands of $\mathcal{L}_s$. If an island born in $\overline{X}^K$ at u is censored, the expected number of censored islands that result at t_s is at most $c_1 e^{\lambda_+(t_s - u)}$, by (2.2) and because $\overline{X}^K$ is stochastically dominated by $\widehat{X}_+$. These islands each have radius at most $(t_s - u)$. Hence, given $\widetilde{\mathcal{F}}_s$, the expected number of intersections resulting from a censored island born at u is at most

$$\begin{aligned} & c_1 e^{\lambda_+(t_s-u)} \sum_{j\in J_s} L^{-1}\nu_{(s-\tau_j)+(t_s-u)} \\ & \quad\le\ c_1 e^{\lambda_+(t_s-u)}\nu(1+\eta_\Lambda)L^{-1}\sum_{j\in J_s}((s-\tau_j)+(t_s-u))^d \\ & \quad\le\ 2^d c_1\nu e^{\lambda_+(t_s-u)}L^{-1}(\overline{M}_s+\overline{N}_s(t_s-u)^d), \end{aligned}$$

in view of (3.6), (1.4) and (3.13); $\overline{N}$ and $\overline{M}$ are as in (3.3). Similarly, using (3.9), the conditional probability $\pi_u^{0,t_s,K}$ of an island born in $\overline{X}^K$ at u being censored for $\mathcal{L}^{0,t_s,K}$, given the history up to u, is bounded above by

$$\begin{aligned} & (1+\eta_\Lambda)\nu L^{-1}\int_{(0,u)}\{2s+(t_s-v)+(t_s-u)\}^d\overline{N}^K(dv) \\ & \quad\le\ 2.3^{d-1}\nu L^{-1}\int_{(0,u)}\{(2s)^d+(2(t_s-u))^d+(u-v)^d\}\overline{N}^K(dv) \\ & \quad=\ 2.3^{d-1}\nu L^{-1}\big\{\overline{N}^K_{u-}\{(2s)^d+(2(t_s-u))^d\}+\overline{M}^K_{u-}\big\}. \end{aligned}$$

Hence, again using $\overline{N}^K$ as an upper bound for the number of uncensored islands, and noting that the birth intensity in $\overline{X}^K$ at time u is at most

$$\rho\sum_{j\in\overline{J}_u}\nu_{u-\bar\tau_j^K}\ \le\ 2\nu\rho\overline{M}^K_u,$$

we have

$$\begin{aligned} & \mathbb{E}\{Z^K_{t,s}\,|\,\widetilde{\mathcal{F}}_s\} \\ & \quad\le\quad \mathbb{E}\left\{\int_0^{t_s} 2.3^{d-1}\nu L^{-1}\big\{\overline{N}^K_{u-}\{(2s)^d+2^d(t_s-u)^d\}+\overline{M}^K_{u-}\big\}\right. \\ & \qquad\qquad \left. 2^d c_1\nu e^{\lambda_+(t_s-u)}L^{-1}(\overline{M}_s+\overline{N}_s(t_s-u)^d)\,\overline{N}^K(du)\,\Big|\,\widetilde{\mathcal{F}}_s\right\} \\ & \quad\le\quad 2^{d+1}3^{d-1}c_1\rho\{\nu\}^3L^{-2} \qquad\qquad (3.21) \\ & \qquad \mathbb{E}\left\{\int_0^{t_s}\{\overline{N}^K_u\{(2s)^d+2^d(t_s-u)^d\}+\overline{M}^K_u\}e^{\lambda_+(t_s-u)}(\overline{M}_s+\overline{N}_s(t_s-u)^d)\overline{M}^K_u\,du\,\Big|\,\widetilde{\mathcal{F}}_s\right\}. \end{aligned}$$

Now, by (2.2), (2.4) and Cauchy–Schwarz, and because $\overline{X}^K$ is stochastically dominated by $\widehat{X}_+$,

$$\mathbb{E}\big\{(\overline{N}^K_u\{(2s)^d+(t_s-u)^d\}+\overline{M}^K_u)\overline{M}^K_u\big\}\ \le\ c_2 d!\lambda_+^{-d}\{(2s)^d+(t_s-u)^d+d!\lambda_+^{-d}\}e^{2\lambda_+u}.$$

Using this in (3.21), and noting that $\lambda_+\le\lambda(1+\eta_\Lambda)$ and that $\rho\nu d! = \lambda^{d+1}$, gives the following bound for (3.19):

$$\begin{aligned} 0\ &\le\ \mathbb{P}[\widetilde{\mathcal{L}}^K_{t,s}\cap\mathcal{L}_s=\emptyset\,|\,\widetilde{\mathcal{F}}_s]-\mathbb{P}[\overline{\mathcal{L}}^K_{t_s}\cap\mathcal{L}_s=\emptyset\,|\,\widetilde{\mathcal{F}}_s]\ \le\ \mathbb{E}\{Z^K_{t,s}\,|\,\widetilde{\mathcal{F}}_s\} \\ &\le\ C_1(d)(1+(\lambda_+s)^d)\lambda_+^{-d}\{\nu\}^2L^{-2}e^{2\lambda_+t_s}(\overline{M}_s+\lambda_+^{-d}\overline{N}_s) \qquad (3.22) \\ &\le\ C_1(d)\Lambda^{-2}(1+(\lambda_+s)^d)e^{2\lambda t_s}(\lambda^d\overline{M}_s+\overline{N}_s). \qquad (3.23) \end{aligned}$$

We now need to bound (3.20). This can be done by introducing a process $\mathcal{L}^{0,t_s,K,K'}$, constructed in the same way as $\mathcal{L}^{0,t_s,K}$, but starting from two initial points K,K' and using a CMJ process $\overline{X}^{K,K'}$, which is the same as using two independent CMJ processes $\overline{X}^{K}$ and $\overline{X}^{K'}$, by the branching property. Now $\mathcal{L}^{0,t_s,K,K'}(t_s) \subset (\widetilde{\mathcal{L}}^{K}_{t,s} \cup \widetilde{\mathcal{L}}^{K'}_{t,s})$, and the conditional expection given $\widetilde{\mathcal{F}}_s$ of the number $Z^{K,K'}_{t,s}$ of intersections between censored islands of $\overline{X}^{K,K'}_{t_s}$ and the islands of $\mathcal{L}_s$ satisfies

$$\mathbb{E}\{Z^{K,K'}_{t,s} \,|\, \widetilde{\mathcal{F}}_s\} \;\le\; C_2(d)\{\nu\}^2 L^{-2}(1+(\lambda_+ s)^d)e^{2\lambda_+ t_s}(\lambda^d \overline{M}_s + \overline{N}_s), \tag{3.24}$$

by an argument exactly as before, but for a larger constant $C_2(d)$ than $C_1(d)$ appearing in (3.23). Since $\mathbb{E}\{Z^{K,K'}_{t,s} \,|\, \widetilde{\mathcal{F}}_s\}$ is a bound for the difference in (3.20), we have enough to prove the lemma. ∎

Remark 3.3. With $s = \alpha_1\lambda^{-1}\log\Lambda$ and $t = \alpha_2\lambda^{-1}\log\Lambda$, where $\alpha_1 < \alpha_2 \le 1$, and since $\mathbb{E}(\lambda^d\overline{M}_s + \overline{N}_s) = O(e^{\lambda_+ s})$, from (2.4), it follows that $\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_s\}$ is typically of order $O\big(\Lambda^{2\alpha_2-\alpha_1-2}(\log\Lambda)^d\big)$.

Our main interest is in approximating the distribution of L_t/L when

$$t \;=\; t_\Lambda(u) \;:=\; \lambda^{-1}\{\log\Lambda + u\}, \tag{3.25}$$

for u fixed. This is because the times $(t_\Lambda(u),\, u \in \mathbb{R})$ asymptotically represent the period in which L_t/L increases from 0 to 1. Taking $\alpha_1 = \alpha < 1$ and $\alpha_2 = 1$ in the remark, it follows that $\mathrm{Var}\,\{L_{t_\Lambda(u)}/L \,|\, \widetilde{\mathcal{F}}_s\}$ is typically of order $O(\Lambda^{-\alpha})$ for $s := \alpha\lambda^{-1}\log\Lambda$. Now pick $v := \alpha_1\lambda^{-1}\log\Lambda$ and $s := \alpha_2\lambda^{-1}\log\Lambda$, with $\alpha_1 < \alpha_2 < 1$. Then

$$\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_v\} \;=\; \mathrm{Var}\,\{\mathbb{E}(L_t/L \,|\, \widetilde{\mathcal{F}}_s) \,|\, \widetilde{\mathcal{F}}_v\} + \mathbb{E}\{\mathrm{Var}\,(L_t/L \,|\, \widetilde{\mathcal{F}}_s) \,|\, \widetilde{\mathcal{F}}_v\}\,,$$

in which the latter term, again by the remark, is typically of order $O(\Lambda^{-\alpha_2})$ if $t = t_\Lambda(u)$. Supposing that $\mathrm{Var}\,\{L_t/L \,|\, \widetilde{\mathcal{F}}_v\}$ is actually of magnitude $\Lambda^{-\alpha_1}$, this indicates that the conditional distribution of L_t/L given $\widetilde{\mathcal{F}}_v$ is essentially that of the conditional distribution of $\mathbb{E}(L_t/L \,|\, \widetilde{\mathcal{F}}_s)$ given $\widetilde{\mathcal{F}}_v$. So the next step is to examine $\mathbb{E}\{(1 - L_t/L) \,|\, \widetilde{\mathcal{F}}_s\}$ in detail, for $t = t_\Lambda(u)$, and to express it in more amenable form.

The next lemma once again uses the backward branching process $\overline{\mathcal{L}}^{K}$ from a randomly chosen point K. We define $\mathcal{F}^K_{s,t} := \widetilde{\mathcal{F}}_s \bigvee \mathcal{F}^K_{t-s;0}$, where $\mathcal{F}^K_{v;0} := \sigma(\overline{N}^K_u,\, 0 \le u \le v)$ contains the information about when the islands of $\overline{\mathcal{L}}^{K}$ were formed, up to time v, but not where they are centred. We then write $Z^{s,t}$ for the number of islands of $\overline{\mathcal{L}}^{K}_{t_s}$ that intersect $\mathcal{L}_s$.

Lemma 3.4. *With the definitions above, there is a constant $C_{3.4} = C_{3.4}(d)$ such that*

$$\begin{aligned}&\big|\mathbb{E}\{(1 - L_t/L) \,|\, \widetilde{\mathcal{F}}_s\} - \mathbb{E}\big\{\exp\{-M^K_{s,t}\} \,|\, \widetilde{\mathcal{F}}_s\big\}\big| \\ &\quad\le\; C_{3.4}\{\Lambda^{-1}\overline{N}_s(\lambda t)^d + \Lambda^{-2}(1+(\lambda_+ s)^d)e^{2\lambda(t-s)}(\lambda^d\overline{M}_s + \overline{N}_s)\},\end{aligned}$$

where $M^K_{s,t} := \mathbb{E}\{Z^{s,t} \,|\, \mathcal{F}^K_{s,t}\}$.

Proof. We start by using (3.14), (3.16) and (3.23) to show that, for $t > s$,

$$|\mathbb{E}\{(1-L_t/L) \,|\, \widetilde{\mathcal{F}}_s\} - \mathbb{P}[\overline{\mathcal{L}}^{K}_{t_s} \cap \mathcal{L}_s = \emptyset \,|\, \widetilde{\mathcal{F}}_s]| \;\le\; C_1(d)\Lambda^{-2}(1+(\lambda_+ s)^d)e^{2\lambda t_s}(\lambda^d\overline{M}_s + \overline{N}_s). \tag{3.26}$$

We now use Poisson approximation to approximate the probability $\mathbb{P}[\overline{\mathcal{L}}^{K}_{t_s} \cap \mathcal{L}_s = \emptyset \,|\, \widetilde{\mathcal{F}}_s]$, using the conditional independence between the locations of the islands of $\overline{\mathcal{L}}^{K}_{t_s}$, given $\mathcal{F}^K_{s,t}$, as the basis of the approximation.

We first observe that the conditional probability that an island of $\overline{\mathcal{L}}^K_{t_s}$ with radius v intersects $\mathcal{L}_s$, given $\mathcal{F}^K_{s,t}$, is at most

$$\sum_{j\in J_s} \nu_{s-\tau_j+v} L^{-1} \ \le\ 2\overline{N}_s \nu L^{-1} t^d \ =\ (1+\eta_\Lambda)\Lambda^{-1}\overline{N}_s(\lambda t)^d, \tag{3.27}$$

in view of (3.6), by (1.4), (3.11) and (3.13), and because $v \le t-s$. This, using $Z^{s,t}$ to denote the number of islands of $\overline{\mathcal{L}}^K_{t_s}$ that intersect $\mathcal{L}_s$, implies that

$$d_{TV}(\mathcal{L}(Z^{s,t}\,|\,\mathcal{F}^K_{s,t}), \mathrm{Po}\,(M^K_{s,t})) \ \le\ 2\Lambda^{-1}\overline{N}_s(\lambda t)^d, \tag{3.28}$$

by Barbour, Holst & Janson [4, (1.23)]), where $M^K_{s,t} := \mathbb{E}\{Z^{s,t}\,|\,\mathcal{F}^K_{s,t}\}$. Hence, from (3.28),

$$\left|\mathbb{P}[Z^{s,t} = 0\,|\,\widetilde{\mathcal{F}}_s] - \mathbb{E}\{\exp(-M^K_{s,t})\,|\,\widetilde{\mathcal{F}}_s\}\right| \ \le\ 2\Lambda^{-1}\overline{N}_s(\lambda t)^d,$$

and combining this with (3.26) gives the lemma. ∎

We now define

$$\widetilde{M}^K_{s,t} \ :=\ \int_0^{t_s} \sum_{j\in J_s} \nu L^{-1}(s-\tau_j+t_s-v)^d\,\overline{N}^K(dv), \tag{3.29}$$

as an approximation to $M^K_{s,t}$. The following lemma bounds the accuracy of the approximation for $t = t_\Lambda(u)$.

Lemma 3.5. *For any $\gamma > 0$, there is an event $B_{3.5}(\gamma,s) \in \widetilde{\mathcal{F}}_s$ with $\mathbb{P}[\{B_{3.5}(\gamma,s)\}^c] \le C_{3.5}\Lambda^{-\gamma}$ such that, for $t = t_\Lambda(u)$,*

$$\mathbb{E}\{|\widetilde{M}^K_{s,t} - M^K_{s,t}|\,|\,\widetilde{\mathcal{F}}_s\} I[B_{3.5}(\gamma,s)] \ \le\ C'_{3.5}\Lambda^\gamma e^u\{\Lambda^{-1}\{\lambda s\}^d e^{\lambda s} + \eta_\Lambda\},$$

where $C_{3.5}$ and $C'_{3.5}$ depend only on d.

Proof. We begin by introducing the censored version $\overline{\mathcal{L}}^{s,s}$ of the process $\overline{\mathcal{L}}$. We denote the indices of islands in $\overline{\mathcal{L}}^{s,s}_s$ by $J^{s,s}_s \subset J_s$, and write $r_{js} := s - \bar\tau_j$. It then follows that

$$\int_0^{t_s} \sum_{j\in J^{s,s}_s} L^{-1}\nu_{r_{js}+t_s-v}\overline{N}^K(dv) \ \le\ M^K_{s,t} \ \le\ \int_0^{t_s} \sum_{j\in \overline{J}_s} L^{-1}\nu_{r_{js}+t_s-v}\overline{N}^K(dv), \tag{3.30}$$

with the lower bound using the separation between the islands of $\overline{\mathcal{L}}^{s,s}$. Now, from (3.10), (3.11) and (3.30),

$$\begin{aligned} M^K_{s,t} &\ge \int_0^{t_s} \sum_{j\in J^{s,s}_s} L^{-1}\nu_{r_{js}+t_s-v}\overline{N}^K(dv) \\ &\ge (1-\eta_\Lambda)\int_0^{t_s} \sum_{j\in J^{s,s}_s} L^{-1}\nu(r_{js}+t_s-v)^d\overline{N}^K(dv), \end{aligned}$$

and

$$\begin{aligned} M^K_{s,t} &\le \int_0^{t_s} \sum_{j\in \overline{J}_s} L^{-1}\nu_{r_{js}+t_s-v}\overline{N}^K(dv) \\ &\le (1+\eta_\Lambda)\int_0^{t_s} \sum_{j\in \overline{J}_s} L^{-1}\nu(r_{js}+t_s-v)^d\overline{N}^K(dv). \end{aligned}$$

Hence

$$M^K_{s,t} - \widetilde{M}^K_{s,t} \;\le\; \eta_\Lambda \widetilde{M}^K_{s,t} + (1+\eta_\Lambda)\int_0^{t_s} \sum_{j\in\overline{J}_s\setminus J_s} L^{-1}\nu(r_{js}+t_s-v)^d\,\overline{N}^K(dv), \tag{3.31}$$

and

$$\widetilde{M}^K_{s,t} - M^K_{s,t} \;\le\; \eta_\Lambda \widetilde{M}^K_{s,t} + (1+\eta_\Lambda)\int_0^{t_s} \sum_{j\in J_s\setminus J^{s,s}_s} L^{-1}\nu(r_{js}+t_s-v)^d\,\overline{N}^K(dv). \tag{3.32}$$

This implies that

$$\begin{aligned} |\widetilde{M}^K_{s,t} - M^K_{s,t}| \;&\le\; \eta_\Lambda \widetilde{M}^K_{s,t} + (1+\eta_\Lambda)\int_0^{t_s} \sum_{j\in\overline{J}_s\setminus J^{s,s}_s} L^{-1}\nu(r_{js}+t_s-v)^d\,\overline{N}^K(dv) \\ &\le\; \eta_\Lambda \widetilde{M}^K_{s,t} + 2^d\left(\overline{N}^K_{t_s}\sum_{j\in\overline{J}_s\setminus J^{s,s}_s} L^{-1}r^d_{js} + L^{-1}(\overline{N}_s - N^{s,s}_s)\overline{M}^K_{t_s}\right), \end{aligned} \tag{3.33}$$

where $N^{s,s}_s := |J^{s,s}_s|$. Thus we need to bound the conditional expectation given $\widetilde{\mathcal{F}}_s$ of the right hand side of (3.33).

Define $B_1(\gamma,s)$ by

$$B_1(\gamma,s) \;:=\; \left\{\lambda^d \sum_{j\in\overline{J}_s\setminus J^{s,s}_s} r^d_{js} + d!(\overline{N}_s - N^{s,s}_s) \;\le\; \Lambda^{-1+\gamma}\{\lambda_+ s\}^d e^{2\lambda_+ s}\right\} \;\in\; \widetilde{\mathcal{F}}_s. \tag{3.34}$$

Since $\overline{\mathcal{L}}^K$ is independent of $\mathcal{L}$ in (3.33), it follows that we can easily take the expectation, given $\widetilde{\mathcal{F}}_s$, of its second term. For $t = t_\Lambda(u)$, and using (2.2) and (2.4), this gives

$$\begin{aligned} &\mathbb{E}\Big\{2^d\left(\overline{N}^K_{t_s}\sum_{j\in\overline{J}_s\setminus J^{s,s}_s} L^{-1}r^d_{js} + L^{-1}(\overline{N}_s - N^{s,s}_s)\overline{M}^K_{t_s}\right)\,\Big|\,\widetilde{\mathcal{F}}_s\Big\} I[B_1(\gamma,s)] \\ &\quad\le\; 2^d c_1\Lambda^{-1}e^{\lambda_+ t_s}\left(\lambda^d\sum_{j\in\overline{J}_s\setminus J^{s,s}_s} r^d_{js} + d!(\overline{N}_s - N^{s,s}_s)\right) I[B_1(\gamma,s)] \\ &\quad\le\; 2^{d+2}c_1\Lambda^{-1+\gamma}e^u\{\lambda_+ s\}^d e^{\lambda s}, \end{aligned} \tag{3.35}$$

where we have twice used $e^{(\lambda_+-\lambda)t} \le 2$ for $t \le t_{\max}(\Lambda)$, as follows from (3.13). For the first term in (3.33), from (3.29), we have

$$\widetilde{M}^K_{s,t} \;\le\; 2^d\nu L^{-1}\left(\overline{N}^K_{t_s}\sum_{j\in\overline{J}_s} r^d_{js} + \overline{N}_s\overline{M}^K_{t_s}\right).$$

Defining

$$B_2(\gamma,s) \;:=\; \left\{\lambda^d\sum_{j\in\overline{J}_s} r^d_{js} + d!\overline{N}_s \;\le\; \Lambda^\gamma e^{\lambda_+ s}\right\},$$

it thus follows from the independence of $\mathcal{L}$ and $\overline{\mathcal{L}}^K$ that, for $t = t_\Lambda(u)$,

$$\eta_\Lambda\mathbb{E}\{\widetilde{M}^K_{s,t}\,|\,\widetilde{\mathcal{F}}_s\}I[B_2(\gamma,s)] \;\le\; 2^{d+1}d!c_1\eta_\Lambda\Lambda^\gamma e^u, \tag{3.36}$$

using (3.13) to bound $e^{(\lambda_+-\lambda)t}$.

To complete the proof of the lemma, we need to show that

$$\mathbb{P}[(B_1(\gamma,s))^c] + \mathbb{P}[(B_2(\gamma,s))^c] \;=\; O(\Lambda^{-\gamma}).$$

For $\mathbb{P}[(B_1(\gamma,s))^c]$, we bound $\mathbb{E}\{\overline{N}_s - N_s^{s,s}\}$ and $\mathbb{E}\{\sum_{j\in\overline{J}_s\setminus J_s^{s,s}} r_{js}^d\}$, and then use Markov's inequality. We begin by bounding the conditional probability $\pi_u^{s,s}$, given the past up to time $u- < s$, that an island of $\overline{X}$, born to an uncensored parent at u, is censored in $\overline{\mathcal{L}}^{s,s}$. Using (3.9), it is no greater than

$$L^{-1}\int_{(0,u)} \nu^d_{s-v+s-u+2s}\,\overline{N}(dv) \;\le\; (1+\eta_\Lambda)\nu L^{-1}(4s)^d\overline{N}_{u-}\,.$$

If it is censored, bounding $\overline{X}^K$ by the branching process $\widehat{X}_+$ and using (2.2) and (2.4), the expected number of its offspring by time s, all of which are also censored, is at most $c_1e^{\lambda_+(s-u)}$, and the expected volume censored at most $c_1 d!\lambda_+^{-d}e^{\lambda_+(s-u)}$. Hence

$$\begin{aligned}
&\mathbb{E}\{\overline{N}_s - N_s^{s,s}\}\\
&\quad\le\; (1+\eta_\Lambda)\nu L^{-1}(4s)^d\,\mathbb{E}\left\{\int_0^s c_1e^{\lambda_+(s-u)}\overline{N}_{u-}\,\overline{N}(du)\right\}\\
&\quad\le\; c_1(1+\eta_\Lambda)\nu L^{-1}(4s)^d\,\mathbb{E}\left\{\int_0^s e^{\lambda_+(s-u)}M_0^+(u)\rho(1+\eta_\Lambda) + \nu M_d^+(u)\,du\right\}\\
&\quad\le\; 4c_1c_2\rho\{\nu\}^2 d!\lambda_+^{-d}L^{-1}(4s)^d\int_0^s e^{\lambda_+(s+u)}\,du\\
&\quad\le\; 4c_1c_2(1+\eta_\Lambda)\Lambda^{-1}(4\lambda s)^d e^{2\lambda_+ s},
\end{aligned}\tag{3.37}$$

again by (2.2) and (2.4), and from Cauchy–Schwarz. Then, by a similar argument,

$$\begin{aligned}
\mathbb{E}\left\{\sum_{j\in\overline{J}_s\setminus J_s^{s,s}} r_{js}^d\right\} \;&\le\; (1+\eta_\Lambda)\nu L^{-1}(4s)^d\,\mathbb{E}\left\{\int_0^s c_1d!\lambda^{-d}e^{\lambda_+(s-u)}\overline{N}_{u-}\,\overline{N}(du)\right\}\\
&\le\; 2c_1c_2d!\Lambda^{-1}(4s)^d e^{2\lambda_+ s}.
\end{aligned}\tag{3.38}$$

Combining (3.37) and (3.38) and using Markov's inequality, $\mathbb{P}[\{B_1(\gamma,s)\}^c] \le c\Lambda^{-\gamma}$, for a constant c depending only on d.

For $\mathbb{P}[(B_2(\gamma,s))^c]$, we again bound $\overline{X}^K$ by the branching process $\widehat{X}_+$ and use (2.2) and (2.4), giving

$$\mathbb{E}\overline{N}_s \;\le\; c_1e^{\lambda_+ s};\qquad \mathbb{E}\left\{\sum_{j\in\overline{J}_s} r_{js}^d\right\} \;\le\; \mathbb{E}M_d^+(s) \;\le\; c_1d!\lambda_+^{-d}e^{\lambda_+ s}. \tag{3.39}$$

Hence, from Markov's inequality, $\mathbb{P}[\{B_1(\gamma,s)\}^c] \le c'\Lambda^{-\gamma}$, for a constant c' depending only on d, and the lemma is proved by taking $B_0(\gamma,s) = B_1(\gamma,s)\cap B_2(\gamma,s)$. ∎

We now replace $\mathbb{E}\{\exp(-\widetilde{M}_{s,t}^K)\,|\,\widetilde{\mathcal{F}}_s\}$ by an expression involving the function ℓ defined in (1.12), and using the quantity $W^*(s)$ defined by

$$W^*(s) \;:=\; e^{-\lambda s}\sum_{l=0}^d\sum_{j\in\overline{J}_s}\frac{(\lambda(s-\overline{\tau}_j))^l}{l!} \;\le\; e^{-\lambda s}\sum_{l=0}^d H_l^+(s) \;=\; e^{(\lambda_+-\lambda)s}W^+(s), \tag{3.40}$$

where the inequality follows from Lemma 3.1, so that, from (3.13), $\mathbb{E}W^*(s) \;\le\; 2$.

Lemma 3.6. *Take* $s \le \lambda^{-1}\log\Lambda$*, and let* $\widetilde{M}^K_{s,t}$ *be defined as in* (3.29)*,* $W^*(s)$ *as in* (3.40) *and* ℓ *as for Lemma 1.12. Then, for any* $\gamma > 0$ *and* $0 < \eta < \zeta(d)$*, there is an event* $B_{3.6}(\gamma,\eta,s) \in \widetilde{\mathcal{F}}_s$ *and constants* $C_{3.6}$ *and* $C'_{3.6}$*, depending only on* d*, such that*

$$\mathbb{P}[\{B_{3.6}(\gamma,\eta,s)\}^c] \;\le\; C_{3.6}(\Lambda^{-\gamma} + \lambda s e^{-2\lambda(\zeta(d)-\eta)s})$$

and that

$$\begin{aligned}\big|\mathbb{E}\{e^{-\widetilde{M}^K_{s,t_\Lambda(u)}} \,|\, \widetilde{\mathcal{F}}_s\} - \big(1 - \ell(\log[\hat{c}_d W^*(s)] + u)\big)\big|\, I[B_{3.6}(\gamma,\eta,s)] \\ \le\; C'_{3.6}(1+e^u)\big(\Lambda^\gamma(\eta_\Lambda \log\Lambda + \Lambda^{-1}e^{\lambda s}) + e^{-\lambda\eta s}\big),\end{aligned}$$

uniformly in $t_\Lambda(u) \le t_{\max}(\Lambda)$*, where* $\hat{c}_d := d!/(d+1)$.

Proof. We first observe, from (3.29) and (1.6) that

$$\begin{aligned}\widetilde{M}^K_{s,t} &= L^{-1}\nu \sum_{j\in J_s}\int_0^{t_s}\sum_{l=0}^{d}\binom{d}{l} r^l_{js}(t_s-u)^{d-l}\overline{N}^K(du) \\ &= \Lambda^{-1}\sum_{l=0}^{d}\binom{d}{l}\left(\sum_{j\in J_s}\{\lambda r_{js}\}^l\right)\left(\int_0^{t_s}\{\lambda(t_s-u)\}^{d-l}\overline{N}^K(du)\right), \qquad (3.41)\end{aligned}$$

with $r_{js} := s - \bar{\tau}_j$ as before. Now realize $\widehat{X}_-$, $\overline{X}$ and $\widehat{X}^+$ together as in Lemma 3.1, so that

$$H^-_l(s) \;\le\; \sum_{j\in J_s}\frac{(\lambda r_{js})^l}{l!} \;\le\; H^+_l(s) \text{ a.s.,} \quad \text{for } 0 \le s \le t_{\max}(\Lambda). \qquad (3.42)$$

Then, for such s, it follows from (2.47), then using Lemma 4.2, (2.40) and (2.41), that, on an event $B^+_1(\eta,s) \in \mathcal{F}^{++}_s$ such that $\mathbb{P}[\{B^+_1(\eta,s)\}^c] \le c(d)(1+\lambda s)e^{-2\lambda(\zeta(d)-\eta)s}$, we have

$$\sum_{l=0}^{d} c_l H^+_l(s) \;\le\; C(1)\left(\frac{1}{d+1}\sum_{l=0}^{d} H^+_l(s) + e^{\lambda_+(1-\eta)s}\right) \qquad (3.43)$$

and

$$\sum_{l=0}^{d} c_l H^-_l(s) \;\ge\; C(1)\left(\frac{1}{d+1}\sum_{l=0}^{d} H^-_l(s) - e^{\lambda_-(1-\eta)s}\right), \qquad (3.44)$$

for *all* choices of $c_0,\ldots,c_d$, where $C(1) := \sum_{l=0}^d c_l$. Define

$$B^+_2(\gamma,s) \;:=\; \left\{\frac{1}{d+1}\sum_{l=0}^{d}(H^+_l(s) - H^-_l(s)) \le e^{\lambda s}\Lambda^\gamma\eta_\Lambda\log\Lambda\right\} \;\in\; \mathcal{F}^{++}_s. \qquad (3.45)$$

Then, on $B^+_1(\eta,s) \cap B^+_2(\gamma,s)$ and for $0 \le s \le t_{\max}(\Lambda)$, we have

$$\begin{aligned}\sum_{j\in J_s}\frac{(\lambda r_{js})^l}{l!} &\le H^+_l(s) \;\le\; \frac{1}{d+1}\sum_{l=0}^{d} H^+_l(s) + e^{\lambda_+(1-\eta)s} \\ &\le \frac{1}{d+1}\sum_{l=0}^{d} H^-_l(s) + e^{\lambda_+(1-\eta)s} + e^{\lambda s}\Lambda^\gamma\eta_\Lambda\log\Lambda \\ &\le \frac{1}{d+1}e^{\lambda s}W^*(s) + \varepsilon(\gamma,\eta,s), \qquad (3.46)\end{aligned}$$

for all $0 \le l \le d$, from (3.42), (3.43), (3.40) and (3.45), where

$$\varepsilon(\gamma,\eta,s) \;:=\; 2e^{\lambda(1-\eta)s} + e^{\lambda s}\Lambda^\gamma\eta_\Lambda\log\Lambda.$$

Arguing analogously, we also deduce that

$$\sum_{j\in J_s} \frac{(\lambda r_{js})^l}{l!} \;\geq\; \frac{1}{d+1} e^{\lambda s} W^*(s) - \varepsilon(\gamma,\eta,s).$$

Now $\mathbb{P}[\{B_1^+(\eta,s)\}^c] \leq c(d)(1+\lambda s)e^{-2\lambda(\zeta(d)-\eta)s}$. Then, since

$$\mathbb{E}\left\{\sum_{l=0}^d H_l(s)\right\} \;=\; e^{\lambda s}\mathbb{E}W(s) \;=\; e^{\lambda s},$$

and using (3.13), we have

$$\mathbb{E}\left\{\sum_{l=0}^d (H_l^+(s) - H_l^-(s))\right\} \;=\; e^{\lambda_+ s} - e^{\lambda_- s} \;\leq\; 8e^{\lambda s}\eta_\Lambda \log \Lambda$$

in $0 \leq s \leq t_{\max}(\Lambda)$, and hence, by Markov's inequality,

$$\mathbb{P}[\{B_2^+(\gamma,s)\}^c] \;\leq\; 8\Lambda^{-\gamma}. \tag{3.47}$$

Thus the event

$$B_3(\gamma,\eta,s) \;:=\; \bigcap_{l=0}^d \left\{\left|\sum_{j\in J_s} \frac{(\lambda r_{js})^l}{l!} - \frac{1}{d+1}e^{\lambda s}W^*(s)\right| \leq \varepsilon(\gamma,\eta,s)\right\} \;\in\; \widetilde{\mathcal{F}}_s \tag{3.48}$$

is such that

$$\mathbb{P}[\{B_3(\gamma,\eta,s)\}^c] \;\leq\; C_1(d)(\Lambda^{-\gamma} + \lambda s e^{-2\lambda(\zeta(d)-\eta)s}), \tag{3.49}$$

for a suitable constant $C_1(d)$.

Now, taking $c_l := \Lambda^{-1}C_l(s,t)$, where

$$C_l(s,t) \;:=\; \int_0^{t_s} \frac{d!\{\lambda(t_s-u)\}^{d-l}}{(d-l)!} \overline{N}^K(du), \tag{3.50}$$

(3.41) implies that

$$\begin{aligned}\left|\widetilde{M}_{s,t}^K - \frac{e^{\lambda s}W^*(s)}{\Lambda(d+1)}\sum_{l=0}^d C_l(s,t)\right| I[B_3(\gamma,\eta,s)] \\ \leq\; \Lambda^{-1}\sum_{l=0}^d C_l(s,t)e^{\lambda s}\{2e^{-\lambda\eta s} + \Lambda^\gamma \eta_\Lambda \log\Lambda\}.\end{aligned}$$

Hence also

$$\begin{aligned}\left|\mathbb{E}\left(e^{-\widetilde{M}_{s,t}^K} \,|\, \widetilde{\mathcal{F}}_s\right) - \mathbb{E}\left(\exp\left\{-\frac{e^{\lambda s}W^*(s)}{\Lambda(d+1)}\sum_{l=0}^d C_l(s,t)\right\} \,\Big|\, \widetilde{\mathcal{F}}_s\right)\right| I[B_3(\gamma,\eta,s)] \\ \leq\; \Lambda^{-1}\mathbb{E}\left\{\sum_{l=0}^d C_l(s,t) \,\Big|\, \widetilde{\mathcal{F}}_s\right\} e^{\lambda s}\{2e^{-\lambda\eta s} + \Lambda^\gamma\eta_\Lambda\log\Lambda\}.\end{aligned} \tag{3.51}$$

Now, because $\overline{X}^K$ can also be bounded between copies $\widehat{X}_-^K$ and $\widehat{X}_+^K$ of $\widehat{X}_-$ and $\widehat{X}_+$, using Lemma 3.1, we have the inequality

$$d!H_{d-l}^{K,-}(t_s) \;\leq\; C_l(s,t) \;\leq\; d!H_{d-l}^{K,+}(t_s), \qquad 0 \leq l \leq d. \tag{3.52}$$

Hence, since the K-processes can be chosen to be independent of $\widetilde{\mathcal{F}}_s$, it follows that

$$\mathbb{E}\left\{\sum_{l=0}^{d} C_l(s,t)e^{\lambda s} \,|\, \widetilde{\mathcal{F}}_s\right\} \;\le\; d!e^{\lambda_+(t-s)+\lambda s}\mathbb{E}\{W^K(t_s)\} \;\le\; 2\,d!e^{\lambda t}, \tag{3.53}$$

for any $0 < s \le t \le t_{\max}(\Lambda)$. Thus, from (3.51), it follows that

$$\left|\mathbb{E}\left(e^{-\widetilde{M}^K_{s,t}} \,|\, \widetilde{\mathcal{F}}_s\right) - \mathbb{E}\left(\exp\left\{-\frac{e^{\lambda s}W^*(s)}{\Lambda(d+1)}\sum_{l=0}^{d} C_l(s,t)\right\} \,\Big|\, \widetilde{\mathcal{F}}_s\right)\right| I[B_3(\gamma,\eta,s)]$$
$$\le\; 2\,d!\Lambda^{-1}e^{\lambda t}\{2e^{-\lambda\eta s} + \Lambda^{\gamma}\eta_\Lambda \log\Lambda\}. \tag{3.54}$$

The next step is to examine the difference

$$\left|\mathbb{E}\left(\exp\left\{-\frac{e^{\lambda s}W^*(s)}{\Lambda(d+1)}\sum_{l=0}^{d} C_l(s,t)\right\} \,\Big|\, \widetilde{\mathcal{F}}_s\right) - \phi^1_{\lambda(t_\Lambda(u)-s)}(\hat{c}_d e^u W^*(s))\right|,$$

where $\phi^1_s(\theta) := \mathbb{E}\{e^{-\theta W^1(s)}\}$. To start with, from (3.52) and Lemma 2.1,

$$d!e^{\lambda_- t_s}W^{K,-}(t_s) \;\le\; \sum_{l=0}^{d} C_l(s,t) \;\le\; d!e^{\lambda_+ t_s}W^{K,+}(t_s).$$

Hence, for any non-negative and $\widetilde{\mathcal{F}}_s$-measurable random variable Θ_s, we have

$$\phi^+_{t_s}(\Theta_s d!e^{\lambda_+ t_s}) \;\le\; \mathbb{E}\left\{\exp\left(-\Theta_s\sum_{l=0}^{d} C_l(s,t)\right) \,\Big|\, \widetilde{\mathcal{F}}_s\right\} \;\le\; \phi^-_{t_s}(\Theta_s d!e^{\lambda_- t_s}), \tag{3.55}$$

where

$$\phi^+_t(\theta) \;:=\; \mathbb{E}\{e^{-\theta W^+(t)}\} \;=\; \phi^1_{\lambda_+ t}(\theta) \quad\text{and}\quad \phi^-_t(\theta) \;:=\; \mathbb{E}\{e^{-\theta W^-(t)}\} \;=\; \phi^1_{\lambda_- t}(\theta), \tag{3.56}$$

and ϕ^1 is as above, with the final equalities a consequence of (2.14). Since $\lambda(1-\eta_\Lambda) \le \lambda_- \le \lambda_+ \le \lambda(1+\eta_\Lambda)$, we conclude from Lemma 4.2 and (3.13) that

$$\max\{|\phi^+_t(\theta e^{\lambda_+ t}) - \phi^+_t(\theta e^{\lambda t})|, |\phi^-_t(\theta e^{\lambda_- t}) - \phi^-_t(\theta e^{\lambda t})|\} \;\le\; 2e^{-1}\eta_\Lambda\lambda t;$$
$$\max\{|\phi^+_t(\theta) - \phi^1_{\lambda t}(\theta)|, |\phi^-_t(\theta) - \phi^1_{\lambda t}(\theta)|\} \;\le\; \theta e^{-1}\eta_\Lambda\,\lambda t e^{-\lambda t}, \tag{3.57}$$

as long as $t \le t_{\max}(\Lambda)$. Taking $\Theta(s) := \{(d+1)\Lambda\}^{-1}e^{\lambda s}W^*(s)$ and $t = t_\Lambda(u)$, and using (3.55), (3.57) and (3.13), this gives

$$\left|\mathbb{E}\left\{\exp\left(-\frac{e^{\lambda s}W^*(s)}{\Lambda(d+1)}\sum_{l=0}^{d} C_l(s,t)\right) \,\Big|\, \widetilde{\mathcal{F}}_s\right\} - \phi^1_{\lambda(t-s)}(\hat{c}_d e^u W^*(s))\right|$$
$$\le\; 4e^{-1}\eta_\Lambda\lambda t_s + e^{-1}\Theta(s)d!e^{\lambda_+ t_s}\eta_\Lambda\,\lambda t_s e^{-\lambda t_s}$$
$$\le\; 4e^{-1}\eta_\Lambda(\log\Lambda + u) + 3\hat{c}_d e^{\lambda s}W^*(s)\Lambda^{-1}\eta_\Lambda\log\Lambda. \tag{3.58}$$

From (3.40), we have $\mathbb{E}\{W^*(s)\} \le 2$. Thus, defining

$$B_4(\gamma,s) \;:=\; \{W^*(s) \le \Lambda^\gamma\} \;\in\; \widetilde{\mathcal{F}}_s, \tag{3.59}$$

it follows that $\mathbb{P}[\{B_4(\gamma,s)\}^c] \le 2\Lambda^{-\gamma})$, and, combining (3.54) and (3.58), that

$$\left|\mathbb{E}\{e^{-\widetilde{M}^K_{s,t_\Lambda(u)}} \,|\, \widetilde{\mathcal{F}}_s\} - \phi^1_{\lambda(t_\Lambda(u)-s)}(\hat{c}_d e^u W^*(s))\right| I[B_3(\gamma,\eta,s) \cap B_4(\gamma,s)]$$
$$\le\; C_2(d)(\Lambda^\gamma\eta_\Lambda\log\Lambda + e^{-\lambda\eta s}\Lambda^{-1}e^{\lambda s}), \tag{3.60}$$

uniformly in $t_\Lambda(u) \le t_{\max}(\Lambda)$. But now, from Lemma 4.2, on the event $B_4(\gamma, s)$,

$$\begin{aligned} |\phi^1_{\lambda(t_\Lambda(u)-s)}(\hat{c}_d e^u W^*(s)) - \phi^1_\infty(\hat{c}_d e^u W^*(s))| &\le \frac{1}{2e}\hat{c}_d e^u W^*(s)\exp\{-\lambda(t_\Lambda(u)-s)\} \\ &\le \frac{1}{2e}\hat{c}_d \Lambda^{\gamma-1}e^{\lambda s}, \end{aligned}$$

and $\phi^1_\infty(\hat{c}_d e^u W^*(s)) = 1 - \ell(\log(\hat{c}_d W^*(s)) + u)$ by (1.12), (2.14) and (2.18). This establishes the lemma, with $B_{3.6}(\gamma, \eta, s) := B_3(\gamma, \eta, s) \cap B_4(\gamma, s)$, in view of (3.49) and (3.59). ∎

3.3 Replacing $W^*(s)$ by $W(s, v)$

Our aim is to approximate the conditional distribution of $L_{t_\Lambda(u)}/L$, given $\widetilde{\mathcal{F}}_v$, for suitably chosen v. After Lemma 3.6, the problem has largely been reduced to considering the conditional distribution of $W^*(s)$. However, in order to use the results of Section 2, it is advantageous to replace $W^*(s)$ by a function of a *flattened* branching process; $W^*(s)$ is constructed from the birth times $\bar{\tau}_j$ of the original branching process $\overline{X}$. Accordingly, we define

$$W(s, v) \;:=\; e^{-\lambda s}\sum_{l=0}^{d} H^0_l(s-v, v), \quad s \ge v, \tag{3.61}$$

for $H^0_l(\cdot, v)$, $0 \le l \le d$, corresponding to the (flattened) branching process $\widehat{X}_0$ of Lemma 3.1, taken to have initial condition $H^0_l(0, v) = \sum_{j \in \overline{J}_v}(\lambda(v - \bar{\tau}_j))^l/l! \in \sigma(\widetilde{H}(v))$, $0 \le l \le d$. Note that $W(v, v) = W^*(v)$. The error involved in replacing $W^*(s)$ by $W(s, v)$ is bounded in the following lemma.

Lemma 3.7. *For $v \le s \le \lambda^{-1}\log\Lambda$, we have*

$$\mathbb{E}\left\{\left|\ell(\log[\hat{c}_d e^u W^*(s)] + u) - \ell(\log[\hat{c}_d e^u W(s, v)] + u)\right| \,\Big|\, \widetilde{\mathcal{F}}_v\right\} \;\le\; 4\hat{c}_d e^u W^*(v)\eta_\Lambda \log\Lambda.$$

Proof. We once more use Lemma 3.1 to justify that both $W^*(s)$ and $W(s, v)$ belong to the interval

$$\Big[e^{-\lambda s}\sum_{l=0}^{d} H^-_l(s-v, v), e^{-\lambda s}\sum_{l=0}^{d} H^+_l(s-v, v)\Big], \tag{3.62}$$

where the processes $\widehat{X}_-(\cdot, v)$ and $\widehat{X}_+(\cdot, v)$ both have the same initial condition as $\widehat{X}_0(\cdot, v)$. Now

$$\mathbb{E}\left\{\sum_{l=0}^{d} H^+_l(s-v, v)\,\Big|\,\widetilde{\mathcal{F}}_v\right\} \;=\; e^{\lambda_+(s-v)}\sum_{l=0}^{d} H^0_l(0, v)$$

and

$$\mathbb{E}\left\{\sum_{l=0}^{d} H^-_l(s-v, v)\,\Big|\,\widetilde{\mathcal{F}}_v\right\} \;=\; e^{\lambda_-(s-v)}\sum_{l=0}^{d} H^0_l(0, v);$$

hence

$$\begin{aligned} &\mathbb{E}\{|W^*(s) - W(s, v)|\,|\,\widetilde{\mathcal{F}}_v\} \\ &\quad\le\; e^{-\lambda s}\mathbb{E}\left\{\sum_{l=0}^{d}\{H^+_l(s-v, v) - H^-_l(s-v, v)\}\,\Big|\,\widetilde{\mathcal{F}}_v\right\} \\ &\quad\le\; e^{-\lambda v}\sum_{l=0}^{d} H^0_l(0, v)\{e^{(\lambda_+-\lambda)(s-v)} - e^{(\lambda_--\lambda)(s-v)}\} \;\le\; 4W^*(v)\eta_\Lambda\log\Lambda, \end{aligned}$$

by (3.13). This, together with (1.12) and Lemma 4.2, implies that

$$\begin{aligned}&\mathbb{E}\left\{\left|\ell(\log[\hat{c}_d e^u W^*(s)]+u)-\ell(\log[\hat{c}_d e^u W(s,v)]+u)\right| \,\Big|\, \widetilde{\mathcal{F}}_v\right\}\\&\quad\le\ \mathbb{E}\{\hat{c}_d e^u|W(s,v)-W^*(s)|\,|\,\widetilde{\mathcal{F}}_v\}\ \le\ 4\hat{c}_d e^u W^*(v)\eta_\Lambda\log\Lambda,\end{aligned}$$

as required. ∎

We now combine the results of Lemmas 3.2–3.7 to give the following result, relating the distribution of $L_{t_\Lambda(u)}/L$ to that of $\ell(\log[\hat{c}_d e^u W(s,v)]+u)$.

Corollary 3.8. *Take* $v:=\alpha_1\lambda^{-1}\log\Lambda$ *and* $s:=\alpha_2\lambda^{-1}\log\Lambda$ *for* $0<\alpha_1<\alpha_2<1$*, and fix* $0<\eta<\zeta(d)$*. Then there is an event* $B_{3.8}(\gamma,\eta,v)\in\widetilde{\mathcal{F}}_v$*, and constants* $C^0_{3.8}:=C^0_{3.8}(u_0,d)$*,* $C^1_{3.8}:=C^1_{3.8}(u_0,d)$ *and* $C^2_{3.8}:=C^2_{3.8}(d)$*, such that*

$$\mathbb{E}\left\{|f(L_{t_\Lambda(u)}/L)-f(\ell(\log[\hat{c}_d W(s,v)]+u))|\,\Big|\,\widetilde{\mathcal{F}}_v\right\}\ \le\ C^0_{3.8}\|f\|_\infty p_\Lambda+C^1_{3.8}\|f'\|_\infty\varepsilon_\Lambda,$$

and such that $\mathbb{P}[\{B_{3.8}(\gamma,\eta,v)\}^c]\ \le\ C^2_{3.8}p_\Lambda$*, where*

$$\begin{aligned}\varepsilon_\Lambda\ &:=\ \Lambda^\gamma\{\Lambda^{-\alpha_2/2}(\log\Lambda)^{d/2}+\Lambda^{\alpha_2-1}(\log\Lambda)^d+\Lambda^{-\alpha_1}+\eta_\Lambda\log\Lambda\}+\Lambda^{-\alpha_2\eta};\\ p_\Lambda\ &:=\ \Lambda^{-\gamma/2}+\Lambda^{-(\zeta(d)-\eta)}\log\Lambda.\end{aligned}$$

Proof. We take the results of Lemmas 3.2–3.7 in turn. Using Lemma 3.1, we have

$$\begin{aligned}\mathbb{E}\{\lambda^d\overline{M}_s+\overline{N}_s\,|\,\widetilde{\mathcal{F}}_v\}\ &\le\ \mathbb{E}\{d!H^+_d(s)+H^+_0(s)\,|\,\widetilde{H}(v)\}\\&\le\ d!\mathbb{E}\{W^+(s)e^{\lambda_+s}\,|\,\widetilde{H}(v)\}\ \le\ 2\,d!W^*(v)e^{\lambda s}.\end{aligned}\tag{3.63}$$

Define the event $B^{(1)}_{3.8}(\gamma,v):=\{W^*(v)\le\Lambda^\gamma\}$, whose probability is at most $\Lambda^{-\gamma}$, by Markov's inequality. Then, from Lemma 3.2 and (3.63), it follows that

$$\mathbb{E}\{\mathrm{Var}\,\{L_t/L\,|\,\widetilde{\mathcal{F}}_s\}\,|\,\widetilde{\mathcal{F}}_v\}\ \le\ C_{3.2}2\,d!W^*(v)\Lambda^{-2}(1+(\lambda s)^d)e^{\lambda(2t-s)},$$

implying that, on $B^{(1)}_{3.8}(\gamma,v)$, we have

$$\mathbb{E}\left\{|1-(L_{t_\Lambda(u)}/L)-\mathbb{E}\{1-(L_{t_\Lambda(u)}/L)\,|\,\widetilde{\mathcal{F}}_s\}|\,\Big|\,\widetilde{\mathcal{F}}_v\right\}\ \le\ C_a(d)\Lambda^{\gamma-\alpha_2/2}e^u(\log\Lambda)^{d/2}.\tag{3.64}$$

Next, from Lemma 3.4 and (3.63) and on the event $B^{(1)}_{3.8}(\gamma,v)$, we have

$$\begin{aligned}&\mathbb{E}\left\{\left|\mathbb{E}\{1-(L_{t_\Lambda(u)}/L)\,|\,\widetilde{\mathcal{F}}_s\}-\mathbb{E}\{\exp\{-M^K_{s,t_\Lambda(u)}\}\,|\,\widetilde{\mathcal{F}}_s\}\right|\,\Big|\,\widetilde{\mathcal{F}}_v\right\}\\&\quad\le\ C_b(d)W^*(v)(\log\Lambda)^d\{\Lambda^{-1}e^{\lambda s}+e^{2u}e^{-\lambda s)}\}\\&\quad\le\ C_b(d)\Lambda^\gamma(\log\Lambda)^d\{\Lambda^{\alpha_2-1}+e^{2u}\Lambda^{-\alpha_1}\}.\end{aligned}\tag{3.65}$$

Turning to Lemma 3.5, we find that

$$\begin{aligned}&\mathbb{E}\left\{\left|\mathbb{E}\{\exp\{-M^K_{s,t_\Lambda(u)}\}\,|\,\widetilde{\mathcal{F}}_s\}-\mathbb{E}\{\exp\{-\widetilde{M}^K_{s,t_\Lambda(u)}\}\,|\,\widetilde{\mathcal{F}}_s\}\right|I[B_{3.5}(\gamma,s)]\,\Big|\,\widetilde{\mathcal{F}}_v\right\}\\&\quad\le\ C_c(d)\Lambda^\gamma e^u\{\Lambda^{-1}\{\log\Lambda\}^d e^{\lambda s}+\eta_\Lambda\}\\&\quad=\ C_c(d)\Lambda^\gamma e^u\{\Lambda^{\alpha_2-1}\{\log\Lambda\}^d+\eta_\Lambda\}.\end{aligned}\tag{3.66}$$

Then, from Lemma 3.6, we have

$$\begin{aligned}&\mathbb{E}\left\{\left|\mathbb{E}\{e^{-\widetilde{M}^K_{s,t_\Lambda(u)}}\,|\,\widetilde{\mathcal{F}}_s\}-\ell(\log[\hat{c}_d W^*(s)]+u)\right|I[B_{3.6}(\gamma,\eta,s)]\,\Big|\,\widetilde{\mathcal{F}}_v\right\}\\&\quad\le\ C'_{3.6}(1+e^u)(\Lambda^\gamma(\eta_\Lambda\log\Lambda+\Lambda^{-1}e^{\lambda s})+e^{-\lambda\eta s})\\&\quad=\ C'_{3.6}(1+e^u)(\Lambda^\gamma(\eta_\Lambda\log\Lambda+\Lambda^{\alpha_2-1})+\Lambda^{-\alpha_2\eta}).\end{aligned}\tag{3.67}$$

Finally, from Lemma 3.7, on the event $B^{(1)}_{3.8}(\gamma,v)$, we have

$$\mathbb{E}\left\{\left|\ell(\log[\hat{c}_d W^*(s)]+u)-\ell(\log[\hat{c}_d W(s,v)]+u)\right|\,\Big|\,\widetilde{\mathcal{F}}_v\right\} \;\le\; 4\hat{c}_d e^u \Lambda^\gamma \eta_\Lambda \log\Lambda. \tag{3.68}$$

Combining (3.64) to (3.68), we deduce that, on the event $B^{(1)}_{3.8}(\gamma,v)$, and uniformly in $u \le u_0$,

$$\begin{aligned}
&\mathbb{E}\left\{\left|(L_{t_\Lambda(u)}/L)-\ell(\log[\hat{c}_d W(s,v)]+u)\right| I[\widehat{B}(\gamma,\eta,s)]\,\Big|\,\widetilde{\mathcal{F}}_v\right\}\\
&\quad\le\; C_*(d,u_0)\big(\Lambda^\gamma\{\Lambda^{-\alpha_2/2}(\log\Lambda)^{d/2}+\Lambda^{\alpha_2-1}(\log\Lambda)^d+\Lambda^{-\alpha_1}+\eta_\Lambda\log\Lambda\}+\Lambda^{-\alpha_2\eta}\big)\\
&\quad=:\; C_*(d,u_0)\varepsilon_\Lambda,
\end{aligned} \tag{3.69}$$

where $\widehat{B}(\gamma,\eta,s) := B_{3.6}(\gamma,\eta,s)\cap B_{3.5}(\gamma,s)$.

For the exceptional set, from Lemmas 3.6 and 3.5, we have

$$\begin{aligned}
\mathbb{P}[\{\widehat{B}(\gamma,\eta,s)\}^c] &\le C_{3.6}\{\Lambda^{-\gamma}+\lambda s e^{-2\lambda(\zeta(d)-\eta)s})+C_{3.5}\Lambda^{-\gamma}\}\\
&\le C_e(d)\{\Lambda^{-\gamma}+\Lambda^{-2(\zeta(d)-\eta)}\log\Lambda\}.
\end{aligned}$$

On the other hand, for any set $B\in\mathcal{F}$ with $\mathbb{P}[B]=p$, and for any σ-field $\mathcal{G}\subset\mathcal{F}$, we have

$$p \;=\; \mathbb{P}[B] \;\ge\; \mathbb{P}[\{\mathbb{P}[B\,|\,\mathcal{G}]>\sqrt{p}\}]\sqrt{p},$$

by the total probability formula, implying that $\mathbb{P}[B\,|\,\mathcal{G}]\le\sqrt{p}$ with probability at least $1-\sqrt{p}$. Hence there is an event $B^{(2)}_{3.8}(\gamma,\eta,v)\in\widetilde{\mathcal{F}}_v$, whose complement has probability at most

$$(C_e(d))^{1/2}\{\Lambda^{-\gamma/2}+\Lambda^{-(\zeta(d)-\eta)}\log\Lambda\} \;=:\; (C_e(d))^{1/2}p_\Lambda, \tag{3.70}$$

on which $\mathbb{P}[\{\widehat{B}(\gamma,\eta,s)\}^c\,|\,\widetilde{\mathcal{F}}_v]\le(C_e(d))^{1/2}p_\Lambda$. Now define $Z_u := \ell(\log[\hat{c}_d W(s,v)]+u)$ and $Y_u := L_{t_\Lambda(u)}/L$. Then, for any bounded Lipschitz function f, we conclude from (3.69) and (3.70) that, for $u\le u_0$ and on the event

$$B_{3.8}(\gamma,\eta,v) \;:=\; B^{(1)}_{3.8}(\gamma,v)\cap B^{(2)}_{3.8}(\gamma,\eta,v),$$

we have

$$\begin{aligned}
&\mathbb{E}\left\{|\mathbb{E}\{f(Y_u)\}-\mathbb{E}\{f(Z_u)\}|\,|\,\widetilde{\mathcal{F}}_v\right\}\\
&\quad\le\; \mathbb{E}\Big\{|\mathbb{E}\{f(Y_u)\}-\mathbb{E}\{f(Z_u)\}|I[\widehat{B}(\gamma,\eta,s)]\\
&\qquad\qquad +|\mathbb{E}\{f(Y_u)\}-\mathbb{E}\{f(Z_u)\}|I[\{\widehat{B}(\gamma,\eta,s)\}^c]\,\Big|\,\widetilde{\mathcal{F}}_v\Big\}\\
&\quad\le\; \|f'\|_\infty\mathbb{E}\{|Y_u-Z_u|I[\widehat{B}(\gamma,\eta,s)]\,|\,\widetilde{\mathcal{F}}_v\}+2\|f\|_\infty\mathbb{P}[\{\widehat{B}(\gamma,\eta,s)\}^c\,|\,\widetilde{\mathcal{F}}_v]\\
&\quad\le\; \|f'\|_\infty C_*(d,u_0)\varepsilon_\Lambda+2\|f\|_\infty(C_e(d))^{1/2}p_\Lambda.
\end{aligned}$$

This proves the corollary. ∎

3.4 The main theorem

We now use Corollary 3.8 to compare the conditional distributions, given $\widetilde{\mathcal{F}}_v$, of the *normalized* random variables $Y(u,v)$ and $Z(u,v)$, where

$$\begin{aligned}
Y(u,v) &:= e^{\lambda v/2}\{(L_{t_\Lambda(u)}/L)-\ell(\log[\hat{c}_d W^*(v)]+u)\};\\
Z(u,v) &:= e^{\lambda v/2}\{\ell(\log[\hat{c}_d W(s,v)]+u)-\ell(\log[\hat{c}_d W^*(v)]+u)\},
\end{aligned} \tag{3.71}$$

for a careful choice of s, with the centring constant $\ell(\log[\hat{c}_d W^*(v)] + u)$ chosen because $W^*(v) = \mathbb{E}\{W(s,v) \,|\, \widetilde{\mathcal{F}}_v\}$. These are the correct standardizations to achieve a non-trivial limit. Thus we wish to compare $\mathbb{E}f(Y(u,v))$ with $\mathbb{E}f(Z(u,v))$, for Lipschitz functions f that have $\|f\|_\infty \le 1$ and $\|f'\|_\infty \le 1$. This corresponds to taking $\|f\|_\infty \le 1$ and $\|f'\|_\infty \le e^{\lambda v/2}$ in Corollary 3.8, because of the pre-factors $e^{\lambda v/2}$ in the definitions of $Y(u,v)$ and $Z(u,v)$. Thus, although p_Λ is already small for large Λ, if $\eta < \zeta(d)$, we need also to show that, for $v = \alpha_1 \lambda^{-1} \log \Lambda$, it is possible to choose α_2, η and γ so as to make $e^{\lambda v/2}\varepsilon_\Lambda = \Lambda^{\alpha_1/2}\varepsilon_\Lambda$ small with Λ. Recalling the definition (3.11) of η_Λ, the expression for ε_Λ in Corollary 3.8 shows that this is the case, for $\gamma > 0$ chosen small enough, if,

$$\alpha_1 \;<\; \alpha_2; \quad \alpha_2 \;<\; 1 - \alpha_1/2; \quad \gamma < \alpha_1/2; \quad \alpha_1 < 2\alpha_2\eta \;\text{ and }\; \alpha_1 \;<\; 2\gamma_g/d.$$

So, for

$$\alpha_1 \;<\; 2\min\{\gamma_g/d, \zeta(d)/(1+\zeta(d)\},$$

choose $0 < \eta < \zeta(d)$ so that $2\eta/(1+\eta) > \alpha_1$ and then α_2 so that $\alpha_1/(2\eta) < \alpha_2 < 1 - \alpha_1/2$; then, if we choose

$$0 \;<\; \gamma \;=\; \tfrac{2}{3}\min\{\gamma_g/d - \alpha_1/2, (\alpha_2-\alpha_1)/2, 1-\alpha_1/2-\alpha_2, \alpha_1/2, \alpha_2\eta - \alpha_1/2\}, \qquad (3.72)$$

it follows that there are constants $C = C(d,u_0)$ and $C' = C'(d)$ such that

$$|\mathbb{E}\{f(Y(u,v)) \,|\, \mathcal{F}_v\} - \mathbb{E}\{f(Z(u,v)) \,|\, \mathcal{F}_v\}| \;\le\; C\{\Lambda^{-\gamma/2}(\log\Lambda)^d + \Lambda^{-(\zeta(d)-\eta)}\}, \qquad (3.73)$$

for all $f \in F_{\mathrm{BW}}$, except on an event of probability at most $C'\{\Lambda^{-\gamma/2} + \Lambda^{-(\zeta(d)-\eta)}\}$. Particular choices are to take

$$\eta \;:=\; \frac{1}{2}\left(\zeta(d) + \frac{\alpha_1}{2-\alpha_1}\right), \quad \text{and} \quad \alpha_2 \;:=\; \frac{1}{2}\left\{1 + \frac{\alpha_1}{2}\left(\frac{1}{\eta} - 1\right)\right\}, \qquad (3.74)$$

in which case we can take any $0 < \gamma' < \min\{\gamma/2, (\zeta(d)-\eta)\}$, and express the error in (3.73) as $C\Lambda^{-\gamma'}$, except on an event of probability at most $C'\Lambda^{-\gamma'}$, albeit with different constants $C = C(u_0,d)$ and $C'(d)$.

Corollary 3.8 and (3.73) compare the distribution of $L_{t_\Lambda(u)}/L$ with that of the quantity $\ell(\log[\hat{c}_d W(s,v)] + u)$, for any $u \le u_0$. The path of $L_{t_\Lambda(u)}/L$ is approximated, to first order, by a time shift of the deterministic path $\ell(u)$, and the shift is the same throughout the path, being determined by the value of the single $\widetilde{\mathcal{F}}_s$-measurable random variable $W(s,v)$. In the remaining argument, we exploit this to show that, to a good approximation, the path after time v is that of the approximation $\ell(\log[\hat{c}_d W^*(v)] + \cdot)$, together with a perturbation that can be expressed in the form $e^{-\lambda v/2} N h_v(\cdot)$, where $h_v(\cdot)$ is an $\widetilde{\mathcal{F}}_v$-measurable function depending on the value of $W^*(v)$, and $\mathcal{L}(N \,|\, \widetilde{\mathcal{F}}_v)$ is the standard normal distribution.

To do so, in view of (3.73), we now need a central limit theorem for $Z(u,v)$ as defined in (3.71). Writing

$$K_2(u,v) \;\;:=\;\; (D\ell)(u + \log[\hat{c}_d W^*(v)])/W^*(v) \;=\; k\frac{d}{dx}\{\ell(\log x)\}\Big|_{kW^*(v)}, \qquad (3.75)$$

where the final equality holds for all $k > 0$, the next lemma shows that $Z(u,v)$ is close in distribution to $K_2(u,v)\,e^{\lambda v/2}\{W(s,v) - W^*(v)\}$.

Lemma 3.9. *Let $Z(u,v)$ be defined as in* (3.71)*, and let $v := \alpha_1\lambda^{-1}\log\Lambda$ and $s := \alpha_2\lambda^{-1}\log\Lambda$; suppose that γ is as for* (3.72) *and $\gamma' = \frac{1}{2}\min\{\gamma/2, (\zeta(d)-\eta)\}$, where η is as in* (3.74)*. Then there is a constant $C = C(d,u_0)$ such that, for all $f \in F_{\mathrm{BW}}$, and on the event $\{W^*(v) \le \Lambda^\gamma\}$,*

$$|\mathbb{E}\{f(Z(u,v)) \,|\, \widetilde{\mathcal{F}}_v\} - \mathbb{E}\{f(K_2(u,v)\,e^{\lambda v/2}\{W(s,v) - W^*(v)\}) \,|\, \widetilde{\mathcal{F}}_v\}| \;\le\; C\Lambda^{-\gamma'},$$

uniformly in $u \le u_0$.

Proof. From (1.12), we have $g(x) := \ell(\log x) = 1 - \mathbb{E}\{e^{-xW}\}$, so that, by Taylor's expansion, for any $x, y > 0$, we can write

$$|g(x+y) - (g(x) + yg'(x))| \;\le\; \tfrac{1}{2}y^2\|g''\|_\infty \;=\; \tfrac{1}{2}y^2\mathbb{E}W^2 \;\le\; \tfrac{1}{2}y^2.$$

from (2.17). Thus, in making a linear approximation to

$$\ell(\log[kW(s,v)]) - \ell(\log[kW^*(v)]) \;=\; g(kW(s,v)) - g(kW^*(v)),$$

the remainder term can be bounded by $\frac{1}{2}k^2(W(s,v) - W^*(v))^2$. Now, because $W^*(v) = \mathbb{E}\{W(s,v) \,|\, \widetilde{\mathcal{F}}_v\}$, we have

$$\mathbb{E}\{(W(s,v) - W^*(v))^2 \,|\, \widetilde{\mathcal{F}}_v\} \;=\; V(s,v) \;:=\; \mathrm{Var}\,(W(s,v) \,|\, \widetilde{\mathcal{F}}_v) \;\le\; W^*(v)e^{-\lambda v},$$

where the inequality follows using (2.17). Hence, for any $k > 0$, and using (3.75), we have

$$\begin{aligned}\mathbb{E}\big\{\big|e^{\lambda v/2}\{\ell(\log[kW(s,v)]) - \ell(\log[kW^*(v)])\} - e^{\lambda v/2}\{W(s,v) - W^*(v)\}\,K_2(u,v)\big| \,|\, \widetilde{\mathcal{F}}_v\big\} \\ \le\; \tfrac{1}{2}k^2V(s,v) \;\le\; \tfrac{1}{2}k^2W^*(v)e^{-\lambda v}. \end{aligned} \tag{3.76}$$

Thus, taking $k = \hat{c}_d e^u$ in (3.76), and on $\{W^*(v) \le \Lambda^\gamma\}$, it follows that

$$\begin{aligned}\mathbb{E}\Big\{\big|e^{\lambda v/2}\{\ell(\log[\hat{c}_dW(s,v)] + u) - \ell(\log[\hat{c}_dW^*(v)] + u)\} \\ - K_2(u,v)e^{\lambda v/2}\{W(s,v) - W^*(v)\}\big| \,\Big|\, \widetilde{\mathcal{F}}_v\Big\} \\ \le\; \tfrac{1}{2}\hat{c}_d^2e^{2u}\Lambda^{\gamma-\alpha_1/2}, \end{aligned} \tag{3.77}$$

and the lemma follows because $\gamma' + \gamma < 3\gamma/2 \le \alpha_1/2$, from (3.72). ∎

We are now in a position to prove a central limit theorem, with an error bound expressed in terms of the bounded Wasserstein distance.

Theorem 3.10. *Suppose that $v = \alpha\lambda^{-1}\log\Lambda$ for $0 < \alpha < 2\min\{\gamma_g/d, \zeta(d)/(1+\zeta(d)\}$, where γ_g is as in* (1.4) *and $\zeta(d)$ as in* (1.11) *(so that $\zeta(d) = 1/2$ for $d \le 6$). Suppose that γ is as for* (3.72) *and $\gamma' = \frac{1}{2}\min\{\gamma/2, (\zeta(d) - \eta'), (\alpha_2 - \alpha)\}$, where η' and α_2 are as in* (3.74)*, with $\alpha_1 = \alpha$. Suppose that Λ is large enough that* (3.13) *is satisfied, and that $\Lambda^{4\alpha\zeta(d)/7} \ge (d+1)^3$ and $\alpha\log\Lambda > c_{1*}$, where c_{1*} is as in Theorem 2.8. Then, for any $u_1 < u_0 \in \mathbb{R}$, there exist constants $C(d,u_1,u_0)$ and $C'(d,u_1,u_0)$ and an event $E^*(v) \in \sigma(\widetilde{H}(v))$ with $\mathbb{P}[E^*(v)^c] \le C'(d,u_1,u_0)\Lambda^{-\gamma'}$ such that*

$$\begin{aligned}d_{\mathrm{BW}}\big(\mathcal{L}\{e^{\lambda v/2}((L_{t_\Lambda(u)}/L) - \ell(\log[\hat{c}_dW^*(v)] + u)) \,|\, \widetilde{\mathcal{F}}_v \cap E^*(v)\}, \\ \mathcal{N}(0, \{K_2(u,v)\}^2W^*(v)/(d+1))\big) \\ \le\; C(d,u_1,u_0)\Lambda^{-\gamma'},\end{aligned}$$

uniformly in $u_1 \le u \le u_0$, where $K_2(u,v)$ is defined in (3.75)*, $\hat{c}_d$ in Lemma 3.6 and $t_\Lambda(u)$ in* (3.25)*.*

Proof. In view of (3.73) and Lemma 3.9, it suffices to show that

$$d_{\mathrm{BW}}\big(\mathcal{L}(e^{\lambda v/2}\{W(s,v) - W^*(v)\} \,|\, \widetilde{\mathcal{F}}_v), \mathcal{N}(0, W^*(v)/(d+1))\big) \;\le\; C_1(d,u_1,u_0)\Lambda^{-\gamma'},$$

with $s = \alpha_2\lambda^{-1}\log\Lambda$ and α_2 as in (3.74). Corollary 2.10, with $\eta = 6\zeta(d)/7$, shows that there is an event $E^\eta(v) \in \widetilde{H}(v)$ with $\mathbb{P}[\{E^\eta(v)\}^c] \le C'(d)\Lambda^{-2\alpha\zeta(d)/7}$ such that, on $E^\eta(v)$,

$$\begin{aligned}d_{\mathrm{BW}}\big(\mathcal{L}(e^{\lambda v/2}\{W(s,v) - W^*(v)\} \,|\, \widetilde{\mathcal{F}}_v), \mathcal{L}(e^{\lambda v/2}X_v^{(3)}(t_v^{-1}(s-v) \,|\, \widetilde{\mathcal{F}}_v)\big) \\ \le\; C(d)\{L^{-\alpha\zeta(d)/7}\},\end{aligned}$$

provided that $\Lambda^{4\alpha\zeta(d)/7} \geq (d+1)^3$. Then, from (2.25) and (2.36),

$$\mathcal{L}\big(e^{\lambda v/2} X_v^{(3)}(t_v^{-1}(s-v))\big) \;=\; \mathcal{N}\Big(0, \frac{W^*(v)}{d+1}(1-e^{-\lambda(s-v)})\Big),$$

and the theorem follows because $d_{\mathrm{BW}}(\mathcal{N}(0,\sigma_1^2), \mathcal{N}(0,\sigma_2^2)) = O(|\sigma_1 - \sigma_2|)$ and

$$W^*(v)e^{-\lambda(s-v)} \;=\; W^*(v)\Lambda^{-(\alpha_2-\alpha)} \;\leq\; L^{\gamma'-(\alpha_2-\alpha)},$$

on $\{W^*(v) \leq \Lambda^{\gamma'}\}$, and $\gamma' \leq \frac{1}{2}(\alpha_2 - \alpha)$, from (3.72). ∎

This theorem is not quite the same as Theorem 1.1, because both mean and variance are expressed in terms of $W^*(v) = W(v,v)$, which, as is seen from its definition in (3.40), is not necessarily determined by knowledge of $\mathcal{L}_v$ alone, because all the birth times of $\overline{X}$ come into its definition. Instead, one can observe $\widehat{W}(v)$ as in (1.10). We now show that this is enough.

We construct a lower bound $\widehat{W}_-(v)$ for $\widehat{W}(v)$ by summing over the subset of the birth times $\widehat{J}_v \subset J_v$ in (1.10) that belong to $J_v \cap \widetilde{J}_v$, where J_v is defined in (3.5), and

$$\widetilde{J}_v \;:=\; \Big\{ j \geq 0 \colon \overline{P}_j \notin \bigcup_{\substack{l \in \overline{J}_v \\ l < j}} \mathcal{K}(\overline{P}_j, 2v) \Big\},$$

with $\overline{J}_v$ the birth times of $\overline{X}$ before v, defined in (3.3). These give rise to non-intersecting neighbourhoods at time v, though not necessarily to all such, and they form a subset more amenable to calculation. Then it is immediate from (1.4) that, for all Λ sufficiently large,

$$\begin{aligned} \mathbb{E}|\overline{J}_v \setminus \widetilde{J}_v| &\leq 2\mathbb{E}\{|\overline{J}_v|(|\overline{J}_v|-1)L^{-1}(2v)^d\nu\} \\ &\leq 2c_2 e^{2\lambda_+ v}\Lambda^{-1}(2\lambda v)^d, \end{aligned}$$

the final inequality following from (2.2). Then, using arguments analogous to those in Lemma 3.2, we have

$$\begin{aligned} \mathbb{E}|\overline{J}_v \setminus J_v| &\leq \mathbb{E}\Big\{\int_0^v L^{-1}\nu M_d^+(u)\, c_1 e^{\lambda_+(v-u)} M_0^+(du)\Big\} \\ &= \mathbb{E}\Big\{\rho\nu \int_0^v L^{-1}\nu (M_d^+(u))^2\, c_1 e^{\lambda_+(v-u)}\, du\Big\} \\ &\leq \rho\nu^2 L^{-1} c_1 (c_2 d! \lambda_+^{-d})^2 \int_0^v e^{\lambda_+(v+u)}\, du \;\leq\; C\Lambda^{-1}e^{2\lambda_+ v}. \end{aligned}$$

Hence, for $v \leq t_{\max}(\Lambda)$,

$$0 \;\leq\; \mathbb{E}\{\widehat{W}(v) - \widehat{W}_-(v)\} \;=\; O\left\{\Lambda^{-1}\sum_{l=0}^{d} e^{\lambda v}(\log\Lambda)^{d+l}\right\},$$

and, for $v = \alpha\lambda^{-1}\log\Lambda$, this is of order $O(\Lambda^{-1+\alpha}(\log\Lambda)^{2d})$. The most sensitive place where this enters is into $\ell(\log[\hat{c}_d W^*(v)] + u)$, when the difference has to be small relative to $\Lambda^{-\alpha/2}$, because of the factor $e^{\lambda v/2}$; but this is the case if $\alpha < 2/3$, as in the statement of the theorem, by Lemma 4.2. The conversion of $E^*(v)$ into an event that can be determined from $\mathcal{L}_v$ can be accomplished in similar fashion, by modifying the definitions of its constituent events in terms of $W_j(v)$, $0 \leq j \leq v$.

Appendix

We note here two technical lemmas that are used in the previous arguments. The first establishes a bound on the extreme fluctuations of an integral with respect to a compensated Poisson process.

Lemma 4.1. *Let $X(t) := \int_0^t F(u)\{Z(du) - du\}$, where Z is a Poisson process and the process F is predictable and a.s. bounded in modulus by the deterministic function G. Define $G_2(s,t) := \int_s^t \{G(u)\}^2\,du$ and $G^*(s,t) := \sup_{s\le u\le t} G(u)$. Then*

$$\mathbb{P}\Big[\sup_{t_1\le t\le t_2} |X(t) - X(t_1)| > a\Big] \;\le\; 2\exp\big\{-a^2/\{2eG_2(t_1,t_2)\}\big\},$$

for all $0 \le a \le eG_2(t_1,t_2)/G^*(t_1,t_2)$. If G is decreasing, we have

$$\mathbb{P}\Big[\sup_{t_1\le t\le t_2} |X(t) - X(t_1)| > a\Big] \;\le\; 2\exp\big\{-a^2/\{2e\{G(t_1)\}^2(t_2-t_1)\}\big\},$$

for all $0 \le a \le eG(t_1)(t_2 - t_1)$.

Proof. For any θ, the process

$$Y(t) \;:=\; \exp\left\{\theta X(t) - \int_0^t \{e^{\theta F(u)} - 1 - \theta F(u)\}\,du\right\}$$

is a supermartingale (van de Geer [8, p. 1795]), and stopping at a easily yields

$$\begin{aligned}\mathbb{P}\Big[\sup_{t_1\le t\le t_2} (X(t) - X(t_1)) > a\Big] \;&\le\; e^{-\theta a}\mathbb{E}\left\{\exp\left(\int_{t_1}^{t_2} \{e^{\theta F(u)} - 1 - \theta F(u)\}\,du\right)\right\}\\ &\le\; e^{-\theta a}\exp\left(\frac{e}{2}\theta^2 G_2(t_1,t_2)\,du\right),\end{aligned}$$

if $0 \le \theta G^*(t_1,t_2) \le 1$. The corresponding bound for $\inf_{t_1\le t\le t_2}(X(t) - X(t_1))$ is proved in analogous fashion. Now, if $a \le eG_2(t_1,t_2)$, choose $\theta = a/\{eG_2(t_1,t_2)\}$, giving the first conclusion of the lemma. The second follows by choosing $\theta = a/\{eG(t_1)^2(t_2-t_1)\}$. ∎

The second lemma establishes some smoothness of the function $\phi_s^1(\theta) := \mathbb{E}\{e^{-\theta W^1(s)}\}$.

Lemma 4.2. *With ϕ_s^1 defined as above, and for any $s,h,\theta > 0$, we have*

$$\begin{aligned}|\phi_{s+h}^1(\theta) - \phi_s^1(\theta)| \;&\le\; \tfrac12\theta e^{-1}e^{-s}(1-e^{-h});\\ |\phi_s^1(\theta(1+\delta)) - \phi_s^1(\theta)| \;&\le\; \delta\min\{e^{-1},\theta\}.\end{aligned}$$

Proof. We note that $W^1(s) \ge 0$ and that $\mathbb{E}W^1(s) = 1$ for all s. Then, writing $X_s(h) := W^1(s+h) - W^1(s)$ and using (2.17), we have

$$\mathbb{E}\{X_s(h)\,|\,\widehat{\mathcal{F}}_s\} \;=\; 0; \qquad \mathbb{E}\{(X_s(h))^2\,|\,\widehat{\mathcal{F}}_s\} \;\le\; W^1(s)e^{-s}(1-e^{-h}), \tag{4.1}$$

for any $s,h > 0$. Hence, using (4.1), and taking expectations first conditional on $\widehat{\mathcal{F}}_s$, we have

$$\begin{aligned}\phi_{s+h}^1(\theta) - \phi_s^1(\theta) \;&=\; \mathbb{E}\big\{e^{-\theta W^1(s)}\big\{(e^{-\theta X_s(h)} - 1 + \theta X_s(h)) - \theta X_s(h)\big\}\big\}\\ &=\; \mathbb{E}\big\{e^{-\theta W^1(s)}\mathbb{E}\{(e^{-\theta X_s(h)} - 1 + \theta X_s(h))\,|\,\widehat{\mathcal{F}}_s\}\big\}.\end{aligned}$$

This implies that

$$|\phi_{s+h}^1(\theta) - \phi_s^1(\theta)| \;\le\; \tfrac12\mathbb{E}\{e^{-\theta W^1(s)}\theta^2 W^1(s)e^{-s}(1-e^{-h})\} \;\le\; \frac{\theta}{2e}e^{-s}(1-e^{-h}),$$

since $xe^{-x} \le e^{-1}$, proving the first inequality.

For the second, since $e^{-x}(1-e^{-\delta x}) \le \delta e^{-1}$ in $x \ge 0$ and $\mathbb{E}W^1(s) = 1$,

$$|\phi_s^1(\theta(1+\delta)) - \phi_s^1(\theta)| \;=\; |\mathbb{E}\{e^{-\theta W^1(s)}(1-e^{-\theta\delta W^1(s)})\}| \;\le\; \delta\min\{e^{-1},\theta\}.$$ ∎

References

[1] D. J. Aldous (2012) When knowing early matters: gossip, percolation and Nash equilibria. In: *Prokhorov and Contemporary Probability Theory*, Eds. A.N. Shiryaev, S.R.S. Varadhan and E.L. Presman, pp. 3–28. Springer Proceedings in Mathematics & Statistics **33**, Springer Verlag, Heidelberg. MR3070464

[2] S. Asmussen & H. Hering (1983) *Branching processes.* Progress in probability and statistics 3, Birkhäuser, Boston. MR0701538

[3] F. G. Ball (1983) The threshold behaviour of epidemic models. *J. Appl. Probab.* **20**, 227–241. MR0698527

[4] A. D. Barbour, L. Holst & S. Janson (1992) *Poisson approximation*. Oxford University Press. MR1163825

[5] A. D. Barbour & G. Reinert (2013). Asymptotic behaviour of gossip processes and small world networks. *Adv. Appl. Probab.* **45**, 981–1010. MR3161293

[6] S. Chatterjee & R. Durrett (2011) Asymptotic behavior of Aldous' gossip process. *Ann. Appl. Probab.* **21**, 2447–2482. MR2895421

[7] E. Z. Ganuza & S. D. Durham (1974) Mean–square and almost–sure convergence of supercritical age–dependent branching processes. *J. Appl. Probab.* **11**, 678–686. MR0359038

[8] S. van de Geer (1995) Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes. *Ann. Statist.* **23**, 1779–1801. MR1370307

[9] P. Jagers (1975) *Branching Processes with Biological Applications.* John Wiley & Sons, London. MR0488341

[10] J. Komlós, P. Major & G. Tusnády (1975) An approximation of partial sums of independent RV'-s, and the sample DF. I *Z. Wahrscheinlichkeitstheorie verw. Geb.* **32**, 111–131. MR0375412

[11] T. G. Kurtz (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* **7**, 49–58. MR0254917

[12] T. G. Kurtz (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Prob.* **8**, 344–356. MR0287609

[13] S. Li (2011) Concise formulas for the area and volume of a hyperspherical cap. *Asian J. Math. Stat.* **4**, 66–70. MR2813331

Acknowledgments. ADB thanks the Department of Statistics and Applied Probability at the National University of Singapore, and the mathematics departments of the University of Melbourne and Monash University, for their kind hospitality while much of the work was undertaken. AR thanks the School of Mathematics and Statistics at the University of Melbourne for their kind hospitality.