Electronic Journal of Probability

Electron. J. Probab. **20** (2015), no. 75, 1–36.
ISSN: 1083-6489 DOI: 10.1214/EJP.v20-3812

On the scaling limits of Galton–Watson processes in varying environments

Vincent Bansaye* Florian Simatos†

Abstract

We establish a general sufficient condition for a sequence of Galton–Watson branching processes in varying environments to converge weakly. This condition extends previous results by allowing offspring distributions to have infinite variance.
Our assumptions are stated in terms of pointwise convergence of a triplet of two real-valued functions and a measure. The limiting process is characterized by a backwards integro-differential equation satisfied by its Laplace exponent, which generalizes the branching equation satisfied by continuous state branching processes. Several examples are discussed, namely branching processes in random environment, Feller diffusion in varying environments and branching processes with catastrophes.

Keywords: Galton–Watson branching processes ; scaling limits ; varying environments.
AMS MSC 2010: 60J80; 60F17; 60K37.
Submitted to EJP on September 19, 2014, final version accepted on July 1, 2015.
Supersedes arXiv:1112.2547.

1 Introduction

For each $n \geq 1$, consider a sequence of offspring distributions $(q_{i,n}, i \geq 0)$, the *environments*, and the corresponding Galton–Watson process $Z_n = (Z_{i,n}, i \geq 0)$ where individuals of the i-th generation reproduce according to $q_{i,n}$. We are interested in the weak convergence of the sequence $(X_n, n \geq 1)$ of scaled processes of the form $X_n(t) = n^{-1} Z_{\gamma_n(t),n}$ for some sequence of time-changes γ_n.

In the Galton–Watson case where $q_{i,n} = q_{0,n}$, this problem has been first considered by Feller [19] and Kolmogorov [33] and later by Lindvall [39] and Lamperti [37, 38]. It was exhaustively solved by Grimvall [24] who provided a necessary and sufficient condition for the convergence of the scaled processes $(X_n, n \geq 1)$ in terms of the sequence of offspring distributions $(q_{0,n}, n \geq 1)$. Our approach is inspired by these works and relies on the convergence of the Laplace transform.

In the Galton–Watson case, the possible limit processes are called continuous state branching processes (CSBP) and were first considered by Jiřina [29]. This class of

*Ecole Polytechnique, France. E-mail: vincent.bansaye@polytechnique.edu
†Institut Supérieur de l’Aéronautique et de l’Espace (ISAE-SUPAERO), Université de Toulouse, France.
E-mail: florian.simatos@isae.fr

processes is well understood thanks to a random time-change transformation exhibited by Lamperti [36]. This transformation also allows for an elegant and conceptual proof of Grimvall's result, see for instance Ethier and Kurtz [18, Chapter 9] or, in the continuous-time setting, Helland [26], but this approach breaks down in the case of varying environment.

In this case, no such exhaustive result is available. Kurtz [35] and Borovkov [13] proved general results in the finite variance case, to which our main result will be compared in details in Section 2.4. The two main points are that: (1) on the upside, we extend these results to the case of offspring distributions with (possibly) infinite variance; (2) on the downside, we assume that a certain function has locally finite variation: as will be seen, this finite variation assumption is intrinsic to our approach.

When offspring distributions vary but have finite variance, the authors in [13, 35] express their limit process as a simple transformation of Feller diffusion, the only CSBP with continuous sample paths. Kurtz [35] for instance uses semigroup techniques developed in [34]. However, these techniques become significantly more demanding in the infinite variance case considered here, where one needs to consider diffusion processes with jumps.

For this reason, we use in this paper a variation of the approach developed in [19, 33, 37, 38, 39]: namely, our main object of investigation is the Laplace exponent u_n of X_n, defined by $u_n(s,t,\lambda) = -\log \mathbb{E}(\exp(-\lambda X_n(t)) \mid X_n(s) = 1)$ for $0 \leq s \leq t$ and $\lambda \geq 0$. We identify a triplet $(\alpha_n, \beta_n, \nu_n)$ with α_n and β_n two real-valued càdlàg functions, respectively of bounded variation and non-decreasing, and ν_n a σ-finite measure on $(0,\infty)^2$, such that informal calculation suggests the approximation

$$u_n(s,t,\lambda) \approx \lambda + \int_{(s,t]} u_n(y,t,\lambda)\alpha_n(dy) - \int_{(s,t]} (u_n(y,t,\lambda))^2 \beta_n(dy) + \int_{(0,\infty)\times(s,t]} h(x, u_n(y,t,\lambda))\nu_n(dx\,dy) \quad (1.1)$$

with $h(x,\lambda) = 1 - e^{-\lambda x} - \frac{\lambda x}{1+x^2} + \frac{(\lambda x)^2}{2(1+x^2)}$ (see Section 3.2). Motivated by this observation, we identify a mild notion of convergence $(\alpha_n, \beta_n, \nu_n) \to (\alpha, \beta, \nu)$ (see Assumption 2.1 below) under which (X_n) converges weakly. Its weak limit is then characterized by its Laplace exponent, which is shown to be the unique solution to the integro-differential equation obtained by letting $n \to \infty$ in (1.1). As alluded to above, in order for this equation to make sense in the limit, we need to assume that α has finite variation: otherwise, it is not clear how to make sense of the integral with respect to α in (1.1). Finally, note that this equation generalizes the branching equation for the Laplace exponent of CSBP obtained by Silverstein [41], see also Caballero et al. [14] for a recent and complete treatment.

In this paper, we assume the convergence $(\alpha_n, \beta_n, \nu_n) \to (\alpha, \beta, \nu)$, where the measure ν is allowed to be non-zero and satisfies classical $1 \wedge x^2$ finite moment. In the Galton–Watson case, our assumption is equivalent to Grimvall's necessary and sufficient condition [24]. But when environments are allowed to vary over time, we need to avoid times where the process goes to zero instantaneously and almost surely, which will be called *bottleneck*. Indeed, in the case of varying environments one must in general allow for non-critical distributions, and in particular subcritical ones which may cause such bottlenecks. Within our approach, we cannot determine in general the behavior of the process at the time of a bottleneck for reasons discussed in details in Section 2.3, where an example of indetermination $\infty \times 0$ is exhibited. Consequently, our main result studies the process (X_n) on some time interval $[\wp(t), t]$ where $0 \leq \wp(t) \leq t$ intuitively corresponds to the last bottleneck before time t. Moreover, we show in Section 2.3 that, within our assumption $(\alpha_n, \beta_n, \nu_n) \to (\alpha, \beta, \nu)$, a bottleneck can only occur if there is an offspring distribution with mean close to 0: otherwise, we have $\wp(t) = 0$, see Proposition 2.3. Finally, we study in Proposition 2.4 the behavior of the process at the time of a bottleneck in the non-explosive case, where an indetermination $\infty \times 0$ is proscribed.

Let us now mention some closely related results. Galton–Watson processes in *random* environment were first introduced and studied in Smith and Wilkinson [42] in the case where the sequence $(q_{i,n}, i \geq 0)$ is i.i.d., and in Athreya and Karlin [4, 3] when this sequence is stationary. These models have recently attracted considerable interest in the literature, see for instance [1, 2, 5, 10, 11, 12, 21, 25] for results on the long-time behavior in the critical and subcritical regimes and on large deviation. Scaling limits in the finite variance case were conjectured by Keiding [31] who introduced Feller diffusion in random environment. This conjecture was proved by Kurtz [35] and Helland [27]. In the same way, our results describe the weak convergence of scaled processes conditionally on the environments (quenched results), when the offspring distributions may have infinite variance. We describe the probabilistic structure of this process in Section 2.5.1 and we shed light on the correct scaling of such processes.

Our results are also related to some results on superprocesses. More precisely, our limit processes are closely related to the mass of superprocesses considered in El Karoui and Roelly [17]. These superprocesses are obtained in Dynkin [15, 16] as the limit of suitable branching particle systems, under some additional assumptions, e.g. finite first moment (conservative case) and no drift. In these works the emphasis is on the limiting superprocesses themselves. As such, Dynkin [15, 16] considers branching particle systems evolving in continuous time, which, in order to establish limit theorems, are technically more convenient than the discrete time setting, which is our motivation here.

Organization of the paper

Theorem 2.2 is the main result of the paper, and is presented in Section 2.2. We compare it with earlier results in Section 2.4 and discuss some applications in Section 2.5, namely to Galton–Watson processes in random environment, to Feller diffusion in varying environments and to CSBP with catastrophes. Section 3 introduces notation, as well as some preliminary results. Theorem 2.2 is proved in Section 4, with some technical proofs deferred to Appendices B and C. Further results that complement Theorem 2.2 are proved in Section 5, and Appendix A is devoted to checking that the assumptions of Theorem 2.2 are necessary and sufficient in the Galton–Watson case.

2 Notation and results

2.1 General notation

In the rest of the paper, if a function g defined on $[0, \infty)$ is càdlàg, we write $\Delta g(t) = g(t) - g(t-)$ for the value of the jump of g at time t. If g is in addition of locally finite variation, we write $\|g\|(t)$ for the total variation of g on $[0, t]$ and $\int f dg$ for the Lebesgue-Stieltjes integral of a measurable function f; note that $|\int f dg| \leq \int |f| d\|g\|$. Moreover, we say that a function f is increasing if $f(x) \geq f(y)$ for every $x \geq y$.

For each $n \geq 1$, we consider a Galton–Watson process in varying environments $Z_n = (Z_{i,n}, i \geq 0)$. We denote by $q_{i,n}$ the offspring distribution in generation i and $\xi_{i,n}$ a random variable distributed according to $q_{i,n}$, so that we can construct Z_n according to the following recursion:

$$Z_{i+1,n} = \sum_{k=1}^{Z_{i,n}} \xi_{i,n}(k), \ i \geq 0,$$

where the random variables $(\xi_{i,n}(k), i, k \geq 0)$ are independent and $\xi_{i,n}(k)$ is equal in distribution to $\xi_{i,n}$. In order to find an interesting scaling of the sequence of processes $(Z_n, n \geq 1)$, the space scale is equal to n while the time scale is allowed to vary over time. More precisely, for $n \geq 1$, we consider an increasing, càdlàg and onto function $\gamma_n : [0,\infty) \to \mathbb{N}$ (here and elsewhere, $\mathbb{N} = \{0, 1, \dots\}$ denotes the set of non-negative integers) and we define the scaled process $(X_n(t), t \geq 0)$ as follows:

$$X_n(t) = \frac{1}{n} Z_{\gamma_n(t),n}, \ t \geq 0.$$

For $i \geq 0$ and $n \geq 1$, we define $t_i^n = \inf\{t \geq 0 : \gamma_n(t) = i\}$ so that $\gamma_n(t_i^n) = i$ and $t^n_{\gamma_n(t)} \leq t < t^n_{\gamma_n(t)+1}$. Since Z_n satisfies the branching property, i.e., Z_n started from $Z_{0,n} = z$ is stochastically equivalent to the sum of z i.i.d. processes distributed according to Z_n started from $Z_{0,n} = 1$, we obtain after scaling

$$\mathbb{E}\left[\exp\left(-\lambda X_n(t)\right) \mid X_n(s) = x\right] = \exp(-x u_n(s,t,\lambda)) \tag{2.1}$$

for all $\lambda, x, s, t \geq 0$ with $s \leq t$ and $x \in \mathbb{N}/n$ and $u_n(s,t,\lambda) \geq 0$ called the Laplace exponent. We will characterize the convergence of X_n through the convergence of u_n. Our assumptions are relying on the convergence of the triplet $(\alpha_n, \beta_n, \nu_n)$, where α_n and β_n are real-valued functions and ν_n is a measure on $\mathbb{R} \times [0,\infty)$. This triplet is defined in terms of the normalized random variables

$$\bar{\xi}_{i,n} = \frac{1}{n}\left(\xi_{i,n} - 1\right), \ i \geq 0, n \geq 1,$$

in the following way. We introduce the numbers

$$\alpha_{i,n} = n\mathbb{E}\left(\frac{\bar{\xi}_{i,n}}{1+\bar{\xi}_{i,n}^2}\right) \ \text{ and } \ \beta_{i,n} = \frac{1}{2} n\mathbb{E}\left(\frac{\bar{\xi}_{i,n}^2}{1+\bar{\xi}_{i,n}^2}\right)$$

and the measures

$$\nu_{i,n}([x,\infty)) = n\mathbb{P}(\bar{\xi}_{i,n} \geq x),$$

so that the triplet can now be defined for $t \geq 0$ by

$$\alpha_n(t) = \sum_{i=0}^{\gamma_n(t)-1} \alpha_{i,n} = \int \frac{x}{1+x^2} \nu_{i,n}(dx), \quad \beta_n(t) = \sum_{i=0}^{\gamma_n(t)-1} \beta_{i,n} = \frac{1}{2}\int \frac{x^2}{1+x^2} \nu_{i,n}(dx)$$

and

$$\nu_n([x,\infty) \times (0,t]) = \sum_{i=0}^{\gamma_n(t)-1} \nu_{i,n}([x,\infty)) \qquad (x \geq 0).$$

Each time, we understand a sum of the form $\sum_0^{-1}$ to be equal to 0, so that $\alpha_n(0) = \beta_n(0) = 0$.

2.2 Main result

Before stating our main result, we first precisely state the assumption $(\alpha_n, \beta_n, \nu_n) \to (\alpha, \beta, \nu)$ alluded to in the introduction. We also give a definition of $\wp(t)$, to be thought of the last bottleneck before time t (see the introduction).

Assumption 2.1. *There exist a càdlàg function of locally finite variation α, an increasing càdlàg function β, and a positive measure ν on $(0,\infty)^2$, such that the two following conditions hold:*

(A1) *For every $t \geq 0$ and every $x > 0$ such that $\nu(\{x\} \times (0,t]) = 0$,*

$$\alpha_n(t) \underset{n\to\infty}{\longrightarrow} \alpha(t),\ \|\alpha_n\|(t) \underset{n\to\infty}{\longrightarrow} \|\alpha\|(t),\ \beta_n(t) \underset{n\to\infty}{\longrightarrow} \beta(t)$$
$$\text{and } \nu_n([x,\infty) \times (0,t]) \underset{n\to\infty}{\longrightarrow} \nu([x,\infty) \times (0,t]).$$

(A2) *For every t such that $\Delta\alpha(t) \neq 0$, $\Delta\beta(t) \neq 0$ or $\nu((0,\infty) \times \{t\}) \neq 0$ and for every $x > 0$ such that $\nu(\{x\} \times \{t\}) = 0$,*

$$\alpha_{\gamma_n(t),n} \underset{n\to\infty}{\longrightarrow} \Delta\alpha(t),\ \beta_{\gamma_n(t),n} \underset{n\to\infty}{\longrightarrow} \Delta\beta(t) \text{ and } \nu_{\gamma_n(t),n}([x,\infty)) \underset{n\to\infty}{\longrightarrow} \nu([x,\infty) \times \{t\}).$$

The following definition of $\wp(t)$ is the most technically convenient and general at this point:

$$\wp(t) = \sup\left\{s \leq t : \lim_{\varepsilon\to 0} \liminf_{n\to\infty} \inf_{s\leq y\leq t} \mathbb{P}\left(X_n(t) > \varepsilon \;\middle|\; X_n(y) = 1\right) = 0\right\} \tag{2.2}$$

with the convention $\sup\emptyset = 0$. A more intuitive definition will be given in Lemma 3.4, while a sufficient condition for $\wp(t) = 0$ under Assumption 2.1 is given in Proposition 2.3.

Theorem 2.2 (Behavior on $[\wp(t), t]$)**.** *Assume that Assumption 2.1 holds, and let α, β and ν the functions and measure defined there. Then, the following properties hold.*

I. *For every $t \geq 0$, we have $\Delta\alpha(t) \geq -1$ and $\int_{(0,\infty)\times(0,t]} (1 \wedge x^2)\nu(dx\,dy) < \infty$. Moreover, the following function $\widetilde{\beta}$ is continuous and increasing:*

$$\widetilde{\beta}(t) = \beta(t) - \int_{(0,\infty)\times(0,t]} \frac{x^2}{2(1+x^2)}\nu(dx\,dy),\ t \geq 0.$$

II. *For every $t, \lambda > 0$ and $s \in [\wp(t), t]$, there exists $u(s,t,\lambda) \in (0,\infty)$ such that for every $s_0 \geq 0$,*

$$\lim_{n\to\infty} \sup_{0\leq s\leq s_0} |u_n(s,t,\lambda) - u(s,t,\lambda)| = 0.$$

Moreover, the function $u_{t,\lambda} : s \in [\wp(t), t] \mapsto u(s,t,\lambda)$ is the unique càdlàg function that satisfies $\inf_{s\leq y\leq t} u_{t,\lambda}(y) > 0$ for every $\wp(t) < s \leq t$ and

$$u_{t,\lambda}(s) = \lambda + \int_{(s,t]} u_{t,\lambda}(y)\alpha(dy) - \int_{(s,t]} u_{t,\lambda}(y)^2 \widetilde{\beta}(dy) + \int_{(0,\infty)\times(s,t]} \left(1 - e^{-xu_{t,\lambda}(y)} - \frac{xu_{t,\lambda}(y)}{1+x^2}\right)\nu(dx\,dy) \tag{2.3}$$

for every $\wp(t) \leq s \leq t$.

III. *Fix $t \geq 0$, $s \in [\wp(t), t]$ and $x \geq 0$. Then for every sequence of initial states (x_n) with $x_n \to x$, every $I \in \mathbb{N} - \{0\}$, every $s \leq t_1 < \cdots < t_I \leq t$ and every $\lambda_1, \ldots, \lambda_I > 0$,*

$$\lim_{n\to\infty} \mathbb{E}\left[\exp\left(-\lambda_1 X_n(t_1) - \cdots - \lambda_I X_n(t_I)\right) \mid X_n(s) = x_n\right] = \exp\Big(-xu\big(s, t_1, \lambda_1 + u\big(t_1, t_2, \lambda_2 + u(\cdots, u(t_{I-1}, t_I, \lambda_I)\cdots)\big)\big)\Big). \tag{2.4}$$

IV. *Fix* $t \geq 0$, $s \in [\wp(t), t]$ *and* $x \geq 0$. *Then for every sequence of initial states* (x_n) *with* $x_n \to x$, *the sequence of processes* $(X_n(y), s \leq y \leq t)$ *under* $\mathbb{P}(\cdot \mid X_n(s) = x_n)$ *is tight on the space* $D([s,t],[0,\infty])$ *of càdlàg functions* $f : [s,t] \to [0,\infty]$ *endowed with the* J_1 *topology, where the space* $[0,\infty]$ *is equipped with the metric* $d(x,y) = |e^{-x} - e^{-y}|$. *In particular, weak convergence holds in view of* (2.4).

In claim IV we consider X_n as a process with range $[0,\infty]$. Although X_n for fixed n cannot explode, for technical reasons we need to specify its behavior started at ∞: in the sequel we assume that ∞ is an absorbing state, so that if $X_n(s) = \infty$ for some s, then $X_n(t) = \infty$ for all $t \geq s$. The proof of claim IV will actually show that $(X_n(y), s \leq y \leq t)$ under $\mathbb{P}(\cdot \mid X_n(s) = x_n)$ is tight for any $0 \leq s \leq t$, not only $\wp(t) \leq s \leq t$.

2.3 Around the bottleneck

We now discuss in more details the notion of bottleneck that we have introduced. As a first guess, we could expect that the process goes to 0 when going through the time of a bottleneck, which would mean that $u_n(s,t,\lambda) \to 0$ if $s < \wp(t)$. We now consider an example which illustrates several things that can go wrong and justify our framework, in particular the definition of the bottleneck $\wp(t)$ and the fact that Theorem 2.1 is stated on $[\wp(t), t]$.

Consider a critical offspring distribution q and $Y_n = (Y_{i,n}, i \geq 0)$ the Galton–Watson process with offspring distribution q, started from n individuals. Assume that q and Γ_n are such that the sequence $(\widehat{Y}_n)$ with $\widehat{Y}_n(t) = Y_{\lfloor \Gamma_n t \rfloor, n}/n$ converges weakly to a non-conservative CSBP $\widehat{Y}$. For each $n \geq 1$, we define Z_n by $Z_{0,n} = n$ and (δ_k denotes the unit mass at $k \in \mathbb{N}$):

$$q_{i,n} = \begin{cases} q & \text{if } 0 < i < \Gamma_n, \\ \delta_1 & \text{if } \Gamma_n \leq i < 2\Gamma_n, \\ (1-p_n)\delta_0 + p_n\delta_1 & \text{if } i = 2\Gamma_n, \\ q & \text{if } i > 2\Gamma_n \end{cases}$$

for some vanishing sequence $p_n \in [0,1]$. Defining $\gamma_n(t) = \lfloor \Gamma_n t \rfloor$ and recalling $X_n(t) = Z_{\gamma_n(t),n}/n$, we see that X_n coincides (in distribution) with Y_n on $[0,1)$, stays constant on $[1,2)$, undergoes a highly subcritical offspring distribution with mean p_n at time 2, referred to as catastrophe, and then resumes evolving according to q after time 2. The catastrophe at time 2 is meant to correspond to a bottleneck, and indeed one can check that $\wp(t) = 0$ if $t < 2$ and $\wp(t) = 2$ if $t \geq 2$. Moreover, since X_n shifted at time 2 is a rescaled Galton–Watson process, the discussion on Galton–Watson processes in the next section will show that Assumption 2.1 is satisfied.

Fix some $t > 2$, and let us now discuss the asymptotic behavior of $u_n(s,t,\lambda)$ for $s < \wp(t) = 2$. First of all, although $p_n \to 0$ suggests that the process gets extinct at time 2, $u_n(s,t,\lambda)$ may actually not converge to 0. Indeed, since $\widehat{Y}_n$ converges weakly to $\widehat{Y}$ and $\widehat{Y}$ is not conservative, there exist $\rho > 0$ and a sequence $y_n \to \infty$ such that $\mathbb{P}(\widehat{Y}_n(2-) \geq y_n) \geq \rho$. In particular, just before the catastrophe X_n is, with probability at least ρ, at least of the order of y_n. In this event, the catastrophe brings X_n to level $p_n y_n$ by the law of large numbers, which diverges if $p_n \gg 1/y_n$, i.e., if p_n vanishes slowly enough. This argument could be made rigorous to show that $u_n(s,t,\lambda)$ does not go to 0 for $s < 1$. Secondly, even if $u_n(s,t,\lambda)$ converges the limit may depend on s: for $s < 1$ we have seen that the limit was > 0, while for $1 \leq s < 2$ the limit is $= 0$. Finally, $u_n(s,t,\lambda)$ may even fail to converge: to see this, one may for instance consider two sequences $p_n^{(1)}$ and $p_n^{(2)}$ with $y_n p_n^{(1)} \to \infty$ and $y_n p_n^{(2)} \to 0$, $X_n^{(1)}$ and $X_n^{(2)}$ the two processes obtained by the above construction using $p_n^{(1)}$ and $p_n^{(2)}$ instead of p_n, respectively, and finally intertwine them by considering $X_{2n} = X_n^{(1)}$ and $X_{2n+1} = X_n^{(2)}$.

This example therefore shows that a wide variety of behavior can happen before the bottleneck. We now give a sufficient condition that ensures $\wp(t) = 0$, i.e., that there is no bottleneck before time t. Intuitively, the following assumption ensures that $\xi_{i,n}$ is not too close to 0, which avoids the almost sure absorption in one generation. For instance, it prevents the catastrophe of the previous example at time 2.

Proposition 2.3 (No bottleneck)**.** *Let $t > 0$. If for every $C > 0$*

$$\liminf_{n\to\infty} \left(\inf_{0\le i\le\gamma_n(t)} \mathbb{E}\left(\xi_{i,n}; \xi_{i,n} \le Cn\right) \right) > 0, \tag{2.5}$$

then $\wp(t) = 0$.

The proof of this proposition and of the next one are deferred to Section 5. We conclude this discussion by giving a condition under which $u_n(s, t, \lambda) \to 0$ along a subsequence, for all $s < \wp(t)$ and $\lambda \ge 0$. Then the process started at the bottleneck goes as expected to zero (along a subsequence) when going through the bottleneck. Note that the example given at the beginning of this section shows that this is not always the case. Roughly speaking, this condition means that the limiting process is conservative.

Proposition 2.4 (No explosion)**.** *Fix some $t > 0$. If the two sequences $(\|\alpha_n\|(t), n \ge 1)$ and $(\beta_n(t), n \ge 1)$ are bounded and*

$$\lim_{A\to\infty} \sup_{n\ge1,\, 0\le s\le y\le t} \mathbb{P}(X_n(y) \ge A \mid X_n(s) = 1) = 0, \tag{2.6}$$

then there exists an increasing sequence of integers $n(k)$ such that for all $s < \wp(t)$ and $\lambda \ge 0$, $u_{n(k)}(s, t, \lambda) \to 0$ as $k \to \infty$.

Moreover, these assumptions are satisfied, i.e., $(\|\alpha_n\|(t), n \ge 1)$ and $(\beta_n(t), n \ge 1)$ are bounded and (2.6) *holds, if the following first moment condition is satisfied:*

$$\sup_{n\ge1} \left(n \sum_{i=0}^{\gamma_n(t)-1} \mathbb{E}\big(|\bar{\xi}_{i,n}|\big) \right) < \infty. \tag{2.7}$$

2.4 Comparison with earlier work

In the Galton–Watson case where $\gamma_n(t) = \lfloor \Gamma_n t \rfloor$ for some integer-valued sequence (Γ_n) and $q_{i,n} = q_{0,n}$ for every $i \ge 0$, necessary and sufficient conditions for the finite-dimensional and weak convergence of $(X_n, n \ge 1)$ are known since Grimvall [24], where weak convergence in the space $D([0,\infty), [0,\infty))$ is considered. The main condition there is the weak convergence of the sequence $(S_n, n \ge 1)$ where S_n is distributed as the sum of $n\Gamma_n$ independent copies of $\bar{\xi}_{0,n}$. Indeed, Theorem 1.4 in Ethier and Kurtz [18] shows that the weak convergence of (S_n) implies the weak convergence of (X_n) in $D([0,\infty), [0,\infty])$, and Grimvall [24] proved the converse provided (X_n) converges to a non-explosive process (and thus in $D([0,\infty), [0,\infty))$; Helland [26] proved similar results for continuous-time branching processes. The following result, proved in the Appendix A, therefore shows that our Assumption 2.1 is sharp in the Galton–Watson case. Note in particular that in the Galton–Watson case, $q_{0,n}$ must be near-critical which implies in view of Proposition 2.3 that $\wp(t) = 0$.

Lemma 2.5. *In the Galton–Watson case, Assumption 2.1 is equivalent to the weak convergence of (S_n).*

In the case of Galton–Watson processes in varying environments, Kurtz [35] used semigroup techniques to study the case where offspring distributions have uniformly bounded third moments, which was later weakened by Borovkov [13] to a $2+\delta$ moment condition. There are three main differences between the assumptions made in [13, 35] and our Assumption 2.1.

First, we do not need to assume uniformly bounded second moments, which make appear new phenomena such as possible indetermination form at the bottleneck (discussed in Section 2.3) and issues related to the correct time scale of Galton–Watson processes in random environment (discussed in Section 2.5.1).

Second, as we already mentioned in the introduction, the function that in [13, 35] essentially plays the role of our α_n is not assumed to have finite variation in [13, 35]. This finite variation assumption is natural in our approach: otherwise it is not clear what meaning should be given to the term $\int_{(s,t]} u_{t,\lambda}(y)\alpha(dy)$ in (2.3). An enticing approach would be to consider α with finite quadratic variations, which would for instance make it possible to use a pathwise construction of Itô's integral such as in Föllmer [20], see also Wong and Zakai [44].

Finally, the functions that in [13, 35] essentially play the role of our α_n and β_n are assumed in [13, 35] to converge in the J_1 topology, whereas here we only assume pointwise convergence.

2.5 Applications

We discuss in this section new results that stem from Theorem 2.2. We keep the discussion at a high level and reserve rigorous results for future work (with the exception of Proposition 2.7).

2.5.1 Scaling limits of Galton–Watson processes in random (i.i.d.) environment

Consider for each $n \geq 1$ a random probability measure Q_n and a sequence $(q_{i,n}, i \geq 0)$ of i.i.d. random variables distributed as Q_n. Then the sequence $((\alpha_{i,n}, \beta_{i,n}, \nu_{i,n}), i \geq 0)$ is an i.i.d. sequence of $\mathbb{R} \times [0,\infty) \times \mathcal{M}$-valued random variables, with $\mathcal{M}$ the space of locally finite measures on $\mathbb{R}$ endowed with the vague topology. Because of the law of large numbers, it is natural to choose γ_n linear in t, i.e., $\gamma_n(t) = \lfloor \Gamma_n t \rfloor$ for some sequence $\Gamma_n \to \infty$. We now discuss conditions under which Assumption 2.1 holds.

We are interested in the convergence of the process $Y_n(t) = (\alpha_n(t), \beta_n(t), M_n(t))$, where $M_n(t) = \sum_{0 \leq i < \gamma_n(t)} \nu_{i,n}$ defines a measure-valued process. Note that the process Y_n has i.i.d. increments, and so we can use classical results on measure-valued processes and random walks, such as [28, Theorem VII.2.35] and [30, Theorem A2.4], to get an explicit condition for its convergence. Namely, a function $h : \mathbb{R}^3 \to \mathbb{R}^3$ is called truncation function if it is continuous, bounded and satisfies $h(x) = x$ in a neighborhood of 0 (in the sequel, vectors are considered to be columns and v' denotes the transposition). Let $\mathcal{C}_K^+$ be the set of non-negative continuous functions with compact support equipped with the uniform norm, and $\mathcal{C} \subset \mathcal{C}_K^+$ be a dense subset closed under addition. For $\varphi \in \mathcal{C}$, let $y_n^\varphi = (\alpha_{0,n}, \beta_{0,n}, \int \varphi d\nu_{0,n})$.

Condition 2.6. *There exist a truncation function h, F^φ a measure on $\mathbb{R}^3$ integrating $1 \wedge |x|^2$ and $b^\varphi \in \mathbb{R}^3$, $c_{ij}^\varphi \geq 0$ such that for every $\varphi \in \mathcal{C}$,*

$$\Gamma_n \mathbb{E}(h(y_n^\varphi)) \underset{n\to\infty}{\longrightarrow} b^\varphi, \ \Gamma_n \left\{ \mathbb{E}\big[h_i(y_n^\varphi)h_j(y_n^\varphi)\big] - \mathbb{E}(h_i(y_n^\varphi))\mathbb{E}(h_j(y_n^\varphi)) \right\} \underset{n\to\infty}{\longrightarrow} c_{ij}^\varphi$$

$$\text{and } \ \Gamma_n \mathbb{E}(g(y_n^\varphi)) \underset{n\to\infty}{\longrightarrow} \int g(x) F^\varphi(dx). \quad (2.8)$$

In the above, the second convergence holds for all $i, j = 1, 2, 3$ and the last convergence holds for all bounded, continuous functions g that are equal to 0 in a neighborhood of 0.

Assuming that this condition holds, it can be proved that Y_n converges to the process $Y(t) = (\alpha(t), \beta(t), M(t))$ such that for every φ continuous with compact support, the process $Y^\varphi = (\alpha, \beta, M^\varphi)$ with $M^\varphi = (\int \varphi(x) M(t)(dx), t \geq 0)$ is the Lévy process with Lévy exponent

$$\psi^\varphi(v) = iv' b^\varphi - \frac{1}{2} v' c^\varphi v + \int \left(e^{ivx} - 1 - iv' h(x)\right) F^\varphi(dx), \ v \in \mathbb{R}^3.$$

Further, using Skorohod's embedding theorem, we can assume that the convergence $Y_n \to Y$ holds almost surely. By definition, there exists a sequence of increasing bijections $(\lambda_n, n \geq 1)$ from $[0, \infty)$ to $[0, \infty)$ such that $\sup_{0 \leq s \leq t} |\lambda_n(s) - s| \to 0$ for every $t \geq 0$, and such that assumptions (A1) (except for the convergence of $\|\alpha_n\|$) and (A2) are satisfied for $\gamma_n' = \gamma_n \circ \lambda_n$ (see, e.g., Proposition VI.2.1 in Jacod and Shiryaev [28]).

Assuming now that α is of finite variation, α being a Lévy process must be of the form $\alpha(t) = \mathrm{d}_\alpha t + S_+(t) - S_-(t)$ where $\mathrm{d}_\alpha \in \mathbb{R}$ and S_+ and S_- are two independent pure-jump subordinators (see, e.g., Bertoin [8]). With this special structure, it is possible to prove that $\|\alpha_n\|(t) \to \|\alpha\|(t)$ so that Assumption 2.1 is fully satisfied and all the conclusions of Theorem 2.2 hold. It would be interesting to delve deeper into the probabilistic structure of the process (α, β, ν), and to understand how it relates to the properties of the limiting process X such as the extinction probability or the speed of extinction. In the literature, only the case of Feller diffusion in random environment where $\nu = 0$ and α is a Brownian motion has begun to be looked at, see, e.g., Böinghoff and Hutzenthaler [12].

We conclude this section by commenting on a question that actually motivated us in the first place: given a sequence of Galton–Watson processes in random environment, how can we find the right scaling in time, i.e., the right sequence (Γ_n)?
Let us focus on the simplest possible case where in each generation we choose at random among one of two possible offspring distributions, i.e., we can write $Q_n = p_n^{(1)} \delta_{q^{(1)}} + p_n^{(2)} \delta_{q^{(2)}}$ where $p_n^{(j)} \in (0, 1)$, $p_n^{(1)} + p_n^{(2)} = 1$ and $q^{(1)}$, $q^{(2)}$ are two offspring distributions. In this discussion, we will call a CSBP with characteristic (b, c, F) the CSBP whose branching mechanism is given by

$$\psi(\lambda) = \lambda b - \frac{1}{2} c \lambda^2 + \int (e^{-\lambda x} - 1 - \lambda x \mathbb{1}_{\{x \leq 1\}}) F(dx).$$

For each $j = 1, 2$ let $Z_n^{(j)} = (Z_n^{(j)}(i), i \geq 0)$ be a Galton–Watson process with offspring distribution $q^{(j)}$ and consider $(\Gamma_n^{(j)})$ a sequence such that $(X_n^{(j)}, n \geq 1)$ converges weakly to the CSBP with characteristic $(b^{(j)}, c^{(j)}, F^{(j)})$, where $X_n^{(j)}(t) = n^{-1} Z_n^{(j)}(\lfloor \Gamma_n^{(j)} t \rfloor)$.
If both $q^{(1)}$ and $q^{(2)}$ have finite variance, then it is well-known that in order to scale the Galton–Watson process with offspring distribution $q^{(i)}$ when the space scale is n, one needs to speed up time with n also, i.e., $\Gamma_n^{(1)} = \Gamma_n^{(2)} = n$. Thus when "mixing" these two processes, it is natural to speed up the resulting process by the common time scale and thus take $\Gamma_n = n$. To our knowledge, only such cases have been considered in the literature so far. When offspring distributions have infinite variance however, the situation becomes more delicate. Indeed, if for instance $q^{(1)}([x, \infty)) \sim x^{-a}$ as $x \to \infty$ for some $a \in (1, 2)$, then one needs to consider $\Gamma_n^{(1)} = n^{a-1}$. Thus there are now two "natural" time scales, namely $\Gamma_n^{(1)} = n^{a-1}$ and $\Gamma_n^{(2)} = n$.

Note that $\Gamma_n^{(j)}$ is the number of generations needed so that the variation of $Z_n^{(j)}$ may be of the order of n. Over Γ_n generations, the law of large numbers implies that $q^{(j)}$

has been used $p_n^{(j)}\Gamma_n$ times. Thus, if $p_n^{(j)}\Gamma_n \ll \Gamma_n^{(j)}$, the offspring distribution $q^{(j)}$ has not been picked sufficiently often in order to have any effect (on the space scale n). This suggests that the correct time scale is $\Gamma_n = \min_j(\Gamma_n^{(j)}/p_n^{(j)})$ and indeed, the following result can be proved using Theorem 2.2:

- if $\Gamma_n^{(1)}/p_n^{(1)} \ll \Gamma_n^{(2)}/p_n^{(2)}$ and $\Gamma_n = \Gamma_n^{(1)}/p_n^{(1)}$, then X_n converges toward the CSBP with branching mechanism $(b^{(1)}, c^{(1)}, F^{(1)})$;
- if $\Gamma_n^{(2)}/p_n^{(2)} \ll \Gamma_n^{(1)}/p_n^{(1)}$ and $\Gamma_n = \Gamma_n^{(2)}/p_n^{(2)}$, then X_n converges toward the CSBP with branching mechanism $(b^{(2)}, c^{(2)}, F^{(2)})$;
- if $\Gamma_n^{(1)}p_n^{(2)}/(\Gamma_n^{(2)}p_n^{(1)}) \to \ell \in (0,\infty)$ and $\Gamma_n = \Gamma_n^{(1)}/p_n^{(1)}$, then X_n converges toward the CSBP with characteristic $(b^{(1)} + \ell b^{(2)}, c^{(1)} + \ell c^{(2)}, F^{(1)} + \ell F^{(2)})$.

This discussion can be easily extended to the case of a finite number of offspring distributions that also vary with n, and it would be very interesting to understand the implications of Theorem 2.2 in more general settings, e.g., when we can choose among uncountably many offspring distributions.

2.5.2 Feller diffusion

Going back to the case of varying environments, the finite variance case is of particular interest. This is the only one that has been studied so far, see in particular [13, 35]. In this case, our approach via the generalized branching equation (2.3) makes it possible to derive an expression of the extinction probability. This extends results already known for linear birth and death branching processes in varying environments from [32] and for particular classes of CSBP in random environment from [6, 12].

Proposition 2.7. *Assume that Assumption 2.1 holds with $\nu = 0$. Then β is continuous and for all $t \geq 0$, we have*

$$u(s,t,\lambda) = \frac{\exp(-\overline{\alpha}(s))}{\lambda^{-1}\exp(-\overline{\alpha}(t)) + \int_{(s,t]}\exp(-\overline{\alpha}(y))\beta(dy)}, \ 0 \leq s \leq t, \lambda \geq 0, \tag{2.9}$$

where $\overline{\alpha}(t) = \alpha(t) + \sum_{0\leq s\leq t}[\log(1+\Delta\alpha(s)) - \Delta\alpha(s)]$. In particular, if $\wp(t) = 0$ for every $t \geq 0$ (for instance, if (2.5) holds), then for any $s \geq 0$ and $x \geq 0$

$$\lim_{t\to\infty} \mathbb{P}(X(t) = 0 \mid X(s) = x) = \exp\left(-\frac{x\exp(-\overline{\alpha}(s))}{\int_{(s,\infty)}\exp(-\overline{\alpha}(y))\beta(dy)}\right) \tag{2.10}$$

where X is the weak limit of the sequence of processes $(X_n(y), y \geq s)$ given by properties III and IV of Theorem 2.2.

Proof. Since Assumption 2.1 holds, all the conclusions of Theorem 2.2 hold. In particular, $\widetilde{\beta}$ is continuous and since $\nu = 0$ by assumption, $\beta = \widetilde{\beta}$ and β itself is continuous.

Let us now prove (2.9). Fix $t, \lambda > 0$: according to Theorem 2.2, it is enough to check that $G(s) = H(s)$, where $G(s)$ is equal to the right-hand side of (2.9) and $H(s) = \lambda + \int_{(s,t]} G d\alpha - \int_{(s,t]} G^2 d\beta$. Observe that G and H may only jump when α does. We first compare the jumps: since β is continuous, $s \mapsto \int_{(s,t]}\exp(-\overline{\alpha}(y))d\beta(y)$ is continuous and so

$$\Delta G(s) = \frac{\exp(-\overline{\alpha}(s)) - \exp(-\overline{\alpha}(s-))}{\lambda^{-1}\exp(-\overline{\alpha}(t)) + \int_{(s,t]}\exp(-\overline{\alpha}(y))\beta(dy)} = G(s)\left(1 - e^{\Delta\overline{\alpha}(s)}\right).$$

Since by definition $\Delta\overline{\alpha}(s) = \log(1+\Delta\alpha(s))$ we obtain $\Delta G(s) = -G(s)\Delta\alpha(s)$ which coincides with $\Delta H(s)$ (since $s \mapsto \int_{(s,t]} G^2 d\beta$ is continuous). Let us now compare the continuous parts of G and H, resp. denoted by G^c and H^c. Starting from the right-hand side of (2.9), the chain rule for functions of bounded variations gives

$$dG^c(s) = \frac{-\alpha(ds)\exp(-\overline{\alpha}(s))}{\lambda^{-1}\exp(-\overline{\alpha}(t)) + \int_{(s,t]}\exp(-\overline{\alpha}(y))\beta(dy)} + \frac{\exp(-2\overline{\alpha}(s))\beta(ds)}{\left(\lambda^{-1}\exp(-\overline{\alpha}(t)) + \int_{(s,t]}\exp(-\overline{\alpha}(y))\beta(dy)\right)^2},$$

i.e., $dG^c(s) = -G(s)\alpha(ds) + G(s)^2\beta(ds) = dH^c(s)$. This proves (2.9) from which (2.10) follows from the facts that $\mathbb{P}(X(t) = 0 \mid X(s) = x) = \lim_{\lambda\to\infty}\mathbb{E}(e^{-\lambda X(t)} \mid X(s) = x)$ and that $\mathbb{E}(e^{-\lambda X(t)} \mid X(s) = x) = \exp(-xu(s,t,\lambda))$. □

2.5.3 Remarks on CSBP with catastrophes

Theorem 2.2 makes it possible to study Galton–Watson processes where only few offspring distributions are not near-critical. The simplest example is given by taking $\gamma_n(t) = \lfloor \Gamma_n t \rfloor$ and $q_{i,n} = q_{0,n}$, in such a way that the corresponding sequence of scaled Galton–Watson processes converges to a CSBP. Then, for some $t_0 \geq 0$, one can change $q_{\gamma_n(t_0),n}$ and take its mean equal to $1 + a$. Thus (X_n) converges to a process X which is a CSBP on $[0, t_0)$ and on $[t_0, \infty)$ and such that $X(t_0) = (1 + a)X(t_0-)$.
Such processes with catastrophes have been studied in [6] with motivations for cell division models. More precisely, CSBP's are multiplied at a constant rate by some random number which yields the impact of the catastrophe. The successive times of catastrophes and their impact follow a Poisson point process with intensity $rdt\mathbb{P}(F \in d\theta)$, whose associated Lévy process has finite variation. Theorem 2.2 thus yields an alternative way to construct the process and characterize its Laplace exponent, whereas [6] uses results on stochastic differential equations with jumps.

Another way to create a discontinuity at a fixed time is to take $q_{\gamma_n(t_0),n} = (1 - 1/n)\delta_0 + (1/n)\delta_n$ as in the example considered in the beginning of Section 2.3. Again, (X_n) converges to a process X which is a CSBP on $[0, t_0)$ and on $[t_0, \infty)$ and such that $X(t_0) = S(X(t_0-))$ with $(S(x), x \geq 0)$ a Poisson process. Theorem 2.2 allows accumulation of such fixed jumps; note that in both cases these jumps may be negative, whereas CSBP's only have positive jumps.

Building on these two simple examples, we expect in general that if X is a time-inhomogeneous Markov process satisfying the branching property, then for each fixed time of discontinuity t, there should exist a subordinator $S_t = (S_t(x), x \geq 0)$ such that $X(t) = S_t(X(t-))$.

3 Additional notation and preliminary results

In this section we gather some notation used throughout the rest of the paper. Of particular importance are the constants and functions defined in Section 3.3, which will be used repeatedly in the proofs.

3.1 Additional notation

From now on we identify any càdlàg function of locally finite variation f with its corresponding signed measure, see for instance Chapter 3 in Kallenberg [30]. For instance, we will write indifferently $f((s,t])$, $f(s,t]$ or $f(t) - f(s)$ for $0 \leq s \leq t$, as well as $\Delta f(t)$ or $f\{t\}$. Let g and h be defined as follows:

$$g(x,\lambda) = 1 - e^{-\lambda x} - \frac{\lambda x}{1+x^2} \quad \text{and} \quad h(x,\lambda) = g(x,\lambda) + \frac{(\lambda x)^2}{2(1+x^2)}, \; x \in \mathbb{R}, \lambda \geq 0. \tag{3.1}$$

For $n \geq 1$ let in the sequel $\mu_n = \|\alpha_n\| + \beta_n$, i.e.,

$$\mu_n(t) = \|\alpha_n\|(t) + \beta_n(t),\ t \geq 0, \tag{3.2}$$

and for $i \geq 0$ and $n \geq 1$, let

$$\psi_{i,n}(\lambda) = u_n(t_i^n, t_{i+1}^n, \lambda) - \lambda = -n \log \left(1 - \frac{1}{n} \int \left(1 - e^{-\lambda x}\right) \nu_{i,n}(dx) \right),\ \lambda \geq 0, \tag{3.3}$$

where $\log$ stands for the natural logarithm and the second equality is derived from the definition of u_n as Laplace exponent of X_n. In order to use the approximation $\psi_{i,n}(\lambda) \approx \int (1 - e^{-\lambda x}) \nu_{i,n}(dx)$, we introduce the function $\epsilon_{i,n}$ such that

$$\psi_{i,n}(\lambda) = (1 + \epsilon_{i,n}(\lambda)) \int \left(1 - e^{-\lambda x}\right) \nu_{i,n}(dx), \tag{3.4}$$

with $\epsilon_{i,n}(\lambda) = 0$ when $\int (1 - e^{-\lambda x}) \nu_{i,n}(dx) = 0$.

For every $n \geq 1$ and every measurable, positive function $f : [0, \infty) \to (0, \infty)$, define the two measures $\Psi(f)$ and $\Psi_n(f)$ as follows:

$$\Psi(f)(A) = \int_A f(y) \alpha(dy) - \int_A f(y)^2 \beta(dy) + \int_{(0,\infty) \times A} h(x, f(y)) \nu(dx\, dy),\ A \in \mathcal{B},$$

with $\mathcal{B}$ the Borel subsets of $\mathbb{R}$, and

$$\Psi_n(f)(A) = \sum_{i \geq 1} \mathbb{1}_{\{t_i^n \in A\}} \psi_{i-1,n}(f(t_i^n)),\ A \in \mathcal{B}.$$

With a slight abuse of notation, we will also consider $\Psi(f)$ and $\Psi_n(f)$ for functions f only defined on a subset of $[0, \infty)$, typically $[\wp(t), t]$. Then we will only consider $\Psi(f)(A)$ or $\Psi_n(f)(A)$ for Borel sets A which are subset of the domain of definition of f.

3.2 Heuristic derivation of (2.3)

We first note that (2.3) can be rewritten as

$$u(s, t, \lambda) = \lambda + \Psi(u(\,\cdot\,, t, \lambda))((s, t]). \tag{3.5}$$

To see that u_n satisfies a similar dynamics, note that from the definition of u_n and the Markov property of X_n, we get the following composition rule:

$$u_n(t_1, t_3, \lambda) = u_n(t_1, t_2, u_n(t_2, t_3, \lambda)), \quad 0 \leq t_1 \leq t_2 \leq t_3,\ \lambda \geq 0. \tag{3.6}$$

Lemma 3.1. *For any $n \geq 1$, $\lambda \geq 0$ and $0 \leq s \leq t$, it holds that*

$$u_n(s, t, \lambda) = \lambda + \sum_{i=\gamma_n(s)+1}^{\gamma_n(t)} \psi_{i-1,n}(u_n(t_i^n, t, \lambda)) = \lambda + \Psi_n(u_n(\,\cdot\,, t, \lambda))((s, t]). \tag{3.7}$$

Proof. The second equality follows readily from the definition of Ψ_n, while the first one can be derived as follows:

$$\begin{aligned} u_n(s, t, \lambda) = u_n(t_{\gamma_n(s)}^n, t, \lambda) &= \lambda + \sum_{i=\gamma_n(s)}^{\gamma_n(t)-1} \left(u_n(t_i^n, t, \lambda) - u_n(t_{i+1}^n, t, \lambda) \right) \\ &\overset{\text{(i)}}{=} \lambda + \sum_{i=\gamma_n(s)}^{\gamma_n(t)-1} \left(u_n(t_i^n, t_{i+1}^n, u_n(t_{i+1}^n, t, \lambda)) - u_n(t_{i+1}^n, t, \lambda) \right) \\ &\overset{\text{(ii)}}{=} \lambda + \sum_{i=\gamma_n(s)}^{\gamma_n(t)-1} \psi_{i,n}(u_n(t_{i+1}^n, t, \lambda)), \end{aligned}$$

where (i) comes from the composition rule (3.6) and (ii) comes from the first equality in (3.3). □

From (3.7) we can now let (3.5) (i.e., (2.3)) appear. Indeed, in view of the second equality in (3.3) and of the approximation $\log(1-x) \approx -x$, it is reasonable to expect

$$\psi_{i,n}(\lambda) \approx \int \left(1-e^{-\lambda x}\right) \nu_{i,n}(dx) = \lambda\alpha_{i,n} - \lambda^2\beta_{i,n} + \int_{(0,\infty)} h(x,\lambda)\nu_{i,n}(dx)$$

(recall the definition (3.1) of h for the last equality) and so summing over $i = 0, \ldots, \gamma_n(t)-1$ yields through (3.7) the approximation

$$u_n(s,t,\lambda) \approx \lambda + \int_{(s,t]} u_n(y,t,\lambda)\alpha_n(dy) - \int_{(s,t]} (u_n(y,t,\lambda))^2\beta_n(dy) + \int_{(0,\infty)\times(s,t]} h(x,u_n(y,t,\lambda))\nu_n(dx\,dy).$$

Since $(\alpha_n, \beta_n, \nu_n)$ is assumed to converge toward (α, β, ν), this last approximation suggests that any limit $u(s,t,\lambda)$ of the sequence $(u_n(s,t,\lambda))$ should indeed satisfy (3.5).

3.3 Key constants and functions

For any $n \geq 1$, $t, \lambda, C \geq 0$, $s \leq t$, $N \geq 1$, $0 < \eta < T$, let:

$$c_1(C) = C + c_1'(C) \ \text{ with } \ c_1'(C) = \sup\left\{\frac{2|g(x,\lambda)|(1+x^2)}{x^2} : x \geq -1, 0 \leq \lambda \leq C\right\}, \tag{3.8}$$

$$c_2(\eta,T) = \sup_{\substack{\eta\leq y,y'\leq T\\ 0<x,\ y\neq y'}} \left|\frac{h(x,y)-h(x,y')}{(y-y')x^2/(1+x^2)}\right| \text{ and } c_3(\eta,T) = 1+T+c_2(\eta,T), \tag{3.9}$$

$$\overline{c}^{\epsilon}_{n,t}(C) = \sup\left\{|\epsilon_{i,n}(\lambda)| : 0 \leq i < \gamma_n(t), 0 \leq \lambda \leq C\right\}, \tag{3.10}$$

$$\overline{c}^{u}_{t,\lambda} = \sup\left\{u_n(s,t,\lambda) : n \geq 1, 0 \leq s \leq t\right\}, \tag{3.11}$$

$$\Delta^u_{t,\lambda} = \left(1 + \sup_{n\geq 1}\left\{\overline{c}^{\epsilon}_{n,t}\left(\overline{c}^u_{t,\lambda}\right)\right\}\right) c_1\left(\overline{c}^u_{t,\lambda}\right), \tag{3.12}$$

$$\underline{c}^u_{s,t,\lambda}(N) = \inf\left\{u_n(y,t,\lambda) : s \leq y \leq t, n \geq N\right\}, \tag{3.13}$$

and $N_{s,t,\lambda} = \inf\left\{N \geq 1 : \underline{c}^u_{s,t,\lambda}(N) > 0\right\}$. When $N_{s,t,\lambda}$ is finite, we also define

$$\underline{c}^u_{s,t,\lambda} = \underline{c}^u_{s,t,\lambda}(N_{s,t,\lambda}), \tag{3.14}$$

in which case $\underline{c}^u_{s,t,\lambda} > 0$. We defer the proofs that these constants and numbers are finite to Appendix B, and we now show how to use them to prove key results. Of particular importance are Lemma 3.3, which controls fluctuations of $u_n(s,t,\lambda)$ in s, and Lemma 3.4 which allows to rewrite the time of last bottleneck $\wp(t)$ in a more convenient form.

Lemma 3.2. *For any $C \geq 0$, $n \geq 1$ and $i \geq 0$,*

$$\sup_{0\leq\lambda\leq C}\left|\int \left(1-e^{-\lambda x}\right)\nu_{i,n}(dx)\right| \leq c_1(C)\mu_n(t^n_i, t^n_{i+1}]. \tag{3.15}$$

Proof. By definition (3.1) of g, we have

$$\int \left(1-e^{-\lambda x}\right)\nu_{i,n}(dx) = \lambda\alpha_{i,n} + \int g(x,\lambda)\nu_{i,n}(dx)$$

so that $|\int(1-e^{-\lambda x})\nu_{i,n}(dx)| \leq \lambda|\alpha_{i,n}| + \int|g(x,\lambda)|\nu_{i,n}(dx)$. Since

$$|g(x,\lambda)| \leq c_1'(C)\frac{x^2}{2(1+x^2)}$$

for all $x \geq -1$ and $0 \leq \lambda \leq C$ by definition of $c_1'(C)$, we get (3.15). □

Lemma 3.3. *For any $n \geq 1$, $\lambda, t > 0$ and $0 \leq s \leq s' \leq t$,*

$$|u_n(s,t,\lambda) - u_n(s',t,\lambda)| \leq \Delta^u_{t,\lambda} \mu_n(s,s']. \tag{3.16}$$

Proof. Lemma 3.1 and the definition of $\epsilon_{i,n}$ give

$$|u_n(s,t,\lambda) - u_n(s',t,\lambda)| \\ \leq \sum_{i=\gamma_n(s)+1}^{\gamma_n(s')} (1 + |\epsilon_{i-1,n}(u_n(t^n_i,t,\lambda))|) \left| \int (1 - e^{-xu_n(t^n_i,t,\lambda)}) \nu_{i-1,n}(dx) \right|.$$

Since $0 \leq t^n_i \leq t$ for any $0 \leq i \leq \gamma_n(t)$, we have $u_n(t^n_i,t,\lambda) \leq \bar{c}^u_{t,\lambda}$ and in particular $|\epsilon_{i-1,n}(u_n(t^n_i,t,\lambda))| \leq \bar{c}^\epsilon_{n,t}(\bar{c}^u_{t,\lambda})$ for all $\gamma_n(s) < i \leq \gamma_n(s')$. Using in addition (3.15) with $C = \bar{c}^u_{t,\lambda}$, we obtain

$$|u_n(s,t,\lambda) - u_n(s',t,\lambda)| \leq \sum_{i=\gamma_n(s)+1}^{\gamma_n(s')} \left(1 + \bar{c}^\epsilon_{n,t}(\bar{c}^u_{t,\lambda})\right) c_1(\bar{c}^u_{t,\lambda}) \mu_n(t^n_{i-1}, t^n_i] = \Delta^u_{t,\lambda} \mu_n(s,s']$$

which gives (3.16). □

In the next lemma we provide an alternative expression for $\wp(t)$, defined so far as $\sup \mathcal{S}(t)$ with

$$\mathcal{S}(t) = \left\{ s \leq t : \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \inf_{s \leq y \leq t} \mathbb{P}\left(X_n(t) > \varepsilon \;\middle|\; X_n(y) = 1 \right) = 0 \right\}.$$

More precisely, we show that $\wp(t) = \sup \mathcal{S}(t,\lambda)$ where for each $\lambda \geq 0$,

$$\mathcal{S}(t,\lambda) = \left\{ s \leq t : \liminf_{n \to \infty} \inf_{s \leq y \leq t} u_n(y,t,\lambda) = 0 \right\}.$$

In particular, we deduce that u_n past $\wp(t)$ is uniformly bounded away from 0 (i.e., $N_{s,t,\lambda}$ is finite and $\underline{c}^u_{s,t,\lambda} > 0$ for $s > \wp(t)$), which is in line with the intuition behind $\wp(t)$ being the last bottleneck before time t.

Lemma 3.4. *Fix $t > 0$ and assume that the sequences $(\|\alpha_n\|(t), n \geq 1)$ and $(\beta_n(t), n \geq 1)$ are bounded. Then $\wp(t) = \sup \mathcal{S}(t,\lambda)$ for every $\lambda > 0$ and $N_{s,t,\lambda}$ is finite for every $s \in (\wp(t), t]$.*

In particular, if $(\|\alpha_n\|(t), n \geq 1)$ and $(\beta_n(t), n \geq 1)$ are bounded for every $t \geq 0$, then the function $t \mapsto \wp(t)$ is increasing.

Proof. Fix in the rest of the proof $t, \lambda > 0$ and let $s \leq t$: the following statements are equivalent, which proves that $\mathcal{S}(t) = \mathcal{S}(t,\lambda)$ and implies $\wp(t) = \sup \mathcal{S}(t,\lambda)$:

(i) $\liminf_{n \to \infty} \inf_{s \leq y \leq t} u_n(y,t,\lambda) = 0$;

(ii) there exist sequences $(n(k))$ and (y_k) such that $y_k \in [s,t]$ for each $k \geq 1$ and

$$\lim_{k \to \infty} n(k) = \infty \text{ and } \lim_{k \to \infty} u_{n(k)}(y_k,t,\lambda) = 0;$$

(iii) there exist sequences $(n(k))$ and (y_k) such that $y_k \in [s,t]$ for each $k \geq 1$ and

$$\lim_{k \to \infty} n(k) = \infty \text{ and for every } v > 0, \lim_{k \to \infty} \mathbb{E}\left(e^{-vX_{n(k)}(t)} \mid X_{n(k)}(y_k) = 1 \right) = 1;$$

(iv) there exist sequences $(n(k))$ and (y_k) such that $y_k \in [s,t]$ for each $k \geq 1$ and for any $\varepsilon > 0$,

$$\lim_{k\to\infty} n(k) = \infty \text{ and } \lim_{k\to\infty} \mathbb{P}\left(X_{n(k)}(t) > \varepsilon \,|\, X_{n(k)}(y_k) = 1\right) = 0;$$

(v) $\lim_{\varepsilon\to 0} \liminf_{n\to\infty} \inf_{s\leq y\leq t} \mathbb{P}\big(X_n(t) > \varepsilon \mid X_n(y) = 1\big) = 0$.

The equivalence between (iii) and (iv) relies on the fact that both conditions are equivalent to the following one: the sequence of random variables $(X_{n(k)}(t), k \geq 1)$ under $\mathbb{P}(\,\cdot \mid X_{n(k)}(y_k) = 1)$ converges in distribution to 0. Let us also explain the last equivalence. The condition (iv) implies that

$$\liminf_{n\to\infty} \inf_{s\leq y\leq t} \mathbb{P}\big(X_n(t) > \varepsilon \mid X_n(y) = 1\big) = 0$$

for every $\varepsilon > 0$, which is stronger than (v). Now, assuming that (v) holds, one can find sequences (ε_k), $(n(k))$ and (y_k) such that $y_k \in [s,t]$ and

$$\lim_{k\to\infty} \varepsilon_k = 0, \lim_{k\to\infty} n(k) = \infty \quad \text{and} \quad \lim_{k\to\infty} \mathbb{P}\left(X_{n(k)}(t) > \varepsilon_k \mid X_{n(k)}(y_k) = 1\right) = 0.$$

Then the sequences $(n(k))$ and (y_k) satisfy (iv) since for any $\varepsilon > 0$,

$$\mathbb{P}\left(X_{n(k)}(t) > \varepsilon \mid X_{n(k)}(y_k) = 1\right) \leq \mathbb{P}\left(X_{n(k)}(t) > \varepsilon_k \mid X_{n(k)}(y_k) = 1\right)$$

for k large enough, since $\varepsilon_k \to 0$. This proves $\wp(t) = \sup \mathcal{S}(t,\lambda)$, which implies that $N_{s,t,\lambda}$ is finite when $\wp(t) < s \leq t$ since from the definition (3.13) of $\underline{c}^u_{s,t,\lambda}(N)$,

$$\lim_{N\to\infty} \underline{c}^u_{s,t,\lambda}(N) = \liminf_{n\to\infty} \inf_{s\leq y\leq t} u_n(y,t,\lambda).$$

We now assume that $(\|\alpha_n\|(t))$ and $(\beta_n(t))$ are bounded for every $t \geq 0$, and prove that $\wp(\,\cdot\,)$ is an increasing function. Let $t' > t$: we will show that $\mathcal{S}(t, \bar{c}^u_{t',\lambda}) \subset \mathcal{S}(t',\lambda)$, which proves that $\wp(t) \leq \wp(t')$. So consider $s \in \mathcal{S}(t, \bar{c}^u_{t',\lambda})$, i.e., $s \leq t$ with

$$\liminf_{n\to\infty} \inf_{s\leq y\leq t} u_n(y,t,\bar{c}^u_{t',\lambda}) = 0.$$

Then $s \leq t'$, and the composition rule (3.6) together with the monotonicity of u_n in λ give for any $s \leq y \leq t$

$$u_n(y,t',\lambda) = u_n(y,t,u_n(t,t',\lambda)) \leq u_n(y,t,\bar{c}^u_{t',\lambda})$$

which entails

$$\liminf_{n\to\infty} \inf_{s\leq y\leq t'} u_n(y,t',\lambda) \leq \liminf_{n\to\infty} \inf_{s\leq y\leq t} u_n(y,t,\bar{c}^u_{t',\lambda}).$$

Since this last quantity is equal to 0 this proves that $s \in \mathcal{S}(t',\lambda)$ and gives the result. □

4 Proof of Theorem 2.2

In this section, we assume that Assumption 2.1 holds and we consider the measures α, β and ν given there. Recall that $\mu_n = \|\alpha_n\| + \beta_n$, and define analogously $\mu = \|\alpha\| + \beta$, in particular we have

$$\lim_{n\to\infty} \mu_n(t) = \mu(t). \tag{4.1}$$

Define also the measure $\widetilde{\mu}$ by

$$\widetilde{\mu}(A) = \mu(A) + \int_{(0,\infty)\times A} \frac{x^2}{1+x^2} \nu(dx\,dy),\ A \in \mathcal{B}.$$

We prove the four claims I–IV in Sections 4.1 to 4.4.

4.1 Proof of claim I

That $\Delta\alpha(t) \geq -1$ is a direct consequence of (A2) since $\alpha_{i,n} \geq -1$ for every $i \geq 0$ and $n \geq 1$. Moreover, note that

$$\int_{(0,\infty)\times I_{s,t}} \frac{x^2}{2(1+x^2)}\nu(dx\,dy) \leq \beta(I_{s,t}), 0 \leq s \leq t, \tag{4.2}$$

where $I_{s,t} = (s,t]$ or $I_{s,t} = (s,t)$. Indeed, for $I_{s,t} = (s,t]$

$$\begin{aligned}\int_{(0,\infty)\times(s,t]} \frac{x^2}{2(1+x^2)}\nu(dx\,dy) &\overset{\text{(i)}}{=} \int_{(0,\infty)} \frac{x}{(1+x^2)^2}\nu([x,\infty)\times(s,t])dx \\ &\overset{\text{(ii)}}{=} \int_{(0,\infty)} \frac{x}{(1+x^2)^2} \liminf_{n\to\infty} \nu_n([x,\infty)\times(s,t])dx \\ &\overset{\text{(iii)}}{\leq} \liminf_{n\to\infty} \int_{(0,\infty)} \frac{x}{(1+x^2)^2}\nu_n([x,\infty)\times(s,t])dx \\ &\overset{\text{(iv)}}{=} \liminf_{n\to\infty} \int_{(0,\infty)\times(s,t]} \frac{x^2}{2(1+x^2)}\nu_n(dx\,dy) \\ &\overset{\text{(v)}}{\leq} \liminf_{n\to\infty} \left(\beta_n((s,t])\right)\end{aligned}$$

using Fubini's theorem for (i) and (iv), the assumption (A1) for (ii) (using also that the set $\{x : \nu(\{x\}\times(s,t]) > 0\}$ has zero Lebesgue measure), Fatou's lemma for (iii) and finally the definition of ν_n and β_n for (v). To get the result for $I_{s,t} = (s,t)$ apply (4.2) to $(s,t']$ and let $t' \uparrow t$. The inequality (4.2) has two direct consequences: $\int_{(0,\infty)\times(0,t]}(1\wedge x^2)\nu(dx\,dy)$ is finite and the function $\widetilde{\beta}$ is increasing. Thus to conclude the proof of claim I, it remains to show that $\widetilde{\beta}$ is continuous, i.e., $\Delta\widetilde{\beta}(t) = 0$.

First, by letting $s \uparrow t$ in (4.2) with $I_{s,t} = (s,t]$ we see that $\Delta\widetilde{\beta}(t) = 0$ when $\Delta\beta(t) = 0$, so we only have to consider the case where $\Delta\beta(t) > 0$. In this case, the assumption (A2) implies that $\beta_{\gamma_n(t),n} \to \Delta\beta(t)$ and that $\nu_{\gamma_n(t),n}([x,\infty)) \to \nu([x,\infty)\times\{t\})$ for every $x > 0$ such that $\nu(\{x\}\times\{t\}) = 0$. Then for any $d > 0$ with $\nu(\{d\}\times\{t\}) = 0$, we get by weak convergence of probability measures (since all the measures restricted to $[d,\infty)$ have finite mass) and the dominated convergence theorem

$$\lim_{n\to\infty} \int_{(d,\infty)} \frac{x^2}{2(1+x^2)}\nu_{\gamma_n(t),n}(dx) = \int_{(d,\infty)\times\{t\}} \frac{x^2}{2(1+x^2)}\nu(dx\,dy). \tag{4.3}$$

On the other hand, we have

$$\int_{[-1/n,d]} \frac{x^2}{1+x^2}\nu_{\gamma_n(t),n}(dx) \leq \left(d+\frac{1}{n}\right)\int_{[-1/n,d]} \frac{|x|}{1+x^2}\nu_{\gamma_n(t),n}(dx)$$

and from the definition of $\alpha_{\gamma_n(t),n}$ we see that

$$\begin{aligned}\int_{[-1/n,d]} \frac{|x|}{1+x^2}\nu_{\gamma_n(t),n}(dx) &= \int_{[-1/n,d]} \frac{x}{1+x^2}\nu_{\gamma_n(t),n}(dx) + \frac{2/n}{1+(1/n)^2}\nu_{\gamma_n(t),n}\{-1/n\} \\ &\leq \alpha_{\gamma_n(t),n} + \frac{2n^2}{1+n^2}\end{aligned}$$

using that $\nu_{\gamma_n(t),n}\{-1/n\} \leq \nu_{\gamma_n(t),n}(\mathbb{R}) = n$ for the last inequality. Since $|\alpha_{\gamma_n(t),n}| \leq \|\alpha_n\|(t)$ and $\|\alpha_n\|(t) \to \|\alpha\|(t)$, we obtain from the two last displays

$$\lim_{d\to 0}\limsup_{n\to\infty} \int_{[-1/n,d]} \frac{x^2}{1+x^2}\nu_{\gamma_n(t),n}(dx) = 0.$$

Combined with (4.3), this gives

$$\lim_{n\to\infty}\int \frac{x^2}{2(1+x^2)}\nu_{\gamma_n(t),n}(dx) = \lim_{d\to 0}\int_{(d,\infty)\times\{t\}} \frac{x^2}{2(1+x^2)}\nu(dx\,dy)$$
$$= \int_{(0,\infty)\times\{t\}} \frac{x^2}{2(1+x^2)}\nu(dx\,dy).$$

Since $2\beta_{\gamma_n(t),n} = \int \frac{x^2}{1+x^2}\nu_{\gamma_n(t),n}(dx)$ and $\beta_{\gamma_n(t),n} \to \Delta\beta(t)$ this concludes the proof of claim I.

4.2 Proof of claim II

First of all, note that under Assumption 2.1, we have the convergence $\mu_n \to \mu$ in the J_1 topology, see for instance [28, Section VI.2]. In view of (3.3) and the Arzelá-Ascoli theorem, for each $t, \lambda \geq 0$ this implies that the sequence of functions $(u_n(\,\cdot\,, t, \lambda), n \geq 1)$ on $[0, t]$ is relatively compact. Thus in order to prove claim 4.2, we only have to prove the pointwise convergence $u_n(s, t, \lambda) \to u(s, t, \lambda)$.

In order to do so, the (classical) idea is to use a Lipschitz property satisfied by Ψ, combined with Gronwall's lemma. The Lipschitz property of Ψ takes the following form, where we use the space $\mathcal{F}$ of measurable and positive functions f such that $0 < \inf f \leq \sup f < \infty$ and we remember the constant c_3 defined in (3.9). For any functions $f_1, f_2 \in \mathcal{F}$ and any $A \in \mathcal{B}$, we have

$$|\Psi(f_1)(A) - \Psi(f_2)(A)| \leq c_3\left(\inf_A f_1 \wedge \inf_A f_2, \sup_A f_1 + \sup_A f_2\right)\int_A |f_1 - f_2| d\widetilde{\mu}. \tag{4.4}$$

Indeed, let $\eta = \inf_A f_1 \wedge \inf_A f_2$ and $T = \sup_A f_1 + \sup_A f_2$: then by definition of Ψ we have

$$|\Psi(f_1)(A) - \Psi(f_2)(A)| \leq \int_A |f_1 - f_2| d\|\alpha\| + \int_A \left|f_1^2 - f_2^2\right| d\beta$$
$$+ \int_{(0,\infty)\times A} |h(x, f_1(y)) - h(x, f_2(y))|\,\nu(dx\,dy).$$

Using $|f_1^2 - f_2^2| = |f_1 - f_2|(f_1 + f_2)$ and plugging in the constant c_2, we obtain

$$|\Psi(f_1)(A) - \Psi(f_2)(A)| \leq \int_A |f_1 - f_2| d\|\alpha\| + T\int_A |f_1 - f_2| d\beta$$
$$+ c_2(\eta, T)\int_{(0,\infty)\times A} |f_1(y) - f_2(y)|\,\frac{x^2}{1+x^2}\nu(dx\,dy) \leq c_3(\eta, T)\int_A |f_1 - f_2| d\widetilde{\mu},$$

which establishes (4.4). We will invoke this property using the following backwards version of Gronwall's lemma. The proof is standard and omitted.

Lemma 4.1. *Let u and R be non-negative, càdlàg functions and let π be a locally finite and positive measure. If*

$$u(s) \leq R(s) + \int_{(s,t]} u(x)\pi(dx)$$

holds for all $0 \leq s \leq t$, then for all $0 \leq s \leq t$ we have

$$u(s) \leq R(s) + e^{\pi(s,t]}\int_{(s,t]} R(x)\pi(dx).$$

The claim II of Theorem 2.2 follows readily from Lemma 4.3 below. The proof of this lemma uses the following result, whose long proof is postponed to Appendix C.

Lemma 4.2. *Fix $t, \lambda > 0$ and consider any sequence (ℓ_n) with $\ell_n \to \lambda$. For $n \geq 1$, let R_n be the function*

$$R_n(s) = \left|\Psi_n(u_n(\,\cdot\,, t, \ell_n))((s,t]) - \Psi(u_n(\,\cdot\,, t, \ell_n))((s,t])\right|, \ 0 \leq s \leq t.$$

Then $R_n(s) \to 0$ for any $\wp(t) < s \leq t$ and $\sup\{R_n(s) : 0 \leq s \leq t, n \geq 1\}$ is finite.

Lemma 4.3. *Fix $t, \lambda > 0$ and a sequence (ℓ_n) with $\ell_n \to \lambda$. Then for any $s \in [\wp(t), t]$, the sequence $(u_n(s, t, \ell_n), n \geq 1)$ converges and the function*

$$u : s \in [\wp(t), t] \mapsto \lim_{n \to \infty} u_n(s, t, \ell_n)$$

is the unique function satisfying the following properties:

1. *$u(s) = \lambda + \Psi(u)((s,t])$ for all $\wp(t) \leq s \leq t$;*
2. *u is càdlàg;*
3. *$\inf_{[s,t]} u > 0$ for any $\wp(t) < s \leq t$.*

In particular, $u(s, t, \lambda)$ does not depend on the sequence ℓ_n.

Proof. In the rest of the proof fix $t, \lambda > 0$ and (ℓ_n) a sequence converging to λ. Let $\ell = \inf_{n \geq 1} \ell_n$ and $L = \sup_{n \geq 1} \ell_n$ and assume without loss of generality, since $\ell_n \to \lambda > 0$, that $\ell > 0$. To ease the notation, we write in the rest of the proof $\wp = \wp(t)$ and $u_n(s) = u_n(s, t, \ell_n)$ for $0 \leq s \leq t$. We decompose the proof in four steps: first we prove that the sequence $(u_n(s), n \geq 1)$ is Cauchy for any $s \in (\wp, t]$, then that it is Cauchy for $s = \wp$, then that u satisfies the claimed properties and finally that it is the only such function.

Before beginning, note that everything is trivial if $\wp = t$, because then $u_n(s) = \ell_n$ and $\Psi(u)((s,t]) = 0$ for any $s \in [\wp, t]$. Hence in the sequel we assume that $\wp < t$.

First step: $(u_n(s))$ is Cauchy for $s \in (\wp, t]$. In the rest of this step fix $s \in (\wp, t]$ and for $s \leq y \leq t$ we define as in Lemma 4.2 $R_n(y) = |\Psi_n(u_n)((y,t]) - \Psi(u_n)((y,t])|$. Then the second equality in (3.7) gives for any $s \leq y \leq t$ and any $m, n \geq 1$

$$|u_n(y) - u_m(y)| \leq R_n(y) + R_m(y) + |\Psi(u_n)((y,t]) - \Psi(u_m)((y,t])|.$$

We get from (4.4)

$$|\Psi(u_n)((y,t]) - \Psi(u_m)((y,t])| \leq c_3\left(\inf_{(y,t]} u_n \wedge \inf_{(y,t]} u_m, \sup_{(y,t]} u_n + \sup_{(y,t]} u_m\right) \int_{(y,t]} |u_n - u_m|\, d\widetilde{\mu}.$$

Since the function $u_n(s, t, \lambda)$ is increasing in λ, we have for any $y \in [s,t]$ and $n \geq N_{s,t,\ell}$ (recall that $N_{s,t,\ell}$ is defined in (3.13) and is finite by Lemma 3.4)

$$u_n(y) = u_n(y, t, \ell_n) \geq u_n(y, t, \ell) \geq \inf_{s \leq y' \leq t} u_n(y', t, \ell) \geq \underline{c}^u_{s,t,\ell} > 0.$$

Similar monotonicity arguments lead to $u_n(y) \leq \bar{c}^u_{t,L}$ for any $y \leq t$ and $n \geq 1$, so that monotonicity properties of $c_3(\eta, T)$ in η and T give for $n, m \geq N_{s,t,\lambda}$

$$|\Psi(u_n)((y,t]) - \Psi(u_m)((y,t])| \leq c_3\left(\underline{c}^u_{s,t,\ell}, 2\bar{c}^u_{t,L}\right) \int_{(y,t]} |u_n - u_m|\, d\widetilde{\mu}.$$

We finally get the bound

$$|u_n(y) - u_m(y)| \leq R_n(y) + R_m(y) + C \int_{(y,t]} |u_n - u_m|\, d\widetilde{\mu}$$

with $C = C_{s,t,\ell,L} = c_3(\underline{c}^u_{s,t,\ell}, 2\overline{c}^u_{t,L})$, which holds for all $s \leq y \leq t$ and all $n, m \geq N_{s,t,\ell}$. Thus Lemma 4.1 implies for those n, m

$$|u_n(s) - u_m(s)| \leq R_n(s) + R_m(s) + Ce^{C\widetilde{\mu}(s,t]} \int_{(s,t]} (R_n + R_m)\, d\widetilde{\mu}$$

so that for any $n_0 \geq N_{s,t,\ell}$,

$$\sup_{n,m \geq n_0} |u_n(s) - u_m(s)| \leq 2 \sup_{n \geq n_0} (R_n(s)) + 2Ce^{C\widetilde{\mu}(s,t]} \sup_{n \geq n_0} \left(\int_{(s,t]} R_n d\widetilde{\mu} \right).$$

Lemma 4.2 combined with the dominated convergence theorem shows that the right hand side of the above inequality goes to 0 as $n_0 \to \infty$ which proves that the sequence $(u_n(s), n \geq 1)$ is Cauchy and completes the proof of this first step.

Second step: $(u_n(\wp))$ is Cauchy. For any $\wp < s' \leq t$, (3.16) entails

$$\begin{aligned} |u_n(\wp) - u_m(\wp)| &\leq |u_n(\wp) - u_n(s')| + |u_m(\wp) - u_m(s')| + |u_n(s') - u_m(s')| \\ &\leq 2\Delta^u_{t,\ell_n} \mu_n(\wp, s'] + |u_n(s') - u_m(s')| \\ &\leq 2\Delta^u_{t,L} \mu_n(\wp, s'] + |u_n(s') - u_m(s')| \end{aligned}$$

using for the last inequality that $\ell_n \leq L$ and that $\Delta^u_{t,y}$ is increasing in y, as can be seen directly from its definition (3.12). Hence for any $n_0 \geq 1$,

$$\sup_{m,n \geq n_0} |u_n(\wp) - u_m(\wp)| \leq 2\Delta^u_{t,L} \sup_{n \geq n_0} \mu_n(\wp, s'] + \sup_{m,n \geq n_0} |u_n(s') - u_m(s')|.$$

By (4.1) and the fact that $(u_n(s'))$ is Cauchy by the first step since $\wp < s' \leq t$, the right hand side of the above inequality goes to $2\Delta^u_{t,L}\mu(\wp, s']$ as n_0 goes to infinity. Since $\mu(\wp, s'] \to 0$ as $s' \downarrow \wp$, letting then $s' \downarrow \wp$ shows that $(u_n(\wp))$ is Cauchy.

Third step: properties of u. Let from now on u denote the function of the statement and consider $s \in [\wp, t]$. First note that the second property follows from the first one, so we only have to prove the first and third ones. Assume first that $s > \wp$. We have seen in the first step that for any $s \leq y \leq t$ and $n \geq N_{s,t,\ell}$

$$0 < \underline{c}^u_{s,t,\ell} \leq u_n(y) \leq \overline{c}^u_{t,L} < \infty.$$

Since $u_n(y) \to u(y)$ for $s \leq y \leq t$ by definition of u, u also satisfies $\underline{c}^u_{s,t,\lambda} \leq u(y) \leq \overline{c}^u_{t,L}$ for $s \leq y \leq t$. In particular the third property $\inf_{[s,t]} u > 0$ is satisfied. Let us now show the first property, still in the case $s > \wp$. Plugging in (3.7), we get

$$\begin{aligned} |u(s) - \lambda - \Psi(u)((s,t])| \leq |u(s) - u_n(s)| &+ |\Psi_n(u_n)((s,t]) - \Psi(u_n)((s,t])| \\ &+ |\Psi(u_n)((s,t]) - \Psi(u)((s,t])|. \end{aligned}$$

Since both u_n and u are bounded uniformly on $[s,t]$ by $\underline{c}^u_{s,t,\ell}$ and $\overline{c}^u_{t,L}$, we get with similar arguments as in the first step

$$|\Psi(u_n)((s,t]) - \Psi(u)((s,t])| \leq c_3 \left(\underline{c}^u_{s,t,\ell}, 2\overline{c}^u_{t,L}\right) \int_{(s,t]} |u_n - u| d\widetilde{\mu}$$

and finally, we have for $n \geq N_{s,t,\ell}$

$$\begin{aligned} |u(s) - \lambda - \Psi(u)((s,t])| \leq |u(s) - u_n(s)| &+ |\Psi_n(u_n)((s,t]) - \Psi(u_n)((s,t])| \\ &+ c_3 \left(\underline{c}^u_{s,t,\ell}, 2\overline{c}^u_{t,L}\right) \int_{(s,t]} |u_n - u| d\widetilde{\mu}. \end{aligned}$$

Let now n go to infinity. The first term of the above upper bound goes to 0 by definition of $u(s)$; the second term goes to 0 by Lemma 4.2. Finally, the last term also goes to 0 using the dominated convergence theorem. Thus u satisfies the first property for $s > \wp$.

To extend this for $s = \wp$, we proceed as in the second step and consider any $\wp < s' \leq t$: then $|u_n(\wp) - u_n(s')| \leq \Delta^u_{t,L}\mu_n(\wp, s']$ and taking the limit $n \to \infty$ gives $|u(\wp) - u(s')| \leq \Delta^u_{t,L}\mu(\wp, s']$. Letting $s' \downarrow \wp$ shows that $u(s') \to u(\wp)$. On the other hand, it is plain that $\lambda + \Psi(u)((s', t]) \to \lambda + \Psi(u)((\wp, t])$ as $s' \downarrow \wp$ and so u satisfies the first property for all $s \in [\wp, t]$. It remains to show uniqueness in order to complete the proof.

Fourth step: uniqueness. Let $\widetilde{u}$ be a function with the same properties than u. Then (4.4) gives

$$\begin{aligned}|u(s) - \widetilde{u}(s)| &= |\Psi(u)((s,t]) - \Psi(\widetilde{u})((s,t])|\\ &\leq c_3\left(\underline{c}_{s,t,\ell} \wedge \inf_{[s,t]} \widetilde{u},\, \bar{c}^u_{t,L} + \sup_{[s,t]} \widetilde{u}\right) \int_{(s,t]} |u - \widetilde{u}| d\widetilde{\mu}\end{aligned}$$

and we conclude that $u = \widetilde{u}$ using Lemma 4.1. □

4.3 Proof of claim III

Fix in the rest of the proof $t \geq 0$, $s \in [\wp(t), t]$, $x_n \to x \in [0, \infty)$, $I \geq 1$, $\lambda_1, \ldots, \lambda_I > 0$ and $s \leq t_1 < \cdots < t_I \leq t$. Consider first the case $I = 2$, so that we must show that

$$\lim_{n\to\infty} \mathbb{E}\left(e^{-\lambda_1 X_n(t_1) - \lambda_2 X_n(t_2)} \mid X_n(s) = x_n\right) = \exp\left(-xu(s, t_1, \lambda_1 + u(t_1, t_2, \lambda_2))\right). \tag{4.5}$$

Using the Markov property of X_n and the definition (2.1) of u_n, we get

$$\begin{aligned}&\mathbb{E}\Big(e^{-\lambda_1 X_n(t_1) - \lambda_2 X_n(t_2)} \mid X_n(s) = x_n\Big)\\ &\qquad = \mathbb{E}\left[e^{-\lambda_1 X_n(t_1)} \mathbb{E}\left(e^{-\lambda_2 X_n(t_2)} \mid X_n(t_1)\right) \mid X_n(s) = x_n\right]\\ &\qquad = \mathbb{E}\left[e^{-\lambda_1 X_n(t_1)} e^{-X_n(t_1) u_n(t_1, t_2, \lambda_2)} \mid X_n(s) = x_n\right]\end{aligned}$$

and so

$$\mathbb{E}\left(e^{-\lambda_1 X_n(t_1) - \lambda_2 X_n(t_2)} \mid X_n(s) = x_n\right) = \exp\left(-x_n u_n(s, t_1, \lambda_1 + u_n(t_1, t_2, \lambda_2))\right). \tag{4.6}$$

Since $\wp$ is an increasing function by Lemma 3.4 and $\wp(t) \leq s \leq t_1 \leq t_2 \leq t$, it holds that $\wp(t_2) \leq t_1 \leq t_2$ and so Lemma 4.3 implies that $u_n(t_1, t_2, \lambda_2) \to u(t_1, t_2, \lambda_2)$. Also, $\wp(t_1) \leq s \leq t_1$ so Lemma 4.3 implies that

$$\lim_{n\to\infty} u_n(s, t_1, \lambda_1 + u_n(t_1, t_2, \lambda_2)) = u(s, t_1, \lambda_1 + u(t_1, t_2, \lambda_2))$$

which proves (4.5) using (4.6). The general case $I \geq 3$ follows in a similar way by induction.

4.4 Proof of claim IV

Fix $t \geq s \geq 0$, $x \geq 0$ and $x_n \to x$: the goal is to show that the sequence $(X_n(y), s \leq y \leq t)$ under $\mathbb{P}(\,\cdot \mid X_n(s) = x_n)$ is tight, and in order to do so we use Theorem 1′ in Bansaye and Simatos [7]. There are two assumptions, A1 and A2', to check.

Assumption A1 is a compact containment condition, and since $[0, \infty]$ endowed with the metric $d(x, y) = |e^{-x} - e^{-y}|$ is compact (in addition to being complete and separable), it is automatically satisfied. Thus we only have to check A2', i.e., we have to show that for each $n \geq 1$, each $s \leq y_0 \leq y \leq t$ with $\mu_n(y_0, y] \leq \Delta^u_{t,2}/2$ and each $x_0 \in [0, \infty]$,

$$\mathbb{E}\left[d(x_0, X_n(y))^2 \mid X_n(y_0) = x_0\right] \leq 2\Delta^u_{t,2}\mu_n(y_0, y]. \tag{4.7}$$

Indeed, in this case assumption A2' is satisfied with $\eta_0 = (\Delta^u_{t,2})^2$, $F^n = 2\Delta^u_{t,2}\mu_n$ and $F = 2\Delta^u_{t,2}\mu$. So let us now prove (4.7). Since ∞ is, by convention, an absorbing state, we only need to prove this inequality for finite x_0. Starting from the left-hand side of (4.7), we get

$$\begin{aligned}\mathbb{E}\big[d(x_0, X_n(y))^2 \mid X_n(y_0) = x_0\big] \\ &= e^{-2x_0} + e^{-x_0 u_n(y_0,y,2)} - 2e^{-x_0 - x_0 u_n(y_0,y,1)} \\ &= e^{-2x_0}\left[2\left(1 - e^{x_0(1-u_n(y_0,y,1))}\right) - \left(1 - e^{x_0(2-u_n(y_0,y,2))}\right)\right].\end{aligned}$$

Since $|1 - e^z| \leq e^{|z|} - 1 \leq 2(e^{|z|} - 1)$ for any $z \in \mathbb{R}$, we obtain further

$$\mathbb{E}\left[d(x_0, X_n(y))^2 \mid X_n(y_0) = x_0\right] \leq 2e^{-2x_0}\left[e^{x_0|1-u_n(y_0,y,1)|} + e^{x_0|2-u_n(y_0,y,2)|} - 2\right].$$

Writing λ as $\lambda = u_n(y, y, \lambda)$, (3.16) gives

$$\begin{aligned}\mathbb{E}\left[d(x_0, X_n(y))^2 \mid X_n(y_0) = x_0\right] &\leq 2e^{-2x_0}\left(e^{x_0\Delta^u_{y,1}\mu_n(y_0,y]} + e^{x_0\Delta^u_{y,2}\mu_n(y_0,y]} - 2\right) \\ &\leq 4e^{-2x_0}\left(e^{x_0\Delta^u_{t,2}\mu_n(y_0,y]} - 1\right)\end{aligned}$$

using for the last inequality that $y \leq t$ and that both maps $z \mapsto \Delta^u_{z,\lambda}$ and $\lambda \mapsto \Delta^u_{z,\lambda}$ are increasing. Further, elementary analysis shows that for any $0 \leq y' \leq 1$

$$\sup_{x' \geq 0}\left(e^{-x'}(e^{x'y'} - 1)\right) = y'\left(1 - y'\right)^{1/y'-1} \leq \frac{1}{e}y'e^{y'} \leq y'.$$

Combining the two last displays with $x' = 2x_0$ and $y' = \Delta^u_{t,2}\mu_n(y_0, y]/2 \leq 1$, we finally get (4.7) which achieves the proof of claim IV.

5 Proof of Propositions 2.3 and 2.4

5.1 Proof of Proposition 2.3

Fix some $t > 0$ and assume that $(\|\alpha_n\|(t), n \geq 1)$ and $(\beta_n(t), n \geq 1)$ are bounded, and that (2.5) holds. In view of Lemma 3.4, in order to prove that $\wp(t) = 0$ it is enough to prove that

$$\liminf_{n\to\infty} \inf_{0\leq y\leq t} u_n(y, t, 1) > 0.$$

The goal is to apply the following lemma to derive a lower bound on $\inf_{0\leq y\leq t} u_n(y, t, 1)$.

Lemma 5.1. *Let $\varepsilon > 0$, $M \geq 0$ and for each $i \geq 0$, $a_i, b_i \geq 0$ such that $a_i^2 - a_i b_i M \geq \varepsilon$. If $(w_i, 0 \leq i \leq I)$ satisfies $w_I > 0$, $0 \leq w_i \leq M$ for $0 \leq i \leq I$ and $w_i \geq w_{i+1}a_i - w_{i+1}^2 b_i$ for $0 \leq i \leq I - 1$, then*

$$w_i \geq \left(\frac{1}{w_I}\pi_{i,I-1} + \sum_{k=i}^{I-1}\pi_{i,k-1}b_k a_k^{-2}\right)^{-1}, \ 0 \leq i \leq I,$$

where $\pi_{i,i-1} = 1$ and $\pi_{i,j} = \prod_{k=i}^{j}\left((1 + b_k^2 M^2\varepsilon^{-1})a_k^{-1}\right)$ for $i \leq j$.

Proof. In the rest of the proof let $\rho_k = b_k^2 M^2/\varepsilon$ and

$$r_i(x) = b_i^2 \frac{x^2}{a_i^2 - a_i b_i x}, \ x \le M.$$

Note that $a_i^2 - a_i b_i M \ge \varepsilon$ by assumption, so $a_i > 0$ and r_i is well-defined and increasing. In particular, $r_i(x) \le r_i(M) \le b_i^2 M^2/\varepsilon = \rho_i$ so that writing

$$a_i w_{i+1} - b_i w_{i+1}^2 = a_i \left(\frac{1 + r_i(w_{i+1})}{w_{i+1}} + \frac{b_i}{a_i} \right)^{-1}$$

we obtain, using $w_i \ge a_i w_{i+1} - b_i w_{i+1}^2$,

$$w_i \ge a_i \left(\frac{1+\rho_i}{w_{i+1}} + \frac{b_i}{a_i} \right)^{-1}.$$

This last inequality can be rewritten as $\overline{w}_i \le (1+\rho_i) a_i^{-1} \overline{w}_{i+1} + b_i a_i^{-2}$ with $\overline{w}_i = 1/w_i$. It follows by induction that $\overline{w}_i \le \overline{w}_I \pi_{i,I-1} + \sum_{k=i}^{I-1} \pi_{i,k-1} b_k / a_k^2$ which is exactly the desired result. □

Let $C = 1/(2\overline{c}_{t,1}^u)$ and

$$\delta = \frac{1}{2} \min \left(\liminf_{n\to\infty} \inf_{0 \le i \le I_n} \mathbb{E}\left(\xi_{i,n}; \xi_{i,n} \le nC + 1\right), \ 1 \right),$$

so that $\delta > 0$ by (2.5). We want to apply the previous lemma for n large enough to $\varepsilon = \delta^2/16$, $M = \overline{c}_{t,1}^u$, $w_{i,n} = u_n(t_i^n, t, 1)$, $I_n = \gamma_n(t)$,

$$a_{i,n} = 1 + (1 + \epsilon_{i,n}(w_{i+1,n})) \int_{[-1/n, C]} x \nu_{i,n}(dx)$$

and

$$b_{i,n} = (1 + \epsilon_{i,n}(w_{i+1,n})) \int_{[-1/n, C]} x^2 \nu_{i,n}(dx).$$

By Lemma B.3, M is finite and by definition, $w_{I_n,n} = 1 > 0$ and $0 \le w_{i,n} \le M$ for all $0 \le i \le I_n$. Since $1 - e^{-x} \ge x - x^2$ for $x \ge -1$ and their exists a finite n_0 such that $1 + \epsilon_{i,n}(w_{i+1,n}) \ge 0$ for all $n \ge n_0$ and $i < I_n$ (by Lemmas B.2 and B.3), we obtain for all $n \ge \max(M, n_0)$ and $i < I_n$

$$(1 + \epsilon_{i,n}(w_{i+1,n})) \int \left(1 - e^{-x w_{i+1,n}}\right) \nu_{i,n}(dx) \ge w_{i+1,n}(a_{i,n} - 1) - w_{i+1,n}^2 b_{i,n}.$$

Note that the left-hand side is by (3.4) equal to $\psi_{i,n}(w_{i+1,n})$, and that $\psi_{i,n}(w_{i+1,n})$ is equal to $w_{i,n} - w_{i+1,n}$ according to the first equality in (3.3) and the composition rule (3.6). Thus we obtain

$$w_{i,n} \ge w_{i+1,n} a_{i,n} - w_{i+1,n}^2 b_{i,n}, \ n \ge \max(M, n_0), i < I_n. \tag{5.1}$$

In order to apply Lemma 5.1 it remains to control the sequences $(a_{i,n})$ and $(b_{i,n})$. By definition of δ, there exists a finite n_1 such that for all $n \ge n_1$ and $i \le I_n$

$$\int_{[-1/n,C]} x \nu_{i,n}(dx) = \mathbb{E}(\xi_{i,n}; \xi_{i,n} \le nC + 1) - \mathbb{P}(\xi_{i,n} \le nC + 1) \ge \delta - 1.$$

Let $\eta > 0$ such that $1 + (1-\eta)(\delta - 1) \ge \delta/2$, and, according to Lemma B.2, there exists n_2 such that $|\epsilon_{i,n}(w_{i+1,n})| \le \eta$ for all $n \ge n_2$ and $i < I_n$. Then from the definition of $a_{i,n}$ (and since $\delta < 1$) it follows that

$$a_{i,n} \ge 1 + (1-\eta)(\delta - 1) \ge \delta/2, \ n \ge \max(n_1, n_2), i \le I_n. \tag{5.2}$$

We now proceed to controlling $a_{i,n}^2 - a_{i,n} b_{i,n} M$. First, we note that $\nu_{i,n}(\{-1/n\}) \leq n$ and

$$\int_{[-1/n,C]} x^2 \nu_{i,n}(dx) = \frac{\nu_{i,n}(\{-1/n\})}{n^2} + \int_{[0,C]} x^2 \nu_{i,n}(dx) \leq \frac{1}{n} + C \int_{[0,C]} x \nu_{i,n}(dx)$$

and so we get $\int_{[-1/n,C]} x^2 \nu_{i,n}(dx) \leq 1/n + C + C \int_{[-1/n,C]} x \nu_{i,n}(dx)$. Then

$$b_{i,n} \leq \frac{2}{n} + C(1 + \epsilon_{i,n}(w_{i+1,n})) + C(a_{i,n} - 1) \leq \kappa_n + C a_{i,n}$$

where $\kappa_n = 2/n + C\bar{c}^{\epsilon}_{n,t}(M)$. In particular, there exists by Lemma B.2 a finite n_3 such that $\kappa_n \leq \delta/(8M)$ for $n \geq n_3$, so that

$$a_{i,n}^2 - a_{i,n} b_{i,n} M \geq a_{i,n}^2 - a_{i,n}(\delta/(8M) + C a_{i,n}) M = \frac{1}{2} a_{i,n}^2 - \frac{\delta}{8} a_{i,n} \geq \frac{\delta^2}{16}$$

for $n \geq \max(n_1, n_2, n_3)$ and $i \leq I_n$ thanks to (5.2). Thus for any $n \geq \max(M, n_0, n_1, n_2, n_3)$, the assumptions of Lemma 5.1 are satisfied and we obtain

$$w_{i,n} \geq \left(\pi_{i,I_n-1,n} + \sum_{k=i}^{I_n-1} \pi_{i,k-1,n} b_{k,n} a_{k,n}^{-2} \right)^{-1}, \; 0 \leq i \leq I_n,$$

where $\pi_{i,i-1,n} = 1$ and $\pi_{i,j,n} = \prod_{k=i}^{j} \left((1 + b_{k,n}^2 M^2 \varepsilon^{-1}) a_{k,n}^{-1} \right)$ for $i \leq j$. To end the proof it remains to show that

$$\limsup_{n \geq 1} \sup_{0 \leq i \leq I_n} \left(\pi_{i,I_n-1,n} + \sum_{k=i}^{I_n-1} \pi_{i,k-1,n} b_{k,n} a_{k,n}^{-2} \right) < \infty. \tag{5.3}$$

Fix some $n \geq \max(M, n_0, n_1, n_2, n_3)$ and $0 \leq i \leq j < I_n$: we derive an upper bound on $\pi_{i,j,n}$, which we write as

$$\pi_{i,j,n} = \exp\left(-\sum_{k=i}^{j} \log a_{k,n} \right) \times \prod_{k=i}^{j} \left(1 + b_{k,n}^2 M^2 \varepsilon^{-1} \right). \tag{5.4}$$

Let in the sequel $C' = \sup \beta_n(t) + \sup \beta_n(t)^3$ and the suprema are taken over $n \geq 1$ and note that

$$\sum_{k=0}^{I_n-1} \beta_{k,n}^2 \leq \sum_{k=0}^{I_n-1} \beta_{k,n} \mathbb{1}_{\{\beta_{i,n} \leq 1\}} + \beta_n(t)^2 \#\{k = 0, \ldots, n-1 : \beta_{i,n} > 1\} \leq C'.$$

Using convexity and $0 \leq b_{k,n} \leq 2(1 + C^2)\beta_{k,n}$, we get

$$\prod_{k=i}^{j} \left(1 + b_{k,n}^2 M^2 \varepsilon^{-1} \right) \leq \exp\left(\sum_{k=i}^{j} b_{k,n}^2 M^2 \varepsilon^{-1} \right) \leq \exp\left(4(1 + C^2)^2 M^2 \varepsilon^{-1} C' \right) \tag{5.5}$$

We now control the sum of the right-hand side of (5.4). Since $n \geq \max(M, n_0)$ we have by (5.2) that $a_{k,n} \geq \delta/2$ for $k < I_n$. In particular, if $\ell = \inf(\log x/(x-1))$ where the infimum is taken over $x \geq \delta/2$, $\ell \in (0, \infty)$ and we have $\log a_{k,n} \geq \ell(a_{k,n} - 1)^-$, with $x^- = \min(x, 0)$. Further,

$$\begin{aligned}
\int_{[-1/n,C]} x \nu_{k,n}(dx) &\geq -\frac{1}{n} \nu_{k,n}(\{-1/n\}) + \int_{[0,C]} \frac{x}{1+x^2} \nu_{k,n}(dx) \\
&= -\frac{1}{n(n^2+1)} \nu_{k,n}(\{-1/n\}) + \alpha_{k,n} - \int_{(C,\infty)} \frac{x}{1+x^2} \nu_{k,n}(dx) \\
&\geq -|\alpha_{k,n}| - (2C^{-1} + 1)\beta_{k,n}.
\end{aligned}$$

Thus $\log a_{k,n} \geq -2\ell|\alpha_{k,n}| - 2\ell(2C^{-1}+1)\beta_{k,n}$ and summing over $k = i, \ldots, j$, we obtain that $\sum_{k=i}^{j} \log a_{k,n}$ is bounded. Recalling (5.5) and (5.4), we get that $\pi_{i,j,n}$ is bounded. Adding that

$$\sum_{k=0}^{I_n-1} \frac{b_{k,n}}{a_{k,n}^2} \leq \frac{8(1+C^2)}{\delta^2} \sum_{k=0}^{I_n-1} \beta_{k,n} \leq \frac{8(1+C^2)}{\delta^2} C',$$

the proof of (5.3) and thus of Proposition 2.3 is finally complete.

5.2 Proof of Proposition 2.4

Let in the rest of the proof $\widetilde{\nu}_n$ be the following increasing, càdlàg function

$$\widetilde{\nu}_n(t) = \int_{\mathbb{R}\times(0,t]} |x|\nu_n(dx\,dy) = n \sum_{i=0}^{\gamma_n(t)-1} \mathbb{E}(|\bar{\xi}_{i,n}|).$$

Assume that (2.7) holds, i.e., $\sup_{n\geq 1} \widetilde{\nu}_n(t) < \infty$: we first show that it implies the two other assumptions. That the sequences $(\|\alpha_n\|(t), n \geq 1)$ and $(\beta_n(t), n \geq 1)$ are bounded comes from (2.7) by summing from $i = 0$ to $\gamma_n(t) - 1$ the two following inequalities:

$$|\alpha_{i,n}| \leq n\mathbb{E}\left(\frac{|\bar{\xi}_{i,n}|}{1+\bar{\xi}_{i,n}^2}\right) \leq n\mathbb{E}\left(|\bar{\xi}_{i,n}|\right) \quad \text{and} \quad \beta_{i,n} = \frac{1}{2}n\mathbb{E}\left(\frac{\bar{\xi}_{i,n}^2}{1+\bar{\xi}_{i,n}^2}\right) \leq \frac{n}{2}\mathbb{E}\left(|\bar{\xi}_{i,n}|\right).$$

We now show that (2.6) also holds. By Lemma B.3 there exists a finite constant C_t such that $u_n(s,y,\lambda) \leq C_t$ for all $y \in [s,t]$ and $\lambda \leq 1$. Further, by Lemma B.2 there exists n_t such that $|\epsilon_{i,n}(v)| \leq 1$ for any $n \geq n_t$, $v \leq C_t$ and $i \leq \gamma_n(t)$. Finally, invoking Lemma 3.1 and using $1 - \exp(-\lambda x) \leq \lambda|x|$ for $x \in \mathbb{R}$ and $\lambda \geq 0$, we get

$$u_n(s,t,\lambda) \leq \lambda + 2 \sum_{i=\gamma_n(s)+1}^{\gamma_n(t)} u_n(t_i^n,t,\lambda) \int |x|\nu_{i-1,n}(dx) = \lambda + \int_{(s,t]} u_n(y,t,\lambda)\widetilde{\nu}_n(dy).$$

Thus Lemma 4.1 implies that $u_n(s,t,\lambda) \leq \lambda + \lambda\widetilde{\nu}_n(s,t]e^{\widetilde{\nu}_n(s,t]}$, and consequently

$$\sup_{n\geq 1, s\leq y\leq t} u_n(s,y,\lambda) \leq \lambda\left[1 + \sup_{n\geq 1}\widetilde{\nu}_n(t)\exp\left(\sup_{n\geq 1}\widetilde{\nu}_n(t)\right)\right]. \tag{5.6}$$

Since $\sup_{n\geq 1} \widetilde{\nu}_n(t)$ is finite, letting $\lambda \to 0$ in (5.6) we see that $\sup u_n(s,y,\lambda) \to 0$ as $\lambda \to 0$, where the supremum is taken over $n \geq 1$ and $s \leq y \leq t$. To see that this implies (2.6), we only have to write for any $A \geq 1$

$$\begin{aligned}\mathbb{P}\left(X_n(y) \geq A \mid X_n(s) = 1\right) &= \mathbb{P}\left(1 - e^{-X_n(y)/A} \geq 1 - 1/e \mid X_n(s) = 1\right) \\ &\leq \frac{1 - e^{-u_n(s,y,1/A)}}{1 - 1/e}.\end{aligned}$$

We now assume that $(\|\alpha_n\|(t), n \geq 1)$ and $(\beta_n(t), n \geq 1)$ are bounded and that (2.6) holds: under these assumptions, we show that there exists $n(k) \to \infty$ such that for every $s < \wp(t)$, $\lim_{n\to\infty} u_{n(k)}(s,t,\lambda) = 0$. First, note that

$$\lim_{\lambda\to 0}\left(\sup_{n\geq 1,\, 0\leq s\leq y\leq t} u_n(s,y,\lambda)\right) = 0. \tag{5.7}$$

Indeed, for any $0 \leq s \leq y \leq t$ and $A > 0$, we can write

$$\begin{aligned}1 - e^{-u_n(s,y,\lambda)} &= \mathbb{E}\left[1 - \exp(-\lambda X_n(y)) \mid X_n(s) = 1\right] \\ &\leq 1 - e^{-\lambda A} + \mathbb{P}(X_n(y) \geq A \mid X_n(s) = 1)\end{aligned}$$

which gives

$$\sup_{n\geq 1, 0\leq s\leq y\leq t} u_n(s,y,\lambda) \leq -\log\left(e^{-\lambda A} - \sup_{n\geq 1, 0\leq s\leq y\leq t} \mathbb{P}(X_n(y)\geq A \mid X_n(s)=1)\right).$$

Letting first $\lambda \to 0$ and then $A \to \infty$ and using (2.6) gives (5.7). Further, Lemma 3.4 guarantees the existence of sequences $(n(k))$ and (y_k) such that $y_k \to \wp(t)$, $n(k) \to \infty$ and $u_{n(k)}(y_k, t, \lambda) \to 0$ as $k \to \infty$. Then, the composition rule (3.6) shows that for every $k \geq 1$ and $s \leq y_k$,

$$u_{n(k)}(s,t,\lambda) = u_{n(k)}\left(s, y_k, u_{n(k)}(y_k,t,\lambda)\right) \leq \sup_{n\geq 1, s\leq y\leq t} u_n\left(s, y, u_{n(k)}(y_k,t,\lambda)\right).$$

Since $u_{n(k)}(y_k, t, \lambda) \to 0$ as $k \to \infty$, (5.7) implies that $\sup\{u_{n(k)}(s,t,\lambda) : s \leq y_k\} \to 0$ which achieves to prove that $u_{n(k)}(s,t,\lambda) \to 0$ for every $s < \wp(t)$.

A Proof of Lemma 2.5

Assume that $q_{i,n} = q_{0,n}$ and that $\gamma_n(t) = \lfloor \Gamma_n t \rfloor$ for some sequence $\Gamma_n \to \infty$. Define $q_n = q_{0,n}$ and $\xi_n = \xi_{0,n}$, and for each $n \geq 1$ let $(\bar{\xi}_n(k), k, \geq 1)$ be i.i.d. random variables distributed as $\bar{\xi}_n$: then we have

$$\alpha_n(t) = n\lfloor \Gamma_n t\rfloor \mathbb{E}\left(\frac{\bar{\xi}_n}{1+\bar{\xi}_n^2}\right),\ \beta_n(t) = \frac{1}{2} n\lfloor \Gamma_n t\rfloor \mathbb{E}\left(\frac{\bar{\xi}_n^2}{1+\bar{\xi}_n^2}\right)$$

and $\nu_n([x,\infty)\times(0,t]) = n\lfloor\Gamma_n t\rfloor \mathbb{P}(\bar{\xi}_n \geq x)$. In this context Assumption 2.1 is therefore equivalent to the following assumption.

Assumption A.1. *There exist $a \in \mathbb{R}$, $b \geq 0$ and a positive, σ-finite measure F with support in $(0,\infty)$ such that*

$$n\Gamma_n \mathbb{E}\left(\frac{\bar{\xi}_n}{1+\bar{\xi}_n^2}\right) \underset{n\to\infty}{\longrightarrow} a,\ n\Gamma_n \mathbb{E}\left(\frac{\bar{\xi}_n^2}{1+\bar{\xi}_n^2}\right) \underset{n\to\infty}{\longrightarrow} b \ \text{ and } \ n\Gamma_n \mathbb{P}(\bar{\xi}_n \geq x) \underset{n\to\infty}{\longrightarrow} F([x,\infty)) \quad \text{(B1)}$$

where the last convergence holds for every $x > 0$ such that $F(\{x\}) = 0$.

Under this assumption we then have $\alpha(t) = at$, $\beta(t) = bt$ and $\nu(dx\,dt) = dtF(dx)$. Thus to prove Lemma 2.5 we have to prove that (B1) is equivalent to the weak convergence of the sum $\sum_{k=1}^{n\Gamma_n} \bar{\xi}_n(k)$.

Assume that (B1) holds. Then by III of Theorem 2.2, the sequence $(X_n, n \geq 1)$ converges in the sense of finite-dimensional distributions. By Grimvall [24, Theorem 3.1], this implies the weak convergence of $\sum_{k=1}^{n\Gamma_n} \bar{\xi}_n(k)$.

Assume now that $\sum_{k=1}^{n\Gamma_n} \bar{\xi}_n(k)$ converges weakly. Then Theorem 1 of § 25 in Gnedenko and Kolmogorov [22] immediately gives the existence of F with $\int(1\wedge x^2)F(dx) < \infty$ such that the last convergence in (B1) holds. Let us prove the two first convergences of (B1). In the rest of the proof let $G \subset (0,\infty)$ denote the set of continuity points of F, for $\kappa = 1$ or 2 let $m_\kappa(x) = |x|^\kappa/(1+x^2)$ and fix some $\kappa \in \{1,2\}$.

Since $n\Gamma_n\mathbb{P}(\bar{\xi}_n > x) \to F((x,\infty))$ for $x \in G$, it follows that

$$n\Gamma_n \mathbb{E}\left(m_\kappa(\bar{\xi}_n); |\bar{\xi}_n| > \varepsilon\right) \underset{n\to\infty}{\longrightarrow} \int_{x>\varepsilon} m_\kappa(x) F(dx)$$

for any $\varepsilon \in G$. This can for instance be seen by considering the weak convergence of the random variables $\bar{\xi}_n$ conditioned on $|\bar{\xi}_n| > \varepsilon$. Moreover, according to Corollary 15.16 in

Kallenberg [30], there exists a finite number d_κ such that $n\Gamma_n \mathbb{E}(\xi_n^\kappa; |\bar\xi_n| \le \varepsilon) \to d_\kappa + L_\kappa(\varepsilon)$ for any $\varepsilon \le 1$ in G, and where $L_1(\varepsilon) = -\int_{\varepsilon < x \le 1} xF(dx)$ and $L_2(\varepsilon) = \int_{x \le \varepsilon} x^2 F(dx)$. Note in particular that

$$\sup_{n\ge 1} n\Gamma_n \mathbb{E}\left(m_2(\bar\xi_n)\right) < \infty,$$

since

$$\begin{aligned} n\Gamma_n \mathbb{E}\left(m_2(\bar\xi_n)\right) &= n\Gamma_n \mathbb{E}\left(m_2(\bar\xi_n); \bar\xi_n \le \varepsilon\right) + n\Gamma_n \mathbb{E}\left(m_2(\bar\xi_n); \bar\xi_n > \varepsilon\right) \\ &\le n\Gamma_n \mathbb{E}\left(\bar\xi_n^2; \bar\xi_n \le \varepsilon\right) + n\Gamma_n \mathbb{E}\left(m_2(\bar\xi_n); |\bar\xi_n| > \varepsilon\right) \end{aligned}$$

(note that $\bar\xi_n > \varepsilon \Leftrightarrow |\bar\xi_n| > \varepsilon$ for n large enough, since $\bar\xi_n \ge -1/n$). We now write

$$\begin{aligned} \mathbb{E}\left(\frac{\bar\xi_n^\kappa}{1+\bar\xi_n^2}\right) &= \mathbb{E}\left(\frac{\bar\xi_n^\kappa}{1+\bar\xi_n^2} - \bar\xi_n^\kappa; |\bar\xi_n| \le \varepsilon\right) + \mathbb{E}\left(\frac{\bar\xi_n^\kappa}{1+\bar\xi_n^2}; |\bar\xi_n| > \varepsilon\right) + \mathbb{E}\left(\bar\xi_n^\kappa; |\bar\xi_n| \le \varepsilon\right) \\ &= -\mathbb{E}\left(\frac{\bar\xi_n^{\kappa+2}}{1+\bar\xi_n^2}; |\bar\xi_n| \le \varepsilon\right) + \mathbb{E}\left(m_\kappa(\bar\xi_n); |\bar\xi_n| > \varepsilon\right) + \mathbb{E}\left(\bar\xi_n^\kappa; |\bar\xi_n| \le \varepsilon\right) \end{aligned}$$

to obtain

$$\begin{aligned} &\left| n\Gamma_n \mathbb{E}\left(\frac{\bar\xi_n^\kappa}{1+\bar\xi_n^2}\right) - n\Gamma_n \mathbb{E}\left(m_\kappa(\bar\xi_n); |\bar\xi_n| > \varepsilon\right) - n\Gamma_n \mathbb{E}\left(\bar\xi_n^\kappa; |\bar\xi_n| \le \varepsilon\right) \right| \\ &\qquad \le \varepsilon^\kappa n\Gamma_n \mathbb{E}\left(m_2(\bar\xi_n)\right) \le \varepsilon^\kappa \sup_{n\ge 1} n\Gamma_n \mathbb{E}\left(m_2(\bar\xi_n)\right). \end{aligned}$$

Letting first $n \to \infty$ with $\varepsilon \in G$ and then $\varepsilon \to 0$, we thus obtain

$$\lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left| n\Gamma_n \mathbb{E}\left(\frac{\bar\xi_n^\kappa}{1+\bar\xi_n^2}\right) - \int_{x>\varepsilon} m_\kappa(x) F(dx) - d_\kappa - L_\kappa(\varepsilon) \right| = 0.$$

For $\kappa = 2$ we have

$$\int_{x>\varepsilon} m_2(x) F(dx) - d_2 - L_2(\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} \int m_2(x) F(dx)$$

which is finite, while for $\kappa = 1$ we have by definition of L_1

$$\int_{x>\varepsilon} m_1(x) F(dx) - d_1 - L_1(\varepsilon) = \int_{\varepsilon < x \le 1} (m_1(x) - x) F(dx) - d_1 + \int_{x>1} m_1(x) F(dx).$$

Since $m_1(x) - x \sim -x^2$ as $x \to 0$, $\int_{\varepsilon < x \le 1} (m_1(x) - x) F(dx)$ converges by the dominated convergence theorem to $\int_{x\le 1} (m_1(x) - x) F(dx)$ as $\varepsilon \to 0$, which is finite since $\int (1 \wedge x^2) F(dx) < \infty$. This completes the proof.

B Proof that the key constants and functions are finite

For $x \in \mathbb{R}$ let

$$\Phi(x) = \frac{e^{-x} - 1 + x}{x^2} \tag{B.1}$$

with $\Phi(0) = 1/2$. If $\Phi_2(x) = 1/2 - \Phi(x)$, note that we can rewrite

$$g(x,\lambda) = \frac{x^2}{1+x^2}\left(1 - e^{-\lambda x} - \lambda^2 \Phi(\lambda x)\right) \text{ and } h(x,\lambda) = \frac{x^2}{1+x^2}\left(1 - e^{-\lambda x} + \lambda^2 \Phi_2(\lambda x)\right). \tag{B.2}$$

Lemma B.1 (Control of c_1, c_2 and c_3)**.** *For any $C \geq 0$ and $0 < \eta < T$, the constants $c_1(C)$, $c_2(\eta, T)$ and $c_3(\eta, T)$ are finite.*

Proof. Since $\lim_{x\to\infty} \Phi(x) = 0$, Φ is bounded on $[-C, \infty)$ for any $C > 0$ and so the constant $c_1'(C)$ of (3.8) is finite in view of (B.2). In particular $c_1(C)$ is finite for every $C \geq 0$. Let $0 < \eta < T$ and $\eta \leq y, y' \leq T$, and $x \geq 0$: $c_2(\eta, T)$ is finite because

$$\left| \frac{h(x,y) - h(x,y')}{(y-y')x^2/(1+x^2)} \right| \leq \sup_{\eta \leq v \leq T} |H_x'(v)|,$$

with $H_x(y) = h(x,y)(1+x^2)/x^2$. One can compute

$$H_x'(y) = xe^{-yx} + \frac{e^{-xy} - 1 + xy}{x}$$

and prove that this function is bounded for $x \geq 0$ and $y \in [\eta, T]$, since

$$xe^{-xy} \leq \frac{1}{\eta} \sup_{v \geq 0} (ve^{-v})$$

and for $x \leq 1$,

$$\left| \frac{e^{-xy} - 1 + xy}{x} \right| = \left| \frac{x}{y^2} \Phi(xy) \right| \leq \frac{1}{\eta^2} \sup_{v\geq 0} |\Phi(v)|,$$

while for $x \geq 1$,

$$\left| \frac{e^{-xy} - 1 + xy}{x} \right| \leq \sup_{v \geq 0} |e^{-v} - 1| + T.$$

We get the finiteness of $c_2(\eta, T)$, and hence of $c_3(\eta, T)$. □

Lemma B.2 (Control of $\overline{c}^{\epsilon}_{n,t}(C)$)**.** *Fix $t \geq 0$ and assume that the sequences $(\|\alpha_n\|(t), n \geq 1)$ and $(\beta_n(t), n \geq 1)$ are bounded. Then for any $C \geq 0$, we have $\overline{c}^{\epsilon}_{n,t}(C) \to 0$ as n goes to infinity.*

Proof. Fix t and $C \geq 0$ and define

$$I_{t,C} = \sup \left\{ \left| \int (1 - e^{-\lambda x}) \nu_{i,n}(dx) \right| : 1 \leq n, 0 \leq i < \gamma_n(t), 0 \leq \lambda \leq C \right\}.$$

Then (3.15) entails

$$\begin{aligned} I_{t,C} &\leq c_1(C) \sup \left\{ \mu_n(t_i^n, t_{i+1}^n] : n \geq 1, 0 \leq i < \gamma_n(t) \right\} \\ &\leq c_1(C) \sup_{n \geq 1} \left(\sum_{i=0}^{\gamma_n(t)-1} \mu_n(t_i^n, t_{i+1}^n] \right) = c_1(C) \sup_{n\geq 1} \mu_n(t). \end{aligned}$$

Since $\mu_n(t) = \|\alpha_n\|(t) + \beta_n(t)$ the sequence $(\mu_n(t))$ is bounded by assumption, showing that $I_{t,C}$ is finite. It follows from the definitions (3.3) and (3.4) of $\psi_{i,n}$ and $\epsilon_{i,n}$ that for any $i \geq 0$

$$\epsilon_{i,n}(\lambda) = \frac{-\log\left(1 - \frac{1}{n}\int (1 - e^{-\lambda x}) \nu_{i,n}(dx)\right) - \frac{1}{n}\int (1 - e^{-\lambda x})\nu_{i,n}(dx)}{\frac{1}{n}\int (1 - e^{-\lambda x}) \nu_{i,n}(dx)}$$

and so

$$\overline{c}^{\epsilon}_{n,t}(C) \leq \sup_{|x| \leq I_{t,C}/n} \left| \frac{-\log(1-x) - x}{x} \right|.$$

Letting $n \to \infty$ achieves the proof of the result. □

Lemma B.3 (Control of $\bar{c}^u_{t,\lambda}$)**.** *Fix $t \geq 0$ and assume that the two sequences $(\|\alpha_n\|(t), n \geq 1)$ and $(\beta_n(t), n \geq 1)$ are bounded. Then for any $\lambda \geq 0$ the constant $\bar{c}^u_{t,\lambda}$ is finite and moreover*

$$\sup\left\{\bar{c}^u_{s,\lambda} : 0 \leq s \leq t, \lambda \leq 1\right\} < \infty.$$

Proof. In the rest of the proof fix t and $\lambda \geq 0$, define $B_t = 2\sup_{n\geq 1}\mu_n(t)$, which is finite by assumption, and $C_{t,\lambda} = (\lambda+2)(1+B_t)e^{B_t}$. Following Lemma B.2 choose $n_{t,\lambda} \geq 1$ such that $\bar{c}^\epsilon_{n,t}(C_{t,\lambda}) \leq 1$ for all $n \geq n_{t,\lambda}$. Since $Z_{i,n}$ is finite for each $i \geq 0$ and $n \geq 1$, it follows that $\sup_{0\leq s\leq t} u_n(s,t,\lambda)$ is finite for each $n \geq 1$. Thus, to conclude, it is enough to get that

$$\sup\left\{u_n(s,t,\lambda) : n \geq n_{t,\lambda}, 0 \leq s \leq t\right\} = \sup\left\{u_n(t^n_i, t^n_{\gamma_n(t)}, \lambda) : n \geq n_{t,\lambda}, 0 \leq i \leq \gamma_n(t)\right\}$$

is finite. To prove that, In the rest of the proof fix $n \geq n_{t,\lambda}$ and define $a_i = u_n(t^n_i, t^n_{\gamma_n(t)}, \lambda)$. We prove by backwards induction on i that $a_i \leq C_{t,\lambda}$ for all $0 \leq i \leq \gamma_n(t)$, and since the bound does not depend on n or i this will show the result. We have $a_{\gamma_n(t)} = \lambda \leq C_{t,\lambda}$ so the initialization is satisfied. Now consider some $1 \leq i < \gamma_n(t)$ and assume that $a_k \leq C_{t,\lambda}$ for all $i \leq k \leq \gamma_n(t)$: we prove that $a_{i-1} \leq C_{t,\lambda}$.

Fix some $i < k \leq \gamma_n(t)$. By definition, we have

$$\psi_{k-1,n}(a_k) = (1 + \epsilon_{k-1,n}(a_k))\left(a_k\alpha_{k-1,n} + \int g(x,a_k)\nu_{k-1,n}(dx)\right).$$

By induction hypothesis, it holds that $a_k \leq C_{t,\lambda}$. Combined with $\bar{c}^\epsilon_{n,t}(C_{t,\lambda}) \leq 1$, this gives $0 \leq 1 + \epsilon_{k-1,n}(a_k) \leq 2$. Together with the inequality $g(x,y) \leq x^2/(1+x^2)$ (note that $\Phi \geq 0$), we finally get

$$\psi_{k-1,n}(a_k) \leq (1+\epsilon_{k-1,n}(a_k))\left(a_k|\alpha_{k-1,n}| + 2\beta_{k-1,n}\right) \leq 2(a_k+2)\mu_n(t^n_{k-1}, t^n_k].$$

Hence for any $i-1 \leq j \leq \gamma_n(t)$, this gives together with Lemma 3.1 for the first equality

$$a_j = \lambda + \sum_{k=j+1}^{\gamma_n(t)} \psi_{k-1,n}(a_k) \leq \lambda + \sum_{k=j+1}^{\gamma_n(t)} 2(a_k+2)\mu_n(t^n_{k-1}, t^n_k].$$

This can be rewritten $a'_j \leq A + S_{j+1}$ if $a'_k = a_k + 2$, $A = \lambda + 2$, $d_k = 2\mu_n(t^n_{k-1}, t^n_k]$ and $S_k = d_k a'_k + \cdots + d_{\gamma_n(t)}a'_{\gamma_n(t)}$. This gives for $j = i-1$

$$a'_{i-1} \leq A + S_i = A + d_i a'_i + S_{i+1} \leq A + d_i(A + S_{i+1}) + S_{i+1} = (1+d_i)(A+S_{i+1}).$$

Then by induction one gets

$$a'_{i-1} \leq (1+d_i)\cdots(1+d_{\gamma_n(t)-1})(A + S_{\gamma_n(t)}) \leq \exp\left(d_1 + \cdots + d_{\gamma_n(t)}\right)\left(A + d_{\gamma_n(t)}a'_{\gamma_n(t)}\right).$$

Since $a'_{\gamma_n(t)} = A = \lambda + 2$ and $d_{\gamma_n(t)} \leq d_1 + \cdots + d_{\gamma_n(t)} = 2\mu_n(t) \leq B_t$, this shows that $a_{i-1} \leq C_{t,\lambda}$ which achieves the proof of the induction and shows that $\bar{c}^u_{t,\lambda} \leq C_{t,\lambda}$. This gives the finiteness of $\bar{c}^u_{t,\lambda}$, and since $C_{t,\lambda}$ is increasing in both t and λ, for any $s \leq t$ and $\lambda \leq 1$ we obtain $\bar{c}^u_{s,\lambda} \leq C_{t,1}$ which gives the second part of the lemma. □

C Proof of Lemma 4.2

This appendix is devoted to the proof of Lemma 4.2. Recall the function $\mu = \|\alpha\| + \beta$ defined at the beginning of Section 4. We will use the following simple result.

Lemma C.1. *For any $\varepsilon > 0$ and $0 \leq s < t$, there exists a partition of the interval $(s,t]$ as*

$$(s,t] = \left(\bigcup_{j=1}^{J}(a_j, b_j]\right) \cup \left(\bigcup_{k=1}^{K}(a'_k, b'_k]\right)$$

such that $\{b'_k, 1 \le k \le K\} = (s,t] \cap \{v \ge 0 : \Delta\mu(v) \ge \varepsilon\}$, $\mu(a_j, b_j] \le \varepsilon$ *for each* $1 \le j \le J$ *and* $\mu(a'_k, b'_k) \le \varepsilon/K$ *for each* $1 \le k \le K$.

In the rest of the proof fix $t, \lambda > 0$, $\wp(t) < s \le t$, (ℓ_n) a sequence converging to λ and let $u_n(y) = u_n(y, t, \ell_n)$. With this notation, we have

$$R_n(y) = |\Psi_n(u_n)((y,t]) - \Psi(u_n)((y,t])|\,,\ 0 \le y \le t.$$

Let $\ell = \inf_{n\ge 1} \ell_n$ and $L = \sup_{n\ge 1} \ell_n$ and assume without loss of generality, since $\ell_n \to \lambda > 0$, that $\ell > 0$. We first show that $R_n(s) \to 0$, the fact that $\sup\{R_n(y) : s \le y \le t, n \ge 1\}$ is finite is proved in Section C.3. From the definitions of Ψ and Ψ_n one can write

$$|\Psi_n(u_n)((s,t]) - \Psi(u_n)((s,t])| \le B_n^\alpha + B_n^\beta + B_n^\nu + B_n^\epsilon$$

with

$$B_n^\alpha = \left| \int_{(s,t]} u_n d\alpha_n - \int_{(s,t]} u_n d\alpha \right|, \quad B_n^\beta = \left| \int_{(s,t]} u_n^2 d\beta_n - \int_{(s,t]} u_n^2 d\beta \right|,$$

$$B_n^\nu = \left| \int_{[-1/n,\infty)\times(s,t]} h(x, u_n(y))\nu_n(dx\,dy) - \int_{(0,\infty)\times(s,t]} h(x, u_n(y))\nu(dx\,dy) \right|$$

and

$$B_n^\epsilon = \sum_{i=\gamma_n(s)+1}^{\gamma_n(t)} |\epsilon_{i-1,n}(u_n(t_i^n))| \left| \int \left(1 - e^{-x u_n(t_i^n)}\right) \nu_{i-1,n}(dx) \right|.$$

We will show that each sequence (B_n^α), (B_n^β), (B_n^ν) and (B_n^ϵ) goes to 0 as n goes to infinity. By (3.15) and by definition of the constants $\overline{c}_{n,t}^\epsilon$, $\overline{c}_{t,L}^u$ and c_1, one can derive similarly as in the proof of (3.16)

$$B_n^\epsilon \le \overline{c}_{n,t}^\epsilon(\overline{c}_{t,\ell_n}^u) c_1(\overline{c}_{t,\ell_n}^u)\mu_n(t) \le \overline{c}_{n,t}^\epsilon(\overline{c}_{t,L}^u) c_1(\overline{c}_{t,L}^u)\mu_n(t)$$

where the last inequality follows from the fact that $\ell_n \le L$ and that the functions $\overline{c}_{n,t}^\epsilon(C)$ and $\overline{c}_{t,y}^u$ are increasing in C and y, respectively. From now on, we will use such monotonicity properties without further comment. This last upper bound is seen to go 0, invoking (4.1) and Lemmas B.2 and B.3. Thus the sequence (B_n^ϵ) goes to 0 and we have to control the three other sequences (B_n^α), (B_n^β) and (B_n^ν). We control the two first sequences in Section C.1 and the last one in Section C.2

C.1 Control of the sequences (B_n^α) and (B_n^β)

We treat in detail the convergence of (B_n^α) to 0. For (B_n^β), one essentially needs to replace α by β and u_n by u_n^2, we mention along the way what modifications need to be done.

Fix $\varepsilon > 0$ and consider the partition $((a_j, b_j], 1 \le j \le J)$ and $((a'_k, b'_k], 1 \le k \le K)$ of $(s,t]$ provided by Lemma C.1. Note that the partition depends on s, t and ε but not on n. We can write $B_n^\alpha \le \sum_{j=1}^J B_{n,j}^{\alpha,1} + \sum_{k=1}^K (B_{n,k}^{\alpha,2} + B_{n,k}^{\alpha,3})$ with

$$B_{n,j}^{\alpha,1} = \left| \int_{(a_j,b_j]} u_n d\alpha_n - \int_{(a_j,b_j]} u_n d\alpha \right|, \quad B_{n,k}^{\alpha,2} = \int_{(a'_k,b'_k)} u_n d\|\alpha_n\| + \int_{(a'_k,b'_k)} u_n d\|\alpha\|$$

and $B_{n,k}^{\alpha,3} = u_n(b'_k)\left|\alpha_{\gamma_n(b'_k),n} - \Delta\alpha(b'_k)\right|$. For $B_{n,j}^{\alpha,1}$ we have

$$B_{n,j}^{\alpha,1} \le \int_{(a_j,b_j]} |u_n(y) - u_n(b_j)|\,\|\alpha_n\|(dy) + \int_{(a_j,b_j]} |u_n(y) - u_n(b_j)|\,\|\alpha\|(dy) + u_n(b_j)\,|\alpha_n(a_j,b_j] - \alpha(a_j,b_j]|\,.$$

By (3.16), $|u_n(y) - u_n(b_j)| \leq \Delta^u_{t,L}\mu_n(a_j, b_j]$ for all $y \in (a_j, b_j]$ and so, using also $u_n(b_j) \leq \bar{c}^u_{t,L}$, we get

$$B^{\alpha,1}_{n,j} \leq \Delta^u_{t,L}\mu_n(a_j, b_j]\big(\|\alpha_n\|(a_j, b_j] + \|\alpha\|(a_j, b_j]\big) + \bar{c}^u_{t,L}\left|\alpha_n(a_j, b_j] - \alpha(a_j, b_j]\right|.$$

For $B^{\beta,1}_n$ one needs to use

$$\left|u_n(y)^2 - u_n(b_j)^2\right| = |u_n(y) - u_n(b_j)|\,(u_n(y) + u_n(b_j)) \leq 2\bar{c}^u_{t,L}\Delta^u_{t,L}\mu_n(a_j, b_j],$$

which leads to a similar upper bound. Since the partition does not depend on n, we have $\alpha_n(a_j, b_j] \to \alpha(a_j, b_j]$ and $\mu_n(a_j, b_j] \to \mu(a_j, b_j]$ by (A1), so that summing over $j = 1, \ldots, J$, letting n go to infinity and using $\|\alpha\|(A) \leq \mu(A)$ gives

$$\limsup_{n\to\infty} \sum_{j=1}^{J} B^{\alpha,1}_{n,j} \leq 2\Delta^u_{t,L} \sum_{j=1}^{J} \big(\mu(a_j, b_j]\big)^2 \leq 2\varepsilon\Delta^u_{t,L}\mu(s,t], \tag{C.1}$$

using also $\mu(a_j, b_j] \leq \varepsilon$, which holds by choice of the partition, to derive the second inequality. To upper bound $B^{\alpha,2}_{n,k}$ we write $B^{\alpha,2}_{n,k} \leq \bar{c}^u_{t,L}\left(\|\alpha_n\|(a'_k, b'_k) + \|\alpha\|(a'_k, b'_k)\right)$ which leads, using $\mu(a'_k, b'_k) \leq \varepsilon/K$, to

$$\limsup_{n\to\infty} \sum_{k=1}^{K} B^{\alpha,2}_{n,k} \leq 2\bar{c}^u_{t,L} \sum_{k=1}^{K} \mu(a'_k, b'_k) \leq 2\varepsilon\bar{c}^u_{t,L}. \tag{C.2}$$

For $B^{\beta,2}_{n,k}$ one can use $B^{\beta,2}_{n,k} \leq (\bar{c}^u_{t,L})^2\left(\beta_n(a'_k, b'_k) + \beta(a'_k, b'_k)\right)$ to obtain a similar upper bound. Finally, for $B^{\alpha,3}_{n,k}$ one has $B^{\alpha,3}_{n,k} \leq \bar{c}^u_{t,L}|\alpha_{\gamma_n(b'_k),n} - \Delta\alpha(b'_k)|$ which goes to 0 by (A2). One can similarly write $B^{\beta,3}_{n,k} \leq (\bar{c}^u_{t,L})^2|\beta_{\gamma_n(b'_k),n} - \Delta\beta(b'_k)|$ for $B^{\beta,3}_{n,k}$. Since K does not depend on n this gives $\sum_{k=1}^{K} B^{\alpha,3}_{n,k} \to 0$ and so (C.1) and (C.2) give

$$\limsup_{n\to\infty} B^{\alpha}_n \leq 2\varepsilon\left(\Delta^u_{t,L}\mu(s,t] + \bar{c}^u_{t,L}\right).$$

Since ε was arbitrary, letting $\varepsilon \to 0$ gives the result.

C.2 Control of the sequence (B^{ν}_n)

For $T \geq 0$ we define the constant

$$c_4(T) = \sup\left\{\left|\frac{h(x,y)}{x^3/(1+x^2)}\right| : x \geq -1, 0 \leq y \leq T\right\} \tag{C.3}$$

which, starting from (B.2), can be seen to be finite. For $d > 0$ we write

$$B^{\nu}_n \leq \widetilde{B}^{\nu}_n + \hat{B}^{\nu}_n + \check{B}^{\nu}_n \tag{C.4}$$

with

$$\widetilde{B}^{\nu}_n = \left|\int_{[d,\infty)\times(s,t]} h(x, u_n(y))\nu_n(dx\,dy) - \int_{[d,\infty)\times(s,t]} h(x, u_n(y))\nu(dx\,dy)\right|.$$

$$\hat{B}^{\nu}_n = \int_{[-1/n,d)\times(s,t]} |h(x, u_n(y))|\nu_n(dx\,dy), \quad \check{B}^{\nu}_n = \int_{(0,d)\times(s,t]} |h(x, u_n(y))|\nu(dx\,dy).$$

Note that $\widetilde{B}^{\nu}_n$ depends on d but, similarly as t or λ, we do not reflect this in the notation because d will be fixed once and for all shortly. Bounding the two last terms thanks to (C.3), we have

$$B^{\nu}_n \leq \widetilde{B}^{\nu}_n + c_4(\bar{c}^u_{t,L})\left(\int_{(0,d)\times(0,t]} \frac{x^3}{1+x^2}\nu(dx\,dy) + \int_{[-1/n,d)\times(0,t]} \frac{|x|^3}{1+x^2}\nu_n(dx\,dy)\right).$$

Since $\int_{(0,\infty)\times(0,t]}(1\wedge x^2)\nu(dx\,dy)$ is finite, we have $\int_{(0,d)\times(0,t]}\frac{x^3}{1+x^2}\nu(dx\,dy)\to 0$ as $d\to 0$. Moreover, proceeding similarly as for the proof of (4.3), we can show that

$$\lim_{d\to 0}\limsup_{n\to\infty}\int_{[-1/n,d)\times(0,t]}\frac{|x|^3}{1+x^2}\nu_n(dx\,dy)=0.$$

Thus letting first $n\to\infty$ and then $d\to 0$, we obtain

$$\limsup_{n\to\infty}B_n^\nu\le\lim_{d\to 0}\limsup_{n\to\infty}\widetilde{B}_n^\nu.$$

Hence to prove $B_n^\nu\to 0$ we only have to show that $\widetilde{B}_n^\nu\to 0$ for every $d>0$. So in the rest of this step we fix an arbitrary $d>0$ and show that $\widetilde{B}_n^\nu\to 0$. Fix $\varepsilon>0$ and consider the partition $((a_j,b_j],1\le j\le J)$ and $((a_k',b_k'],1\le k\le K)$ of $(s,t]$ given by Lemma C.1, which does not depend on n. Then we can write $\widetilde{B}_n^\nu\le\sum_{j=1}^J\widetilde{B}_{n,j}^{\nu,1}+\sum_{k=1}^K(\widetilde{B}_{n,k}^{\nu,2}+\widetilde{B}_{n,k}^{\nu,3})$ with

$$\widetilde{B}_{n,j}^{\nu,1}=\left|\int_{[d,\infty)\times(a_j,b_j]}h(x,u_n(y))\nu_n(dx\,dy)-\int_{[d,\infty)\times(a_j,b_j]}h(x,u_n(y))\nu(dx\,dy)\right|,$$
$$\widetilde{B}_{n,k}^{\nu,2}=\int_{[d,\infty)\times(a_k,b_k')}|h(x,u_n(y))|\nu_n(dx\,dy)+\int_{[d,\infty)\times(a_k,b_k')}|h(x,u_n(y))|\nu(dx\,dy)$$

and

$$\widetilde{B}_{n,k}^{\nu,3}=\left|\int_{[d,\infty)}h(x,u_n(b_k'))\nu_{\gamma_n(b_k'),n}(dx)-\int_{[d,\infty)\times\{b_k'\}}h(x,u_n(b_k'))\nu(dx\,dy)\right|.$$

Further we write $\widetilde{B}_{n,j}^{\nu,1}\le\widetilde{B}_{n,j}^{\nu,4}+\widetilde{B}_{n,j}^{\nu,5}$ with

$$\widetilde{B}_{n,j}^{\nu,4}=\int_{[d,\infty)\times(a_j,b_j]}\left|h(x,u_n(y))-h(x,u_n(b_j))\right|\nu_n(dx\,dy)$$
$$+\int_{[d,\infty)\times(a_j,b_j]}\left|h(x,u_n(y))-h(x,u_n(b_j))\right|\nu(dx\,dy)$$

and

$$\widetilde{B}_{n,j}^{\nu,5}=\left|\int_{[d,\infty)\times(a_j,b_j]}h(x,u_n(b_j))\nu_n(dx\,dy)-\int_{[d,\infty)\times(a_j,b_j]}h(x,u_n(b_j))\nu(dx\,dy)\right|.$$

We derive, in order, upper bounds on $\widetilde{B}_{n,k}^{\nu,2}$, $\widetilde{B}_{n,j}^{\nu,4}$, $\widetilde{B}_{n,k}^{\nu,3}$ and finally on $\widetilde{B}_{n,j}^{\nu,5}$.

To control $\widetilde{B}_{n,k}^{\nu,2}$ we introduce the constant

$$c_5(T)=\sup\left\{\frac{|h(x,y)|}{x^2/(1+x^2)}:0\le y\le T,x\ge 0\right\}$$

which can be seen to be finite, starting from instance from (B.2). Thus

$$\widetilde{B}_{n,k}^{\nu,2}\le c_5(\bar{c}_{t,L}^u)\left(\int_{[d,\infty)\times(a_k',b_k')}\frac{x^2}{1+x^2}\nu_n(dx\,dy)+\int_{[d,\infty)\times(a_k,b_k')}\frac{x^2}{1+x^2}\nu(dx\,dy)\right)$$
$$\le 2c_5(\bar{c}_{t,L}^u)\left(\beta_n(a_k',b_k')+\beta(a_k',b_k')\right)$$

using (4.2) for the last inequality. Using $\beta_n(a_k',b_k')\to\beta(a_k',b_k')\le\mu(a_k',b_k')\le\varepsilon/K$ (the convergence $\beta_n(a_k',b_k')\to\beta(a_k',b_k')$ comes from Assumptions (A1) and (A2) by writing

$\beta_n(a'_k, b'_k) = \beta_n(a'_k, b'_k] - \Delta\beta_n(b'_k)$ and observing that b'_k is by construction an atom of μ), this leads to

$$\limsup_{n\to\infty} \sum_{k=1}^{K} \widetilde{B}_{n,k}^{\nu,2} \leq 4\varepsilon c_5(\bar{c}_{t,L}^u). \tag{C.5}$$

To derive an upper bound on $\widetilde{B}_{n,j}^{\nu,4}$, we use the constant $c_2(\eta, T)$ defined in (3.9). Since $0 < \underline{c}_{s,t,\ell}^u \leq u_n(y) \leq \bar{c}_{t,L}^u$ for $n \geq N_{s,t,\ell}$ and $a_j < y \leq b_j$, we have for such n

$$\int_{[d,\infty)\times(a_j,b_j]} |h(x,u_n(y)) - h(x,u_n(b_j))| \, \nu_n(dx\,dy) \\ \leq c_2\left(\underline{c}_{s,t,\ell}^u, \bar{c}_{t,L}^u\right) \int_{[d,\infty)\times(a_j,b_j]} |u_n(y) - u_n(b_j)| \, \frac{x^2}{1+x^2} \nu_n(dx\,dy).$$

Since $|u_n(y) - u_n(b_j)| \leq \Delta_{t,L}^u \, \mu_n(a_j, b_j]$ for $a_j < y \leq b_j$ by (3.16), we obtain

$$\int_{[d,\infty)\times(a_j,b_j]} |h(x,u_n(y)) - h(x,u_n(b_j))| \, \nu_n(dx\,dy) \\ \leq c_2\left(\underline{c}_{s,t,\ell}^u, \bar{c}_{t,L}^u\right) \Delta_{t,L}^u \, \mu_n(a_j, b_j] \int_{[d,\infty)\times(a_j,b_j]} \frac{x^2}{1+x^2} \nu_n(dx\,dy).$$

Since

$$\int_{[d,\infty)\times(a_j,b_j]} \frac{x^2}{1+x^2} \nu_n(dx\,dy) \leq 2\beta_n(a_j, b_j] \leq 2\mu_n(a_j, b_j], \tag{C.6}$$

we finally get

$$\int_{[d,\infty)\times(a_j,b_j]} |h(x,u_n(y)) - h(x,u_n(b_j))| \, \nu_n(dx\,dy) \leq C_{s,t,\ell,L} (\mu_n(a_j, b_j])^2$$

with $C_{s,t,\ell,L} = 2c_2(\underline{c}_{s,t,\ell}^u, \bar{c}_{t,L}^u)\Delta_{t,L}^u$. The exact same reasoning with ν instead of ν_n, using the inequality (4.2) instead of (C.6), leads to

$$\widetilde{B}_{n,j}^{\nu,4} \leq C_{s,t,\ell,L} \left[(\mu_n(a_j, b_j])^2 + (\mu(a_j, b_j])^2 \right].$$

Hence (4.1) gives

$$\limsup_{n\to\infty} \sum_{j=1}^{J} \widetilde{B}_{n,j}^{\nu,4} \leq 2C_{s,t,\ell,L} \sum_{j=1}^{J} (\mu(a_j, b_j])^2 \leq 2\varepsilon C_{s,t,\ell,L} \mu(s,t] \tag{C.7}$$

using $\mu(a_j, b_j] \leq \varepsilon$ to get the second inequality.

The arguments to control $\widetilde{B}_{n,k}^{\nu,3}$ and $\widetilde{B}_{n,j}^{\nu,5}$ are very similar: we treat the case $\widetilde{B}_{n,j}^{\nu,5}$ in detail and mention necessary changes needed for $\widetilde{B}_{n,k}^{\nu,3}$. We need the constant c_6

$$c_6(T) = \sup_{\substack{0\leq y\leq T \\ 0\leq x,x'}} \left| \frac{h(x,y) - h(x',y)}{x - x'} \right| \tag{C.8}$$

which is finite because

$$\frac{\partial h}{\partial x}(x,y) = ye^{-xy} + y\frac{x^2 + xy - 1}{(1+x^2)^2}$$

and so for $x, x' \geq 0$ and $0 \leq y \leq T$,

$$\left| \frac{\partial h}{\partial x}(x,y) \right| \leq T + T \sup_{v\geq 0} \left(\frac{v^2 + Tv + 1}{(1+v^2)^2} \right).$$

Let $\pi_{n,j}$ be the signed measure defined for $A \in \mathcal{B}$ by

$$\pi_{n,j}(A) = \nu_n(A \times (a_j, b_j]) - \nu(A \times (a_j, b_j]).$$

For $\widetilde{B}^{\nu,3}_{n,k}$ one needs to consider the measure $\pi_{n,k}$ defined similarly but with $A \times \{b'_k\}$ instead of $A \times (a_j, b_j]$. With this notation we have

$$\widetilde{B}^{\nu,5}_{n,j} \leq \sup_{0 \leq y \leq \overline{c}^u_{t,L}} \left| \int_{[d,\infty)} h(x,y) \pi_{n,j}(dx) \right|.$$

Fix $Y, \eta > 0$ and consider a subdivision $d = \tau_1 < \cdots < \tau_N < \tau_{N+1} = \infty$ with the following three properties: (1) $\tau_{\ell+1} - \tau_\ell \leq \eta$ for all $1 \leq \ell < N$; (2) $\tau_N = Y$; and (3) $\nu(\{\tau_\ell\} \times (a_j, b_j]) = 0$ for all $1 \leq \ell \leq N$. For $\widetilde{B}^{\nu,3}_{n,k}$ the third condition should be $\nu(\{\tau_\ell\} \times \{b'_\ell\}) = 0$ for all $1 \leq \ell \leq N$. Then for any $y \geq 0$,

$$\left| \int_{[d,\infty)} h(x,y) \pi_{n,j}(dx) \right| \leq \sum_{\ell=1}^{N-1} \int_{[\tau_\ell, \tau_{\ell+1})} |h(x,y) - h(\tau_\ell, y)| \|\pi_{n,j}\|(dx) + \int_{[Y,\infty)} |h(x,y) - h(Y,y)| \|\pi_{n,j}\|(dx) + \sum_{\ell=1}^{N} |h(\tau_\ell, y)| \left| \pi_{n,j}([\tau_\ell, \tau_{\ell+1})) \right|.$$

By choice of the partition (τ_ℓ) and by definition (C.8) of c_6, we have for any $y \leq \overline{c}^u_{t,L}$

$$\sum_{\ell=1}^{N-1} \int_{[\tau_\ell, \tau_{\ell+1})} |h(x,y) - h(\tau_\ell, y)| \|\pi_{n,j}\|(dx) \leq c_6(\overline{c}^u_{t,L}) \sum_{\ell=1}^{N-1} \int_{[\tau_\ell, \tau_{\ell+1})} |x - \tau_\ell| \|\pi_{n,j}\|(dx) \leq \eta c_6(\overline{c}^u_{t,L}) \|\pi_{n,j}\|([d,\infty)).$$

Thus introducing the constant

$$\overline{c}^h_{t,L} = \sup \left\{ |h(x,y)| : x \geq 0, 0 \leq y \leq \overline{c}^u_{t,L} \right\}$$

which in view of (B.2) can be seen to be finite, one gets for any $y \leq \overline{c}^u_{t,L}$,

$$\left| \int_{[d,\infty)} h(x,y) \pi_{n,j}(dx) \right| \leq \eta c_6(\overline{c}^u_{t,L}) \|\pi_{n,j}\|([d,\infty)) + 2\overline{c}^h_{t,L} \|\pi_{n,j}\|([Y,\infty)) + \overline{c}^h_{t,L} \sum_{\ell=1}^{N} \left| \pi_{n,j}([\tau_\ell, \tau_{\ell+1})) \right|.$$

Since no (τ_ℓ) is an atom of the measure $\int_{\cdot \times (a_j, b_j]} \nu(dx\, dy)$, it follows from (A1) that $\pi_{n,j}([\tau_\ell, \tau_{\ell+1})) \to 0$ as n goes to infinity for each ℓ. Moreover, one has

$$|\pi_{n,j}(A)| \leq \nu_n(A \times (a_j, b_j]) + \nu(A \times (a_j, b_j])$$

and finally, for any $\eta > 0$ we have, using also the fact that $\limsup_{n\to\infty} \|\pi_{n,j}\|([c,\infty)) \leq 2\nu([c,\infty) \times (a_j, b_j])$ for any $c \geq 0$,

$$\limsup_{n\to\infty} \sup_{0 \leq y \leq \overline{c}^u_{t,L}} \left| \int_{[d,\infty)} h(x,y) \pi_{n,j}(dx) \right| \leq 2\eta c_6(\overline{c}^u_{t,L}) \nu([d,\infty) \times (a_j, b_j]) + 4\overline{c}^h_{t,L} \nu([Y,\infty) \times (a_j, b_j]).$$

Thus letting $\eta \to 0$ and $Y \to \infty$ finally shows that $\widetilde{B}^{\nu,5}_{n,j} \to 0$ for each $1 \leq j \leq J$ and also $\widetilde{B}^{\nu,3}_{n,k} \to 0$ for each $1 \leq k \leq K$. Hence combining (C.5) and (C.7) finally gives

$$\limsup_{n\to\infty} \widetilde{B}^{\nu}_n \leq \varepsilon \left[c_5(\overline{c}^u_{t,L}) + 2C_{s,t,\lambda} \mu(s,t] \right]$$

and since ε is arbitrary, letting $\varepsilon \to 0$ achieves to prove that $R_n(s) \to 0$.

C.3 Boundedness of $(R_n(y))$

We now complete the proof of the lemma by showing that $\sup\{R_n(y) : 0 \le y \le t, n \ge 1\}$ is finite. We have $R_n(y) \le |\Psi_n(u_n)((y,t])| + |\Psi(u_n)((y,t])|$, so that it is enough to prove that

$$\sup\{|\Psi_n(u_n)((y,t])| : 0 \le y \le t, n \ge 1\} < \infty \tag{C.9}$$

and similarly with Ψ instead of Ψ_n. Using (3.7) for the first equality and (3.16) for the second inequality, we get for any $0 \le y \le t$

$$|\Psi_n(u_n)((y,t])| = |u_n(y) - u_n(t)| \le \Delta^u_{t,L}\, \mu_n(y,t] \le \Delta^u_{t,L} \sup_{n\ge1} \mu_n(t)$$

so that (C.9) holds. On the other hand, starting from the definition of Ψ we get

$$\begin{aligned}|\Psi(u_n)((y,t])| &\le \int_{(s,t]} |u_n| d\|\alpha\| + \int_{(s,t]} u_n^2 d\beta + \int_{(0,\infty)\times(s,t]} |h(x, u_n(y))| \nu(dx\, dy)\\ &\le \bar{c}^u_{t,L} \|\alpha\|(t) + (\bar{c}^u_{t,L})^2 \beta(t) + c_5(\bar{c}^u_{t,L}) \int_{(0,\infty)\times(0,t]} \frac{x^2}{1+x^2} \nu(dx\, dy)\end{aligned}$$

which ends the proof of the lemma, since this upper bound is finite (invoking (4.2) for the finiteness of the integral term).

References

[1] V. I. Afanasyev. Invariance principle for the critical branching process in a random environment attaining a high level. *Theory Probab. Appl.*, 54(1):1–13, 2010. MR-2766644

[2] V. I. Afanasyev, J. Geiger, G. Kersting, and V. A. Vatutin. Criticality for branching processes in random environment. *Ann. Probab.*, 33(2):645–673, 2005. MR-2123206

[3] Krishna B. Athreya and Samuel Karlin. On branching processes with random environments. I. Extinction probabilities. *Ann. Math. Statist.*, 42:1499–1520, 1971. MR-0298780

[4] Krishna B. Athreya and Samuel Karlin. Branching processes with random environments. II. Limit theorems. *Ann. Math. Statist.*, 42:1843–1858, 1971. MR-0298781

[5] V. Bansaye and J. Berestycki. Large deviations for branching processes in random environment. *Markov Process. Related Fields*, 15(4):493–524, 2009. MR-2598126

[6] Vincent Bansaye, Juan Carlos Pardo Millan, and Charline Smadi. On the extinction of Continuous State Branching Processes with catastrophes. arXiv:1301.6093.

[7] V. Bansaye and F. Simatos. A sufficient condition for the tightness of time-inhomogeneous Markov processes. arXiv:1409.5215.

[8] Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996. MR-1406564

[9] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. MR-1700749

[10] Matthias Birkner, Jochen Geiger, and Götz Kersting. Branching processes in random environment – a view on critical and subcritical cases. In *Interacting stochastic systems*, pages 269–291. Springer, Berlin, 2005. MR-2118578

[11] C. Böinghoff, E. E. Dyakonova, G. Kersting, and V. A. Vatutin. Branching processes in random environment which extinct at a given moment. *Markov Process. Related Fields*, 16(2):329–350, 2010. MR-2666857

[12] C. Böinghoff and M. Hutzenthaler. Branching diffusions in random environment. *Markov Process. Related Fields*, 18(2):269–310, 2012. MR-2985138

[13] K. Borovkov. A note on diffusion-type approximation to branching processes in random environments. *Teor. Veroyatnost. i Primenen.*, 47(1):183–188, 2002. MR-1978708

[14] Ma. Emilia Caballero, Amaury Lambert, and Gerónimo Uribe Bravo. Proof(s) of the Lamperti representation of continuous-state branching processes. *Probab. Surv.*, 6:62–89, 2009. MR-2592395

[15] E. B. Dynkin. Branching particle systems and superprocesses. *Ann. Probab.*, 19(3):1157–1194, 1991. MR-1112411

[16] E. B. Dynkin. Superprocesses and partial differential equations. *Ann. Probab.*, 21(3):1185–1262, 1993. MR-1235414

[17] Nicole El Karoui and Sylvie Roelly. Propriétés de martingales, explosion et représentation de Lévy-Khintchine d'une classe de processus de branchement à valeurs mesures. *Stochastic Process. Appl.*, 38(2):239–266, 1991. MR-1119983

[18] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes. Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. MR-0838085

[19] William Feller. Diffusion processes in genetics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pages 227–246, Berkeley and Los Angeles, 1951. University of California Press. MR-0046022

[20] H. Föllmer. Calcul d'Itô sans probabilités. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, volume 850 of *Lecture Notes in Math.*, pages 143–150. Springer, Berlin, 1981.

[21] Jochen Geiger, Götz Kersting, and Vladimir A. Vatutin. Limit theorems for subcritical branching processes in random environment. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(4):593–620, 2003. MR-1983172

[22] B. V. Gnedenko and A. N. Kolmogorov. *Limit distributions for sums of independent random variables.* Translated from the Russian, annotated, and revised by K. L. Chung. With appendices by J. L. Doob and P. L. Hsu. Revised edition. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills., Ont., 1968. MR-0233400

[23] D. R. Grey. Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probability*, 11:669–677, 1974. MR-0408016

[24] Anders Grimvall. On the convergence of sequences of branching processes. *Ann. Probab.*, 2:1027–1045, 1974. MR-0362529

[25] Yves Guivarc'h and Quansheng Liu. Propriétés asymptotiques des processus de branchement en environnement aléatoire. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(4):339–344, 2001. MR-1821473

[26] Inge S. Helland. Continuity of a class of random time transformations. *Stochastic Process. Appl.*, 7(1):79–99, 1978. MR-0488203

[27] Inge S. Helland. Minimal conditions for weak convergence to a diffusion process on the line. *Ann. Probab.*, 9(3):429–452, 1981. MR-0614628

[28] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003. MR-1943877

[29] Miloslav Jiřina. Stochastic branching processes with continuous state space. *Czechoslovak Mathematical Journal*, 8(83):292–321, 1958. MR-0101554

[30] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002. MR-1876169

[31] Niels Keiding. Extinction and exponential growth in random environments. *Theoretical Population Biology*, 8(1):49–63, 1975.

[32] David G. Kendall. On the generalized "birth-and-death" process. *Ann. Math. Statist.*, 19:1–15, 1948. MR-0024091

[33] A. N. Kolmogorov. The transition of branching processes into diffusion processes and associated problem of genetics. *Theory Probab. Appl.*, 4, 218-220, 1959.

[34] Thomas G. Kurtz. Semigroups of conditioned shifts and approximation of Markov processes. *Ann. Probab.*, 3(4):618–642, 1975. MR-0383544

[35] Thomas G. Kurtz. Diffusion approximations for branching processes. In *Branching processes (Conf., Saint Hippolyte, Que., 1976)*, volume 5 of *Adv. Probab. Related Topics*, pages 269–292. Dekker, New York, 1978. MR-0517538

[36] John Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73:382–386, 1967. MR-0208685

[37] John Lamperti. The limit of a sequence of branching processes. *Z. Wahrsch. Verw. Gebiete*, 7:271–288, 1967. MR-0217893

[38] John Lamperti. Limiting distributions for branching processes. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 2, pages 225–241. University of California Press, 1967. MR-0219148

[39] Torgny Lindvall. Convergence of critical Galton–Watson branching processes. *J. Appl. Probability*, 9:445-450, 1972. MR-0345227

[40] Sylvie Roelly-Coppoletta. A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, 17(1–2):43–65, 1986. MR-0878553

[41] M. L. Silverstein. A new approach to local times. *J. Math. Mech.*, 17:1023–1054, 1967/1968. MR-0226734

[42] Walter L. Smith and William E. Wilkinson. On branching processes in random environments. *Ann. Math. Statist.*, 40:814–827, 1969. MR-0246380

[43] Ward Whitt. Some useful functions for functional limit theorems. *Math. Oper. Res.*, 5(1):67–85, 1980. MR-0561155

[44] Eugene Wong and Moshe Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.*, 36:1560–1564, 1965. MR-0195142

Acknowledgments. We thank Julien Berestycki for interesting discussions at the earliest stage of this work. We are also very grateful to the anonymous referees for their careful reading of this paper, which has led to a deep improvement of the presentation.

This work was funded by project MANEGE "Modèles Aléatoires en Écologie, Génétique et Évolution" 09-BLAN-0215 of ANR (French national research agency), Chair Modélisation Mathématique et Biodiversité VEOLIA-Ecole Polytechnique-MNHN-F.X. and the professoral chair Jean Marjoulet. While most of this work was carried out, the second author was affiliated with CWI and his research was funded by an NWO-VIDI grant.