CrossMark

J. Appl. Prob. **58,** 637–676 (2021)
*doi:*10.1017/jpr.2020.112

EXTINCTION TIME OF THE LOGISTIC PROCESS

ERIC FOXALL,* *University of British Columbia, Okanagan Campus*

Abstract

The logistic birth and death process is perhaps the simplest stochastic population model that has both density-dependent reproduction and a phase transition, and a lot can be learned about the process by studying its extinction time, τ_n, as a function of system size n. A number of existing results describe the scaling of τ_n as $n \to \infty$ for various choices of reproductive rate r_n and initial population $X_n(0)$ as a function of n. We collect and complete this picture, obtaining a complete classification of all sequences (r_n) and $(X_n(0))$ for which there exist rescaling parameters (s_n) and (t_n) such that $(\tau_n - t_n)/s_n$ converges in distribution as $n \to \infty$, and identifying the limits in each case.

Keywords: SIS model; stochastic logistic model; diffusion approximation; large deviations; phase transition

2010 Mathematics Subject Classification: Primary 60G99
Secondary 60K35

1. Introduction

Fix $n \in \mathbb{N}$ and $r \in \mathbb{R}_+$ and consider the continuous-time Markov chain with

$$X \to \begin{cases} X+1 & \text{at rate} \quad rX(1-X/n), \\ X-1 & \text{at rate} \quad X. \end{cases} \tag{1}$$

This birth and death process has a simple interpretation in terms of infection spread: in a population of n individuals, X of them are type I (infectious), and the remaining $n-X$ individuals are type S (susceptible). Each type I individual, at rate r, selects an individual uniformly at random from the population and infects them. In addition, each type I individual recovers at rate 1 and is once again immediately susceptible. Since the state transitions at the individual level are $S \to I \to S$, this is called the SIS model.

This model first appears, in a somewhat more general form, in a probabilitistic treatment of logistic population growth [10, Section 4] (see [11, pp. 471–495] for an English translation), where it is shown that the expected value of the process at each point in time is bounded above by the solution to the corresponding logistic differential equation, namely $X' = rX(1-X/n) - X$. Later, [17] gives limit theorems that apply to the sequence of processes $x_n = X_n/n$ (using the subscript to emphasize dependence on n) when r is fixed and $\lim_{n\to\infty} x_n(0)$ exists, demonstrating convergence, on compact time intervals, to the solution of the logistic ordinary differential equation (ODE) $x' = rx(1-x) - x$ with $x(0) = \lim_{n\to\infty} x_n(0)$, with Gaussian fluctuations of order $\frac{1}{\sqrt{n}}$ that solve an explicit stochastic differential equation

Received 14 November 2018; revision received 29 November 2020.
* Postal address: 3333 University Way, Kelowna BC Canada V1V 1V7. Email: efoxall@mail.ubc.ca

© The Author(s), 2021. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

(SDE). Extensions of the ODE convergence to longer (growing slowly with n) time intervals are possible under certain conditions; see, for example, [3].

Later work has focused on understanding the long-term behaviour of the process for large n, in particular (i) the time to extinction $\tau_n := \inf\{t\colon X_n(t) = 0\}$ and (ii) the behaviour prior to extinction. In this article we shall focus on the distribution of τ_n, as $n \to \infty$, as a function of the reproductive parameter r_n and the initial value $X_n(0)$ when both are allowed to vary with n. We do so by identifying *distributional limits* of (τ_n), i.e. deterministic sequences (s_n), (t_n) and random variables T such that $(\tau_n - t_n)/s_n$ converges in distribution to T as $n \to \infty$. Much of the work on this problem already exists; here, we provide a survey and concise proofs of known results, then fill in the remaining gaps to produce a complete theory on the topic.

There is a parallel line of inquiry on model (1) that concerns the so-called quasi-stationary distribution (QSD), which is the unique stationary distribution conditioned on non-extinction; see the comparatively recent monograph of Nasell [20] for a comprehensive survey on this topic, which also includes a discussion of the three phases of the model that we explain below and that are key to understanding the behaviour of the model. As explained in [20], if $X_n(0)$ has the QSD then τ_n is exponentially distributed for each n, and by studying the QSD we can obtain estimates of $\mathbb{E}[\tau_n]$ when $X_n(0)$ has either the QSD or the deterministic value 1. Generally speaking, $\mathbb{E}[\tau_n]$ corresponds to (t_n) in the rescaling described above; then, further methods are needed to find (s_n) and T. These methods appear progressively over time in works that are not surveyed in [20]. So, by gathering these methods and previous works and completing the picture, we obtain a complementary perspective to the path taken in [20].

As is easily shown, the ODE $x' = rx(1-x) - x$ undergoes a transcritical bifurcation at $(x, r) = (0, 1)$, and for $r > 1$ has the endemic equilibrium $x_\star = 1 - 1/r$. An important question is whether this is mirrored by the behaviour of the stochastic process. Kryscio and LefÈvre [16] estimate $\mathbb{E}[\tau_n]$, as well as the QSD for large n; in particular, they find that for $r < 1$, $\mathbb{E}[\tau_n] = O(\log n)$ uniformly over $X_n(0)$ and that the QSD is concentrated near 0, while for $r > 1$, $\mathbb{E}[\tau_n]$ grows exponentially with n uniformly over $X_n(0)$ and the QSD of x_n is concentrated near $x_\star$. Later, [1] refines results on extinction time to limit theorems for the distribution of τ_n when $r < 1$ and $r > 1$ in the following cases: (i) $X_n(0)$ is a constant independent of n, and (ii) $\lim_{n\to\infty} x_n(0)$ exists and is positive. When $r < 1$ and $X_n(0)$ is constant, for large n, $X_n(t)$ is well approximated by a branching process up until extinction, so τ_n converges in distribution to the extinction time of the branching process. When $r < 1$ and $\lim_{n\to\infty} x_n(0) > 0$, $(1-r)\tau_n - t_n$ converges in distribution to a standard Gumbel G, with $\mathbb{P}(G \le w) = \mathrm{e}^{-\mathrm{e}^{-w}}$ for some deterministic correction t_n of order roughly $\log n$. As pointed out in [6], the formula given in [1] has an error in this case; the corrected formula in the case $r < 1$ can be found, for example, in [5]. When $r > 1$ and $\lim_{n\to\infty} x_n(0) > 0$, $\tau_n/\mathbb{E}[\tau_n]$ converges in distribution to exponential with rate 1. The basic reasoning is that $x_n(t)$ is metastable around $x_\star$ with Ornstein–Uhlenbeck-type fluctuations of order $\frac{1}{\sqrt{n}}$, which can be deduced from the results of [17], so a rare and rather sudden event is needed to cause extinction.

The behaviour near $r = 1$ is more delicate. Nasell [18] identifies the transition region $r_n - 1 = O\big(\frac{1}{\sqrt{n}}\big)$, also known as the transition window or critical regime, and finds that for $r_n = 1 + \frac{c}{\sqrt{n}}$, $\mathbb{E}[\tau_n] \sim f(c)\sqrt{n}$ for some function f when the distribution of $X_n(0)$ is the QSD. For the same scale of r_n, [7] proves convergence in distribution of $Y_n(t) := X_n(\sqrt{n}t)/\sqrt{n}$ to the solution of the modified Feller diffusion appearing below in Theorem 2, when $\lim_{n\to\infty} Y_n(0)$ exists. Brightwell, House, and Luczak [5] obtain the distribution of τ_n throughout the subcritical

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

regime, i.e. when $\lim_{n\to\infty} r_n \le 1$ and $\sqrt{n}(1-r_n)\to\infty$, for initial values satisfying $X_n(0)(1-r_n)\to\infty$, with particular focus on the barely subcritical case $1-r_n = o(1)$. Until now, the barely supercritical case, where $r_n - 1 = o(1)$ and $\sqrt{n}\,(r_n-1)\to\infty$, has remained unsolved.

In this article we collect and complete existing results concerning the distribution of τ_n for large n in order to obtain a complete understanding of the extinction time for all possible choices of r_n and $X_n(0)$. Since the existing results form a somewhat overlapping patchwork of the different cases, I have opted to include a full and self-contained proof of all the different cases, which allows us to more accurately view the true extent of the different methods used. Generally speaking, we shall follow existing methods, although in some cases some improvement was possible; I shall point out when each is the case. In addition, the barely supercritical case is new, and in fact it requires the most effort.

The breadth of our results can be summarized as follows. Let $\delta_n = r_n - 1$, $a_n = \big(|\delta_n| + \frac{1}{\sqrt{n}}\big) X_n(0)$, and $c_n = \sqrt{n}\,\delta_n$. Suppose that $X_n(0)\to X_\infty(0)\in\mathbb{N}\cup\{\infty\}$, $r_n\to r_\infty\in[0,\infty)$, $a_n\to a_\infty\in[0,\infty]$, and $c_n\to c_\infty\in[-\infty,\infty]$ as $n\to\infty$. Then there exist sequences (s_n), (t_n) and a non-degenerate distribution function $F{:}[0,\infty)\to[0,1]$ (that we identify in every case) such that $\mathbb{P}((\tau_n - t_n)/s_n \le w)\to F(w)$ as $n\to\infty$ at continuity points w of F. Moreover, F depends only on the values of the limits $X_\infty(0)$, r_∞, a_∞, and c_∞. As explained below, a_n measures whether initial values are small or large, and c_n determines the phase of the process.

2. Phase diagram and main results

Let us formally state the setting and the assumptions that hold throughout the paper. Given $n\in\mathbb{N}$, $r_n\in[0,\infty)$, and $X_n(0)\in\{0,\dots,n\}$, X_n denotes the Markov chain $(X_n(t))_{t\ge0}$ with initial value $X_n(0)$ and transition rates

$$X_n \to \begin{cases} X_n + 1 & \text{at rate} \quad r_n X_n(1 - X_n/n), \\ X_n - 1 & \text{at rate} \quad X_n. \end{cases} \tag{2}$$

The extinction time of X_n is $\tau_n = \inf\{t\colon X_n(t) = 0\}$. We fix a sequence of values $(X_n(0))_{n\ge1}$ and $(r_n)_{n\ge1}$ and seek deterministic sequences (t_n) and (s_n) and a random variable T such that $(\tau_n - t_n)/s_n$ converges in distribution to T as $n\to\infty$; we refer to this as a *distributional limit* for (τ_n), subject to the following assumptions.

Assumption 1. *Let* $\delta_n = r_n - 1$, $a_n = \big(|\delta_n| + \frac{1}{\sqrt{n}}\big) X_n(0)$, *and* $c_n = \sqrt{n}\,\delta_n$. *As* $n\to\infty$, *it is assumed that* $X_n(0)\to X_\infty(0)\in\mathbb{N}\cup\{\infty\}$, $r_n\to r_\infty\in[0,\infty)$, $a_n\to a_\infty\in[0,\infty]$, *and* $c_n\to c_\infty\in[-\infty,\infty]$.

The behaviour of τ_n depends on r_n and $X_n(0)$, or equivalently r_n and $x_n(0) = X_n(0)/n$. So, we can organize our results into a phase diagram for $(r,x)\in\mathbb{R}_+\times[0,1]$. As observed in [18], the parameter values are effectively partitioned into three phases:

1. Subcritical: $c_n\to-\infty$ as $n\to\infty$
2. Critical: $c_n\to c_\infty\in\mathbb{R}$ as $n\to\infty$
3. Supercritical: $c_n\to\infty$ as $n\to\infty$.

It is important to note that since $c_n = \sqrt{n}\,(r_n - 1)$, the critical phase has width $O\big(\frac{1}{\sqrt{n}}\big) = o(1)$ around the point $r = 1$. Along with parameter values, there is a dependence on the initial value

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

$X_n(0)$. The following more refined partition into six regions delineates the different regimes for the limit behaviour of τ_n.

1. Discrete: $X_\infty(0) \in \mathbb{N}$ and $r_\infty \leq 1$
2. Linear diffusive: $X_\infty(0) = \infty$ and either
   (a) subcritical phase and $a_\infty < \infty$, or
   (b) critical/supercritical phase and $a_\infty = 0$
3. Subthreshold cutoff: subcritical phase and $a_\infty = \infty$
4. Non-linear diffusive: critical phase and $a_\infty > 0$
5. Threshold: supercritical phase and $0 < a_\infty < \infty$
6. Metastable: supercritical phase and $a_\infty = \infty$.

The following methods are used in each region to establish the distributional limit for (τ_n).

1. Discrete: As observed in [1], in this region, for each n, X_n can be coupled to the linear birth and death process Z_n with $Z_n(0) = X_n(0)$ and transition rates
$$Z_n \to \begin{cases} Z_n + 1 & \text{at rate} \quad r_n Z_n, \\ Z_n - 1 & \text{at rate} \quad Z_n \end{cases} \tag{3}$$
in such a way that $\mathbb{P}(X_n(t) = Z_n(t) \text{ for all } t \in [0, \tau_n]) = 1 - o(1)$.
2. Linear diffusive: As observed in [7], in this region $X_n(X_n(0)\, t)/X_n(0)$ converges in distribution to the diffusion Y with $Y(0) = 1$ and
$$\mathrm{d}Y = -a_\infty Y \mathrm{d}t + \sqrt{2Y}\mathrm{d}B, \tag{4}$$
where B is standard Brownian motion, and it is not hard to strengthen this result to show that $\tau_n/X_n(0)$ converges in distribution to $\inf\{t\colon Y(t) = 0\}$. This explains nicely the form of the limit. The method we will use instead, since it applies across several regions, is that of [5], which is essentially the same coupling method as [1] but extends the approach by not requiring that $X_n(t) = Z_n(t)$ for all $t \leq \tau_n$ with high probability, but instead using both upper and lower bounds (Z_n) and (Z_n'), both linear birth and death processes, such that the extinction time of both is comparable to (τ_n). To define the lower bound (Z_n'), a ceiling M_n is defined such that the coupling is valid so long as $X_n \leq M_n$ and such that X_n hits 0 before M_n with high probability.
3. Subthreshold cutoff: As observed in [5], x_n is well approximated by the logistic ODE $x' = r_\infty\, x\,(1 - x)$ until $|\delta_n|\, X_n = O(1)$, at which point X_n enters the linear diffusive regime and the above coupling method can be used. In practice, [5] shows the coupling method can be applied once $|\delta_n|\, X_n \leq \omega_n$ for some (ω_n) tending to ∞ slowly enough as $n \to \infty$ as a function of (c_n).
4. Non-linear diffusive: As observed in [7], the process $X_n(\sqrt{n}\, t)/\sqrt{n}$ converges in distribution to the diffusion Y with $Y(0) = \lim_{n\to\infty} X_n(0)/\sqrt{n}$ and $\mathrm{d}Y = (c_\infty Y - Y^2)\,\mathrm{d}t + \sqrt{2Y}\mathrm{d}B$. In this case we make use of this diffusion limit combined with a result that says $\limsup_n \mathbb{P}\big(\frac{\tau_n}{\sqrt{n}} > t \mid X_n(0) \leq \alpha\sqrt{n}\big) \to 0$ for each $t > 0$ as $\alpha > 0$ to establish convergence in distribution of $\frac{\tau_n}{\sqrt{n}}$ to $\inf\{t\colon Y(t) = 0\}$.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

5. Threshold: Let $x_n^\star = 1 - 1/r_n$, $X_n^\star = \lfloor nx_n^\star \rfloor$, and $\tau_n^\star = \inf\{t\colon X_n(t) \in \{0, X_n^\star\}\}$. Then, as observed in [1] when $r_n = r_\infty > 1$ and $X_n(0) = X_\infty(0)$ for all n, and extended in this article to the whole supercritical phase,

$$\lim_{n\to\infty} \mathbb{P}(X_n(\tau_n^\star) = 0) = \begin{cases} r_\infty^{-X_\infty(0)} & \text{if } X_\infty(0) < \infty, \\ \mathrm{e}^{-a_\infty} & \text{if } X_\infty(0) = \infty. \end{cases}$$

   In particular, the limiting probability is in (0,1). Conditioned on $\{X_n(\tau_n^\star) = 0\}$, the process behaves effectively as though it is in the linear diffusive regime with parameter $1 - 1/r_\infty$, and the same approximation applies. On the event $\{X_n(\tau_n^\star) = X_n^\star\}$, the process enters the metastable regime. This dichotomy appears to have been first observed in [1].

6. Metastable: In this case, the process X_n reaches $X_n^\star$ relatively quickly, with typical fluctuations described as follows: the process $(X_n(t/\delta_n) - X_n^\star)/\sqrt{n}$ converges in distribution to the Ornstein–Uhlenbeck process Y described by $\mathrm{d}Y = -Y\mathrm{d}t + \sqrt{2/r_\star}\mathrm{d}B$. General results of this type can be found in [17]. As observed in [1], when $r_n = r_\infty > 1$ for each n, on each excursion from $X_n^\star$ the probability of hitting 0 can be computed quite precisely using well-known formulae for birth and death processes. In this article we extend these calculations to the whole supercritical phase, using essentially the same method but exercising greater care in the barely supercritical ($c_\infty = \infty$ and $r_\infty = 1$) case.

In stating the results for the subcritical phase it is helpful to define $\gamma_n = 1 - r_n = -\delta_n$. We now give precise statements of the results, taking care to cite previous work in the relevant cases.

Theorem 1. *(Discrete case, Theorem 1 (B2) in [1]). Suppose $r_\infty \le 1$ and $X_\infty(0) \in \mathbb{N}$. For each $t > 0$, as $n \to \infty$,*

1. *if $r_\infty = 1$ then $\mathbb{P}(\tau_n \le t) \to (1 + 1/t)^{-X_\infty(0)}$;*
2. *if $r_\infty < 1$ then $\mathbb{P}(\tau_n \le t) \to (1 + \gamma_\infty/(\mathrm{e}^{\gamma_\infty t} - 1))^{-X_\infty(0)}$.*

Theorem 2 is proved by first studying the extinction time of the linear birth and death process Z_n from (3), then transferring the result to X_n using a coupling with the property that $\mathbb{P}(X_n(t) = Z_n(t) \text{ for all } t \in [0, \tau_n]) = 1 - o(1)$.

The next result appears not to have been proved yet in any of the listed references, though it follows straightforwardly from the coupling method already discussed from [5].

Theorem 2. *(Linear diffusive.) Suppose $r_\infty \le 1$, $X_\infty(0) = \infty$, and either $c_\infty = -\infty$ and $a_\infty < \infty$, or $c_\infty > -\infty$ and $a_\infty = 0$. Then $\tau_n/X_n(0) \xrightarrow{(\mathrm{d})} H_{a_\infty}$ with*

$$\mathbb{P}(H_{a_\infty} \le w) = \begin{cases} \mathrm{e}^{-1/w} & \text{if} \quad a_\infty = 0, \\ \mathrm{e}^{-a_\infty/(\mathrm{e}^{a_\infty w} - 1)} & \text{if} \quad a_\infty > 0. \end{cases}$$

Theorem 2 is proved in a similar way to Theorem 1. In this case, we cannot enforce $X_n(t) = Z_n(t)$ for all $t \in [0, \tau_n]$ with high probability, so instead we bound X_n above and below by linear birth and death processes with different parameters. This 'sandwiching' idea appears in [5] to prove the final stage of Theorem 3 (see below). A particular case of Theorem 2, where $r = 1 + Cn^{-\alpha}$ for $\alpha < 1/2$, is discussed in [7]; they prove convergence to the diffusion limit in

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

(4), but not the convergence of τ_n. Although we do not do it here, we could also prove Theorem 2 via the diffusion limit (4) by the same methods we use to prove Theorem 4.

The next result is the main result of [5].

Theorem 3. *(Subthreshold cutoff, Theorem 1.1 in [5].) Suppose that $c_\infty = -\infty$ and $a_\infty = \infty$. Let G denote the standard Gumbel: $\mathbb{P}(G \le w) = \exp(-e^{-w})$. Then $\gamma_n \tau_n - g_n(X_n(0)) \overset{(d)}{\to} G$ as $n \to \infty$, where $g_n(X) = \log(\gamma_n^2 n) - \log(r_n + \gamma_n n/X)$. In particular,*

$$g_n(X_n(0)) = \begin{cases} \log(\gamma_n X_n(0)) + o(1) & \text{if} \quad X_n(0) = o(\gamma_n n), \\ \log(\gamma_n^2 n) - \log(r_\infty + 1/b_\star) & \text{if} \quad X_n(0)/\gamma_n n \to b_\star \in (0, \infty]. \end{cases}$$

In [5] the parameter r is denoted λ, and the rate of $X \to X - 1$ is equal to μX instead of X; since μ can be set to 1 by a uniform time change, no generality is lost. In [5], Theorem 3 is proved in three stages, which we describe in the present notation.

Initial stage: A differential inequality for x_n (incidentally, the same one originally proved in [10]) is used to show that from any initial value, x_n drops to $\gamma_n |c_n|^\epsilon$ for some $\epsilon > 0$ within $o(1/\gamma_n)$ amount of time (see [5, Lemma 4.2] and note that since we take $\mu = 1$, in our notation $\mu - \lambda$ is just γ_n, and since $\mu - \lambda \to 0$, $\lambda \to 1$ as $n \to \infty$).

Intermediate stage: From any initial value $x_n(0) \le \gamma_n |c_n|^\epsilon$ until the first time that $x_n \le \gamma_n |c_n|^{-\epsilon}$, x_n remains close to the solution of the corresponding logistic equation $x' = rx(1-x) - x$.

Final stage: From any initial value $x_n(0) \le \gamma_n |c_n|^{-\epsilon}$, a coupling argument is used to bound X_n both above and below by linear birth and death processes with different parameters, as discussed earlier.

Our proof of Theorem 3 in the initial and final stages is basically identical to [5]; in the intermediate stage we take a somewhat different approach as discussed in Section 6.

Theorem 4. *(Non-linear diffusive.) Suppose $c_\infty \in \mathbb{R}$ and $a_\infty > 0$, and let Y_n denote the rescaled process $Y_n(t) = X_n(\sqrt{n}\, t)/\sqrt{n}$. Let $y = \lim_{n\to\infty} Y_n(0) \in (0, \infty]$. Then, Y_n converges in distribution to the diffusion Y that solves the SDE*

$$dY = Y(c_\infty - Y)\, dt + \sqrt{2Y} dB, \qquad Y(0) = y,$$

and $\tau_n/\sqrt{n} \overset{(d)}{\to} T$, where $T = \inf\{t : Y(t) = 0\}$.

The convergence of $X_n(\sqrt{n}\, t)/\sqrt{n}$ to the diffusion limit Y is proved in [7], although they do not prove convergence of τ_n. To obtain the latter we employ the continuous mapping theorem to obtain convergence of the hitting time of ϵ for small $\epsilon > 0$, then show that once $Y_n \le \epsilon$, Y_n is likely to hit 0 in a short time.

The following result is proved in [1] in the particular case where $r_n = r_\infty > 1$ for all n, although the rapid extinction result is expressed without conditioning on hitting 0 before $X_n^\star$, which is just a difference in the presentation. Our proof takes the same basic approach as [1], with the caveats that the calculations become more delicate when $r_\infty = 1$, requiring a finer analysis, and that we have found a couple of gaps in their proof of the exponential limit (see

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

the discussion below the statement of Theorem 5), so we took the steps described below to bridge those gaps.

Theorem 5. *(Threshold and metastable.) Suppose $c_\infty = \infty$ and $a_\infty > 0$. Let $x_n^\star = 1 - 1/r_n$, $X_n^\star = \lfloor n x_n^\star \rfloor$, $\tau_n^\star = \inf\{t\colon X_n(t) \in \{0, X_n^\star\}\}$, $A_n^\star = \{X_n(\tau_n^\star) = X_n^\star\}$, and $B_n^\star = (A_n^\star)^{\mathrm{c}} = \{X_n(\tau_n^\star) = 0\}$.*

1. *Probability of rapid extinction:*

$$\lim_{n\to\infty} \mathbb{P}(B_n^\star) = \begin{cases} r_\infty^{-X_\infty(0)} & \text{if } X_\infty(0) < \infty, \\ \mathrm{e}^{-a_\infty} & \text{if } X_\infty(0) = \infty. \end{cases}$$

   In particular, $\mathbb{P}(A_n^\star) \to 1$ if $a_\infty = \infty$.

2. *Scaling of τ_n on rapid extinction: If $a_\infty < \infty$ and*

   (a) *$r_\infty = 1$ then, for each $t \geq 0$, $\mathbb{P}(\tau_n/X_n(0) \leq t \mid B_n^*) \to \mathbb{P}(H_{a_\infty} \leq t)$ as $n \to \infty$, as in the linear diffusive regime, and*

   (b) *if $r_\infty > 1$ then, conditioned on $(A_n^*)^{\mathrm{c}}$, τ_n/r_n has the same limit as τ_n in Case 2 of Theorem 1, except with δ_∞/r_∞ in place of γ_∞.*

3. *Scaling of τ_n at metastability:*

   (a) *Expected time to extinction:*

$$\mathbb{E}[\,\tau_n \mid A_n^\star\,] \sim \sqrt{\frac{2\pi}{n}}\,\frac{r_n}{\delta_n^2}\,\mathrm{e}^{n(\log r_n + 1/r_n - 1)}.$$

   (b) *Exponential limit: For each $t \geq 0$, as $n \to \infty$, $\mathbb{P}(\tau_n/\mathbb{E}[\tau_n \mid A_n^\star] \leq t \mid A_n^\star) \to 1 - \mathrm{e}^{-t}$.*

Having access to the prefactors (the stuff in front of the exponential function) in the expected time to extinction allows us to see how it blends into the non-linear diffusive limit as c_n approaches $O(1)$. For r_n near 1, a Taylor expansion gives $\log r_n + 1/r_n - 1 \sim \delta_n^2/2$, so the exponential term is $\exp((1 + o(1))n\delta_n^2/2) = \exp((1 + o(1))c_n^2/2)$ since $c_n = \sqrt{n}\delta_n$, and the prefactor is $\sqrt{2\pi}\, r_n/(\sqrt{n}\delta_n^2) = \sqrt{2\pi n}\, r_n/c_n^2$. Thus, when $r_\infty = 1$,

$$\mathbb{E}[\tau_n \mid A_n^\star] \sim \frac{\sqrt{2\pi n}}{c_n^2}\mathrm{e}^{(1+o(1))c_n^2/2},$$

which is of order $\sqrt{n}$ when $c_n = O(1)$.

Much of the proof of Theorem 5 revolves around precise estimation of various sums that all seem to involve the function $\nu(j, k) := \prod_{i=j+1}^{k} q_-(i)/q_+(i)$, where $q_-(i) = i$, $q_+(i) = rX(1 - X/n)$ is the rate at which X_t decreases, respectively increases, by 1 when $X_t = i$. We discuss the basic approach for each part:

Probability of rapid extinction: An explicit formula for $h_+(j) := \mathbb{P}(A_n^\star \mid X_0 = j)$ is available in terms of the transition rates of the process, and can be estimated. Here is where we first encounter $\nu(j, k)$.

Scaling on τ_n on rapid extinction: Using the Doob h-transform applied to the function $h_- := 1 - h_+$, we study X_n conditioned on $B_n^\star$. The transition rates can be written using h_-. By estimating transition rates we show the conditioned process corresponds to the setting of Theorem 1 or Theorem 2.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Scaling of τ_n at metastability: Conditioning on $A_n^\star$, we break up τ into three epochs: the time to hit $X_n^\star$ (approach), the time spent on excursions that return to $X_n^\star$ (sojourn), and the last excursion from $X_n^\star$ to 0 (fall). We show that the approach and fall time are small compared to the sojourn time, in expectation; so far this is the same approach as in [1]. We then use a coupling argument that shows the process is 'forgetful' in order to derive the exponential limit for the sojourn time; this method is a departure from [1].

We have also found the following gaps in the proofs in [1]:

Exponential limit: The exponential limit for $\tau_n/\mathbb{E}[\tau_n]$ is obtained in [1] via the same approach taken here: showing that the approach and fall time are small and that the sojourn time converges to exponential. Using their notation, for each n the sojourn time can be written as the sum $\sum_{m=1}^{K_n} \sigma_n(m)$, where $(\sigma_n(m))_{m=1}^\infty$ are independent and identically distributed (i.i.d.) and K_n is geometric independent of $(\sigma_n(m))$, with parameter $1/\mathbb{E}[K_n]$ that $\to 0$ as $n \to \infty$. In claiming that the normalized sojourn time has an exponential limit, [1] appeals to [15, Theorem 8.1A]. However, the theory discussed there, including that result, pertains only to a single i.i.d. sequence (T_m) of random variables. On the other hand, in this case the fact that the process itself depends on n implies that a doubly indexed sequence must be considered. I did some calculations (not included) that suggest that uniform integrability of $\sigma_1(n)$ with respect to n is sufficient in order to adapt the result from [15]. However, in [1] this is not done.

Expected excursion length: When expressing the time to extinction as the sum of the approach, sojourn, and fall times, it must be assumed that during the sojourn the process necessarily returns to $X_n^\star$ before hitting zero, which requires conditioning on $A_n^\star$. However, the formula used to compute the expected duration of each excursion from $X_n^\star$, namely $\mathbb{E}[\sigma_n(1)]$, in the proof of [1, Lemma 2], does not take account at all of conditioning on returning to $X_\star$. The answer comes out correct, since the conditioning does not have much effect, but it cannot be ignored outright.

The rest of the paper is organized as follows. In Section 3 we discuss some basic coupling results that are used a few times throughout the paper. Section 4 treats the linear birth and death process, essentially proving Theorems 1 and 2 for the linearization of X_n. In Section 6 we prove Theorems 1 and 2, and the 'final stage' of Theorem 3. In Section 7 we prove the rest of Theorem 3. In Section 6 we prove Theorem 4. Finally, in Section 8 we prove Theorem 5. In most of our proofs we suppress the dependence on n, to avoid writing subscripts everywhere; I will point this out as we go along.

3. Coupling of birth and death processes

Recall that a birth and death (b–d) process is a right-continuous, continuous-time Markov chain with state space $\mathbb{N} = \{0, 1, 2, \dots\}$ that jumps by ± 1 at each transition. There is a natural way to construct such a process, or even multiple such processes with different transition rates, from all initial conditions on a single probability space. We will focus on b–d processes with state space $\{0, \dots, N\}$ for some N, since this is all that is needed for this paper. We begin with the case of a single process.

3.1. Natural coupling for a single process

A b–d process is defined by its transition rates. For $x \in \{0, \dots, N\}$ let $b(x)$ and $d(x)$ be the transition rates from x to $x+1$ and $x-1$ respectively, assume that $d(0) = 0$ and $b(N) = 0$, and

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

define independent Poisson point processes $B(x)$ and $D(x)$ on $\mathbb{R}_+$ with respective rates $b(x)$ and $d(x)$. As a function of the collection $(B(x), D(x)\colon x \in \{0, \dots, N\})$ we shall define, for each $x \in \{0, \dots, N\}$, a process $(\Phi(x, t)\colon t \in [0, \infty))$ which is a copy of the b–d process with the given transition rates, and with initial value x at time 0. Let $\Phi(x, 0) = x$, and define $(t_i(x))_{i \le I(x)}$ and $(\Phi(x, t_i(x)))_{i \le I(x)}$ recursively by

$$t_0(x) = 0, \quad t_{i+1}(x) = \inf\{t > t_i(x)\colon t \in B(\Phi(x, t_i(x))) \cup D(\Phi(x, t_i(x)))\},$$

$$\Phi(x, t_{i+1}(x)) = \begin{cases} \Phi(x, t_i(x)) + 1 & \text{if } t_{i+1}(x) \in B(\Phi(x, t_i(x))), \\ \Phi(x, t_i(x)) - 1 & \text{if } t_{i+1}(x) \in D(\Phi(x, t_i(x))), \end{cases}$$

$$I(x) = \inf\{i\colon b(\Phi(x, t_i(x))) + d(\Phi(x, t_i(x))) = 0\}.$$

Then, for $i < I(x)$ let $\Phi(x, t) = \Phi(x, t_i(x))$ for $t \in [t_i(x), t_{i+1}(x))$, and if $I(x) < \infty$ let $\Phi(x, t) = \Phi(x, t_{I(x)}(x))$ for $t \in [t_{I(x)}(x), \infty)$. This defines the process on $[0, \zeta(x))$ where

$$\zeta(x) := \begin{cases} \lim_{i \to \infty} t_i(x)) & \text{if } I(x) = \infty, \\ \infty & \text{if } I(x) < \infty. \end{cases}$$

To verify that $\zeta(x) = \infty$ when $I(x) = \infty$, note that $\{t_i(x)\colon i \ge 1\} \subset U := \bigcup_{y \le N} B(y) \cup D(y)$. Since U is a Poisson point process with finite total intensity $\sum_{y \le N} b(y) + d(y)$, with probability 1 $U_N \cap [0, T]$ is finite for each fixed $T > 0$, which implies that $\lim_{i \to \infty} t_i$ must be infinite.

Having constructed the coupling, we verify the following desirable properties.

Lemma 1. *Let $\Phi(x, t)$ be as defined above. If $\Phi(x, s) = \Phi(y, s)$ then $\Phi(x, t) = \Phi(y, t)$ for all $t \ge s$, and if $x \le y$ then $\Phi(x, t) \le \Phi(y, t)$ for all $t \ge 0$.*

Proof. Suppose $\Phi(x, s) = \Phi(y, s)$, which we denote by z. Then, for some i,j, $s \ge \max(t_i(x), t_j(y))$ and $\Phi(x, t_i(x)) = \Phi(y, t_j(y)) = z$. If $b(z) + d(z) = 0$ then $I(x) = i$ and $I(y) = j$, and $\Phi(x, t) = \Phi(y, t) = z$ for all $t \ge s$. Otherwise, $I(x) > i$, $I(y) > j$, and $\min(t_{i+1}(x), t_{j+1}(y)) > s$. We then have $t_{i+1}(x) = t_{j+1}(y)$ since both are equal to $\inf\{t > s\colon t \in B(z) \cup D(z)\}$, and the construction ensures that $\Phi(x, t) = \Phi(y, t)$ for all $t \ge s$.

For the second statement, first note that since $(B(x) \cup D(x)\colon x \in \{0, \dots, N\})$ are independent, from standard properties of Poisson point processes it follows that if $x \ne y$ then $B(x) \cup D(x)$ and $B(y) \cup D(y)$ are disjoint. It follows from the construction that if $\Phi(x, t^-) \ne \Phi(y, t^-)$ then either $\Phi(x, t) = \Phi(x, t^-)$ or $\Phi(y, t) = \Phi(y, t^-)$, i.e. both cannot jump simultaneously. Since $\Phi(x, 0) - \Phi(y, 0)$ is integer valued, and since

- for any x, $t \mapsto \Phi(x, t)$ is piecewise constant,
- for any x and fixed $T > 0$, the set $\{t \in [0, T]\colon \Phi(x, t) \ne \Phi(x, t^-)\}$ is almost surely (a.s.) finite, and
- for any x,t, $|\Phi(x, t) - \Phi(x, t^-)| \in \{-1, 0, 1\}$,

if $x \le y$ and $\Phi(x, t) > \Phi(y, t)$ then, since $\Phi(x, 0) = x$ and $\Phi(y, 0) = y$, for some $s < t$, $\Phi(x, s) = \Phi(y, s)$. Using the first statement, we then have $\Phi(x, t) = \Phi(y, t)$, contradicting $\Phi(x, t) > \Phi(y, t)$. □

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

3.2. Natural coupling for two or more ordered processes

Next we describe a similar construction for multiple b–d processes with different transition rates on a common state space $\{0, \dots, N\}$. Index the processes $1, \dots, k$ and let $b_i(x)$, $d_i(x)$ denote the transition rates. We will consider only the case in which the rates are ordered such that if $i < j$ then $b_i(x) \le b_j(x)$ and $d_i(x) \ge d_j(x)$ for each x. Let $\beta_1(x) = b_1(x)$ and $\delta_k(x) = d_k(x)$; then, for $1 < i \le k$ let $\beta_i(x) = b_i(x) - b_{i-1}(x)$ and for $1 \le i < k$ let $\delta_i(x) = d_i(x) - d_{i+1}(x)$. Define Poisson point processes $B_i(x)$, $D_i(x)$ on $\mathbb{R}_+$ with respective rates $\beta_i(x)$, $\delta_i(x)$. Then, for each $i \in \{1, \dots, k\}$ define $(\Phi_i(x, t)\colon t \in [0, \infty))$ in the same way as Φ from the previous subsection, but using $\bigcup_{j=1}^{i} B_j(x)$, $\bigcup_{j=i}^{k} D_j(x)$, $\sum_{j=1}^{i} b_j(x)$, and $\sum_{j=i}^{k} d_j(x)$ in place of $B(x)$, $D(x)$, $b(x)$, and $d(x)$, respectively. Then $\Phi_i(x, t)$ is a copy of the b–d process with transition rates b_i, d_i and initial value x.

For fixed i, the above construction amounts to the same as in the previous subsection, so Lemma 1 applies to Φ_i. Another useful property is summarized in the following result.

Lemma 2. *Let $\Phi_i(x, t)$, $i = 1, \dots, k$, $x \in \{0, \dots, N\}$, $t \in [0, \infty)$ be as defined above. If $x \le y$ and $i \le j$ then $\Phi_i(x, t) \le \Phi_j(y, t)$ for all $t \ge 0$.*

Proof. For the same reason as in the proof of Lemma 1, if $x \le y$ and $\Phi_i(x, t) > \Phi_j(y, t)$ then, for some $s < t$, $\Phi_i(x, s) = \Phi_j(y, s)$. Moreover, s can be chosen so that in addition, for some $u \in (s, t]$, $\Phi_i(x, u^-) = \Phi_j(y, u^-) =: z$ and either

- $\Phi_i(x, u) = z + 1$ and $\Phi_j(y, u) = z$ or
- $\Phi_i(x, u) = z$ and $\Phi_j(y, u) = z - 1$.

The first case implies that $u \in \bigcup_{m=1}^{i} B_m(z)$ and $u \notin \bigcup_{m=1}^{j} B_m(z)$, which is impossible since $i \le j$. Similarly, the second case implies $u \in \bigcup_{m=j}^{k} D_m(z)$ and $u \notin \bigcup_{m=i}^{k} D_m(z)$, which again is impossible. □

4. Linear birth and death process

The linear birth and death process $\mathbb{Z}_+$ with parameter r is defined by the transitions

$$Z \to \begin{cases} Z + 1 & \text{at rate} \quad rZ, \\ Z - 1 & \text{at rate} \quad Z. \end{cases}$$

Z can be thought of as the number of cells in a process in which each cell independently dies at rate 1 and splits into two cells at rate r. For this process we are interested in how the extinction time $\tau = \inf\{t : Z(t) = 0\}$ scales with r and Z_0 in cases where extinction is asymptotically certain, i.e. $\mathbb{P}(\tau < \infty) \to 1$. The results are by now routine, but as I could not find a reference that states them all together, and since they are easy to prove, I have included the proof.

Theorem 6. *Let Z_n denote a sequence of copies of the above process, with respective initial condition and parameter $Z_n(0)$, r_n. Let $\tau_n = \inf\{t\colon Z_n(t) = 0\}$, $\gamma_n = 1 - r_n$, and $a_n = \gamma_n Z_n(0)$. Suppose that $Z_n(0) \to Z_\infty(0) \in \mathbb{N} \cup \{\infty\}$, $r_n \to r_\infty \le 1$, and $a_n \to a_\infty \in [0, \infty]$. Let $\gamma_\infty = \lim_{n\to\infty} \gamma_n$.*

1. *Suppose $Z_\infty(0) < \infty$ and fix any value of $t \ge 0$.*
   (a) *If $r_\infty = 1$ then $\mathbb{P}(\tau_n \le t) \to (1 + 1/t)^{-Z_\infty(0)}$.*
   (b) *If $r_\infty < 1$ then $\mathbb{P}(\tau_n \le t) \to (1 + \gamma_\infty/(e^{\gamma_\infty t} - 1))^{-Z_\infty(0)}$.*

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

2. *Suppose* $Z_\infty(0) = \infty$.

   (a) *If* $a_\infty \in [0, \infty)$ *then* $\tau_n / Z_n(0) \xrightarrow{(\mathrm{d})} H_{a_\infty}$ *with*

$$\mathbb{P}(H_{a_\infty} \le w) = \begin{cases} \mathrm{e}^{-1/w} & \text{if} \quad a_\infty = 0, \\ \mathrm{e}^{-a_\infty/(\mathrm{e}^{a_\infty w}-1)} & \text{if} \quad a_\infty > 0. \end{cases}$$

   (b) *If* $a_\infty = \infty$ *and* $a_n = b_n + o(b_n/\log b_n)$ *for some sequence* (b_n) *such that* $b_n \to \infty$ *then* $\gamma_n \tau_n - \log b_n \xrightarrow{(\mathrm{d})} G$, *where* G *has* $\mathbb{P}(G \le w) = \mathrm{e}^{-\mathrm{e}^{-w}}$ *for* $w \in \mathbb{R}$.

Proof of Theorem 6. For compactness of notation we will suppress the dependence on n in all variables and write Z_t instead of $Z(t)$. In [2, Chapter III] a more general model is considered in which each particle independently dies at some rate α (they call it a but that notation is already in use for us) and is replaced with k particles with probability p_k; the present model corresponds to $\alpha = 1 + r$ and $p_0 = 1/(1+r)$, $p_2 = r/(1+r)$. Using the Kolmogorov backward equation of the process [2, equation (4) of III.2], for $\rho(t) = \mathbb{P}(Z_t = 0 \mid Z_0 = 1)$ the differential equation $\rho' = \alpha(-\rho + \sum_{k=0}^{\infty} p_k \rho^k)$ is obtained [2, equation (2) of III.4]. In our case this gives $\rho' = -(1+r)\rho + 1 + r\rho^2$, which can be conveniently factored to give

$$\rho' = (1 - r\rho)(1 - \rho), \qquad \rho_0 = 0. \tag{5}$$

Separating variables,

$$\frac{\mathrm{d}\rho}{(1-r\rho)(1-\rho)} = \mathrm{d}t.$$

Letting $\gamma = 1 - r$ we note that

$$\frac{1}{1-\rho} - \frac{r}{1-r\rho} = \frac{\gamma}{(1-r\rho)(1-\rho)}.$$

Solving the differential equation and using $\rho(0) = 0$,

$$\gamma\, t = \log\left(\frac{1 - r\rho(t)}{1 - \rho(t)}\right).$$

If $\gamma \neq 0$, solving for $\rho(t)$ then gives

$$\rho(t)^{-1} = \frac{\mathrm{e}^{\gamma t} - r}{\mathrm{e}^{\gamma t} - 1} = 1 + \gamma/(\mathrm{e}^{\gamma t} - 1). \tag{6}$$

Part 1 of Theorem 6 then follows easily; the case $r_\infty = 1$ follows by taking the limit of $\rho(t)$ as $\gamma \to 0$, or it could be computed directly by solving the ODE in the special case $r = 1$.

We now tackle the second part of Theorem 6, using (6) to determine $\mathbb{P}(\tau \le t \mid Z_0) = \rho(t)^{Z_0}$ when $Z_0 \to \infty$ under various limits of γZ_0. As $Z_0 \to \infty$, $\rho(t)^{Z_0} \to \mathrm{e}^{-c}$ if $\gamma/(\mathrm{e}^{\gamma t} - 1) = c/Z_0$. Solving this gives $\gamma\, t = \log(1 + \gamma Z_0/c)$. Recalling that $a = \gamma Z_0$, there are then two cases.

Case 1: $a \to a_\infty \in [0, \infty)$. In this case, if $\rho(t)^{Z_0} \to \mathrm{e}^{-c}$ then $t = (Z_0/a)(\log(1 + a/c))$. Setting $w = \frac{1}{a}\log(1 + a/c)$ gives $c = a/(\mathrm{e}^{aw} - 1)$. If $a \to 0$ then $a/(\mathrm{e}^{aw} - 1) \to 1/w$ and $\mathbb{P}(\tau \le wZ_0) \to \mathrm{e}^{-1/w}$, while if $a \to a_\infty > 0$ then $\mathbb{P}(\tau \le wZ_0) \to \mathrm{e}^{-a_\infty/(\mathrm{e}^{a_\infty w}-1)}$.

Case 2: $a \to \infty$. Taking $\gamma\, t = \log a - \log c$ gives $\rho(t)^{Z_0} \to \mathrm{e}^{-c}$ as $Z_0 \to \infty$. Letting $w = -\log c$ so that $c = \mathrm{e}^{-w}$, $\mathbb{P}(\tau \le \gamma^{-1}(\log \gamma Z_0 + w)) \to \mathrm{e}^{-\mathrm{e}^{-w}}$ as $Z_0 \to \infty$. In other words, $\gamma\tau - \log \gamma Z_0$ has a standard Gumbel distribution.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

The following short lemma concludes the proof of case 2, and thus the proof of Theorem 6. □

Lemma 3. *Let Z_n be a sequence of copies of the linear birth and death process with initial value $Z_n(0) \to \infty$ and parameter $r_n = 1 - a_n/Z_n(0)$ with $a_n \to \infty$. Let (b_n) be a sequence with $b_n \to \infty$ and let $\tau_n = \inf\{t{:}Z_n(t) = 0\}$. Then*

$$\lim_{n\to\infty} \mathbb{P}\left(\tau_n \le \frac{Z_n(0)}{b_n}(\log b_n + w)\right) \to \mathrm{e}^{-\mathrm{e}^{-w}}$$

for all $w \in \mathbb{R}$ if and only if $a_n - b_n = o(b_n/\log b_n)$.

Proof. As before, suppress the dependence on n. Fix $w \in \mathbb{R}$ and let v be such that $\frac{1}{b}(\log b + w) = \frac{1}{a}(\log a + v)$. Since $\log a + v = \frac{a}{b}(\log b + u)$, subtracting $\log a + w$ from both sides gives

$$v - w = \log(b/a) + \frac{a-b}{b}\log b + \frac{a-b}{b}w. \tag{7}$$

If $a - b = o(b/\log b)$ the second term is $o(1)$. Since $b \to \infty$, $a = b + o(b)$ so $\log(b/a) = o(1)$, $\frac{a-b}{b}w = o(1)$, and so $v - w = o(1)$. On the other hand, suppose that $v - w = o(1)$ for every $w \in \mathbb{R}$. Setting $w = 0$ gives

$$\log(b/a) + \frac{a-b}{b}\log b = o(1). \tag{8}$$

Then, setting $w = 1$ and using (7) and (8) gives $\frac{a-b}{b} = o(1)$. This in turn implies that $\log(b/a) = o(1)$. Using (8) once more gives $\frac{a-b}{b}\log b = o(1)$, as required. □

5. Approximation by linear birth and death processes

In this section we use the method of [5], namely approximation by linear birth and death processes, to prove Theorem 1, Theorem 2, and the 'final stage' part of Theorem 3. As in the previous section, we suppress dependence on n and write X_t, Z_t, etc.

For $M > X_0 > 0$ (both depending on n) to be determined, let $r' = r(1 - M/n)$ and let Z and Z' be linear birth and death processes with respective parameters r and r' and common initial value X_0. Let $\tau_M = \inf\{t{:}X_t \in \{0, M\}\}$. Then, applying the natural coupling of Section 3 to Z',X,Z, labelled 1,2,3 respectively, and using Lemma 2 gives $Z'_t \le X_t \le Z_t$ for $t \le \tau_M$. Let $\tau_Z = \inf\{t{:}Z_t = 0\}$ and $\tau_{Z'} = \inf\{t{:}Z_t' = 0\}$. The goal is to take M large enough that $\mathbb{P}(X_{\tau_M} = M) = o(1)$, and small enough that $\mathbb{P}(X_t = Z_t \text{ for } t \le \tau_M) = 1 - o(1)$ if X_0 is fixed and τ_Z and $\tau_{Z'}$ have a common rescaled limit if $X_0 \to \infty$.

We begin with a general observation. Let $p_-(X)$ and $p_+(X)$ denote the probability that $X \to X - 1$ or $X \to X + 1$ in the embedded discrete-time Markov chain, or jump chain, of X_t. Then $p_-(X) = 1/(1 + r(1 - X/n))$ and $p_+(X) = r(1 - X/n)/(1 + r(1 - X/n))$; in particular, $p_-(X)/p_+(X) = 1/(r(1 - X/n)) \ge 1/r$ for all X, so it follows easily that r^{-X_t} is a supermartingale. Recall that $\tau = \inf\{t{:}\ X_t = 0\}$. Using optional stopping, which is applicable since X is bounded and $\mathbb{P}(\tau < \infty) = 1$, it is easy to show that

$$\mathbb{P}(\tau \ne \tau_M \mid X_0) = \mathbb{P}(X_{\tau_M} = M \mid X_0) \le \frac{r^{-X_0} - 1}{r^{-M} - 1}. \tag{9}$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

If $r < 1$ then using the estimate $a/b \le (a+c)/(b+c)$ that holds for $0 < a \le b$ and $c > 0$, we find that if $X_0, M > 0$ then

$$\frac{r^{-X_0}-1}{r^{-M}-1} \le \frac{r^{-X_0}}{r^{-M}} = (1-\gamma)^{M-X_0} \to 0 \qquad \text{if } \gamma\,(M-X_0) \to \infty. \tag{10}$$

Recall that $\gamma = 1 - r$. If $\gamma \to 0$ and if, a fortiori, $\gamma M \to 0$ then

$$\frac{r^{-X_0}-1}{r^{-M}-1} = \frac{(1-\gamma)^{-X_0}-1}{(1-\gamma)^{-M}-1} \sim \frac{X_0}{M} \to 0 \qquad \text{if } X_0 = o(M). \tag{11}$$

We proceed by cases. The first case is Theorem 1.

Proof of Theorem 1. In this setting, $\gamma \to \gamma_\infty \ge 0$ and we can assume X_0 is constant. For any $t \ge 0$,

$$|\mathbb{P}(\tau \le t \mid X_0) - \mathbb{P}(\tau_Z \le t \mid X_0)| \le \mathbb{P}(\tau \ne \tau_M \mid X_0) + \mathbb{P}(\tau = \tau_M,\, \tau_M \wedge t \ne \tau_Z \wedge t \mid X_0).$$

Theorem 1 then follows from Theorem 6 if M can be chosen so that, for any $t \ge 0$, the above $\to 0$ as $n \to \infty$.

We let $p_1 = \mathbb{P}(\tau \ne \tau_M \mid X_0)$ and $p_2 = \mathbb{P}(\tau = \tau_M,\, \tau_M \wedge t \ne \tau_Z \wedge t \mid X_0)$. We first find M such that $p_1 \to 0$. If $\gamma_\infty > 0$, using (10) we find that $p_1 \to 0$ if $M \to \infty$. If $\gamma_\infty = 0$, then using (11) we find that $p_1 \to 0$ if $M \to \infty$ and $\gamma M \to 0$.

We next find M such that $p_2 \to 0$. If $\tau = \tau_M$ then since X and Z are both absorbed at 0, $p_2 \le \mathbb{P}(X_s \ne Z_s$ for some $s \le t \wedge \tau_M \mid X_0)$. If $X, Z \in [0, M]$ their transition rates differ by at most $(r - r')M = rM^2/n$, so using the exponential distribution and noting that $\mathrm{e}^{-x} \ge 1 - x$ we find that $\mathbb{P}(X_s \ne Z_s$ for some $s \le t \wedge \tau_M \mid X_0) \le 1 - \mathrm{e}^{-rM^2 t/n} \le rM^2 t/n$. Since r is bounded in n, the right-hand side $\to 0$ for fixed t provided $M = o(\sqrt{n})$.

Thus, both $p_1, p_2 \to 0$ if we take $M \to \infty$ sufficiently slowly. □

Next we prove Theorem 2.

Proof of Theorem 2. We first reinterpret somewhat the conditions of the theorem.

First option: $c_\infty = -\infty$ and $a_\infty < \infty$. Since $c_\infty = -\infty$, $\frac{1}{\sqrt{n}} = o(|\delta_n|)$ so $a_\infty = \lim_{n\to\infty} |\delta_n| X_n(0)$. Since $a_\infty < \infty$, $X_n(0) = O(1/|\delta_n|) = o(\sqrt{n})$. In shorthand: $\gamma X_0 \to a_\infty \in [0, \infty)$ and $X_0 = o(\sqrt{n})$.

Second option: $c_\infty > -\infty$ and $a_\infty = 0$. If $c_\infty \in \mathbb{R}$ then $a_\infty = 0$ is equivalent to $X_n(0)/\sqrt{n} \to 0$, and implies $|\delta_n| X_n(0) \to 0$. If $c_\infty = \infty$ then, in the same way as above, $a_\infty = \lim_{n\to\infty} |\delta_n| X_n(0)$ and $X_n(0) = o(\sqrt{n})$. So, in this case as well, $\gamma X_0 \to a_\infty \in [0, \infty)$ and $X_0 = o(\sqrt{n})$.

Thus the theorem is proved if it can be proved when $\gamma X_0 \to a_\infty \in [0, \infty)$ and $X_0 = o(\sqrt{n})$. First note that, with p_1 as in the proof of Theorem 1,

$$|\mathbb{P}(\tau'_Z \le \tau \le \tau_Z \mid X_0)| \ge 1 - p_1. \tag{12}$$

Since $X_0 \to \infty$ and $\gamma X_0 \to a_\infty < \infty$, $\gamma \to 0$. For $W \in \{X, Z, Z'\}$, and writing τ as τ_X, let $F_W(t) = \mathbb{P}(\tau_W / X_0 \le t)$. From (12),

$$F_Z(t) - p_1 \le F_X(t) \le F_{Z'}(t) + p_1. \tag{13}$$

By Theorem 6, if $\gamma X_0,\ \gamma' X_0 \to a_\infty$ then, for all $t \ge 0$, $F_Z(t), F_{Z'}(t) \to \mathbb{P}(H_{a_\infty} \le t)$. Since $\gamma X_0 \to a_\infty$ by assumption, it is enough to find M such that (i) $p_1 \to 0$ and (ii) $(\gamma' - \gamma) X_0 \to 0$.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

For (i), using (10) and (11) we need either $\gamma M \to \infty$ or both $X_0/M \to 0$ and $\gamma M \to 0$. For (ii) we compute $(\gamma' - \gamma)X_0 = (r - r')X_0 = rX_0M/n$, so it is enough that $X_0M/n \to 0$.

Subcase 1: $a_\infty = 0$. Define $\beta = |\gamma| \vee \frac{1}{\sqrt{n}}$ and let $M = \beta^{-1}(\beta X_0)^{1/2}$. Since $\gamma X_0 \to 0$ and $X_0 = o(\sqrt{n})$, $\beta X_0 = \max(|\gamma|X_0,\ X_0/\sqrt{n}) \to 0$, so $X_0/M = (\beta X_0)^{1/2} \to 0$ and $\gamma M \le \beta M = (\beta X_0)^{1/2} \to 0$, satisfying (i). Since $\beta \ge \frac{1}{\sqrt{n}}$, $\beta^{-1} \le \sqrt{n}$ so $M = o(\sqrt{n})$ and $X_0M/n \to 0$, satisfying (ii).

Subcase 2: $a_\infty > 0$. Since $X_0 = o(\sqrt{n})$ and γX_0 has a positive limit, $\sqrt{n}\gamma \to \infty$. Let $M = \gamma^{-1}(\sqrt{n}\gamma)^{1/2}$. Then $\gamma M = (\sqrt{n}\gamma)^{1/2} \to \infty$, satisfying (i). Since $X_0 = o(\sqrt{n})$ and $M = (\sqrt{n}/\gamma)^{1/2} = \sqrt{n}/(\sqrt{n}\gamma)^{1/2} = o(\sqrt{n})$, $X_0M/n \to 0$, satisfying (ii). □

Now we prove Theorem 3 in the 'final stage', i.e. when $X_n(0) \le n\gamma_n|c_n|^{-\epsilon}$ for some $\epsilon > 0$. The proof is broadly the same as in [5], although here the result is applicable to a somewhat larger set of initial values.

Proof of Theorem 3 when $X_n(0) \le n\gamma_n|c_n|^{-\epsilon}$ for some $\epsilon > 0$.

Using our abbreviated notation, we shall prove the result under the slightly less restrictive condition $X_0 \log(\gamma X_0) = o(\gamma n)$. To see that this is less restrictive, suppose $X_0 \le \gamma n|c|^{-\epsilon}$, so that $X_0|c|^\epsilon \le \gamma n$. By assumption in Theorem 3, $|c| \to \infty$ and $a = \gamma X_0 \to \infty$. Using $|c| \to \infty$, $\log|c| = o(|c|^\epsilon)$ so $X_0 \log|c| = o(X_0|c|^\epsilon) = o(\gamma n)$. Using $a \to \infty$ and $X_0 = o(\gamma n)$, $\log(\gamma X_0) = o(\log(\gamma^2 n)$. Since $|c| = \gamma\sqrt{n}$, $\log(\gamma X_0) = o(\log(|c|)$. Therefore, $X_0 \log(\gamma X_0) = o(X_0 \log|c|) = o(\gamma n)$ as desired.

If $X_0 \log(\gamma X_0) = o(\gamma n)$ then, since $\gamma X_0 \to \infty$, $X_0 = o(\gamma n)$. As noted in the statement of Theorem 3, in this case $g = \log(\gamma X_0) + o(1)$. So, we want to show that $\gamma\tau - \log(\gamma X_0)$ converges to standard Gumbel.

Let $a = \gamma X_0$ and for $W \in \{X, Z, Z'\}$ let $F_W(t) = \mathbb{P}(\gamma\ \tau_W - \log a \le t)$. Then, (13) also holds for this choice of $F_W(t)$. According to Theorem 6, for all $t \ge 0$ $F_Z(t) \to \mathbb{P}(G \le t)$, and if $(\gamma' - \gamma)X_0 = o(a/\log a)$ then, for all $t \ge 0$, $F_{Z'}(t) \to \mathbb{P}(G \le t)$. Following again the logic of the proof of Theorem 2, it is enough to find M such that (i) $p_1 \to 0$ and (ii) $(\gamma' - \gamma)\,X_0 = o(a/\log a)$.

For (i) we need $\gamma\,(M - X_0) \to \infty$, for which $M = 2X_0$ suffices (and which, assuming only $\gamma X_0 \to \infty$, cannot be improved beyond a factor of 2). For (ii) we need $X_0M/n = o(\gamma X_0/\log(\gamma X_0))$. Using $M = 2X_0$, the condition becomes $X_0 = o(\gamma n/\log(\gamma X_0))$, which is the condition given. □

6. Subthreshold cutoff

In this section we prove Theorem 3 in the case where $X_n(0) \ge n\gamma_n|c_n|^{-\epsilon}$ for any $\epsilon > 0$. This corresponds to the intermediate and initial stages as described in [5]; the final stage, where $X_n(0) \le n\gamma_n|c_n|^{-\epsilon}$, was proved in the previous section. The proof for the initial stage is just about identical to their proof. For the intermediate stage, the setup of the problem follows their approach, then I use a different method to handle the error term between the process and its deterministic approximation (denoted e_N in [5] and W here). In [5], an auxiliary result (Lemma 3.2 there) is used to show that the maximum of the error term is (deterministically) bounded by twice the maximum of the compensator of X, then the latter is estimated using the corresponding exponential martingale. Here, we compute the drift and diffusivity of the error term, then after making a time change we use a so-called drift barrier estimate, Lemma 19, which is proved in [4], to show that the error term remains small on the desired time interval. Neither approach seems to be strictly simpler or more efficient than the other.

Proof of Theorem 3 when $X_n(0) \ge n\gamma_n|c_n|^{-\epsilon}$. As before, suppress n from the notation and write X_t, Y_t, etc. when it is convenient. For this proof only, let $c = \sqrt{n}\gamma$ which amounts to

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

a change of sign; this saves us from always writing $|c|$. Fix a small $\epsilon > 0$, to be determined, and let $Y_t = X(t/\gamma)/(\gamma n)$. Then $\gamma X_0 = c^2 Y_0$. Let us rewrite $g_n(X_n(0))$ from the statement of Theorem 3 in terms of Y. Abusing notation slightly, we have $g(Y_0) = 2\log c - \log(r + 1/Y_0)$.

Let $t_\star = \inf\{t\colon Y_t \le c^{-\epsilon}\} = \inf\{t\colon X(t/\gamma) \le n\gamma c^{-\epsilon}\}$. The final stage corresponds to $t_\star = 0$. Since X jumps by ± 1, Y jumps by $\pm 1/\gamma n = \pm 1/c\sqrt{n} = o(1/c)$, so $g(Y_0) - g(Y(t_\star)) = -\log(r + 1/Y_0) + \log(c^{-\epsilon} \pm o(1/c)) = -\log(r + 1/Y_0) - \epsilon \log c + o(1)$. Thus, to prove the result it remains to show that

$$t_\star - (-\log(r + 1/Y_0) - \epsilon \log c) \to 0. \tag{14}$$

For the intermediate stage, suppose $c^{-\epsilon} \le Y_0 \le c^{\epsilon}$. We will use the notation for drift and diffusivity discussed in the Appendix. Using the transition rates from (2) in (54) and noting that $\gamma = -(r-1)$, for X_t we compute $\mu(X) = rX(1 - X/n) - X = -\gamma X - rX^2/n$ and $\sigma^2(X) = (1 + r(1 - X/n))X \le (1 + r)X$. From (54) we infer that if $Y(t) = \alpha X(\beta t)$ then $\mu(Y) = \alpha\beta\mu(X)$ and $\sigma^2(Y) = \alpha^2\beta\sigma^2(X)$. So,

$$\mu(Y) = \frac{1}{\gamma^2 n}\mu(X) = -\frac{1}{\gamma n}X - r\frac{1}{(\gamma n)^2}X^2 = -Y - rY^2$$

and, using $\gamma^2 n = c^2$,

$$\sigma^2(Y) = \frac{1}{\gamma^3 n^2}\sigma^2(X) \le \frac{1}{\gamma c^2}(1 + r)X = \frac{1}{c^2}(1 + r)Y.$$

As in [5], we directly estimate the distance between Y and its deterministic approximation. Let $y(t)$ denote the solution to the initial value problem $y' = -y - ry^2$, $y(0) = Y_0$, and let $t^{\pm} = \inf\{t : y(t) = (1 \pm c^{-\epsilon})c^{-\epsilon}\}$. Solving by separation of variables,

$$\begin{aligned} t^{\pm} &= \log\frac{Y_0}{1 + rY_0} - \log\frac{(1 \pm c^{-\epsilon})c^{-\epsilon}}{1 + r(1 \pm c^{-\epsilon})c^{-\epsilon}} \\ &= -\log(r + 1/Y_0) - \epsilon\log c + o(1). \end{aligned}$$

Thus, $t^{\pm}$ both have the desired limit for $t_\star$ as in (14). The result will be proved if we show that $t^- \le t_\star \le t^+$ with probability $1 - o(1)$. Define the error process $W_t = Y_t - y(t)$. If $\sup_{t \le t^-} |W_t| \le c^{-2\epsilon}$ then $t^- \le t_\star \le t^+$, so we will show that the former has probability $1 - o(1)$. We first compute the drift and diffusivity of W. Factoring the difference of squares, $\mu(W) = \mu(Y) - y' = -Y - rY^2 + y - ry^2 = -W(1 + r(Y + y))$. Since $(y(t))$ is continuous and has finite variation it has zero quadratic variation, so

$$\sigma^2(W) = \sigma^2(Y) \le \frac{1}{\gamma^2}(1 + r)Y.$$

Since $r, y, Y > 0$,

$$\operatorname{sgn}(\mu(W)) = -\operatorname{sgn}(W), \qquad \frac{|\mu(W)|}{\sigma^2(W)} \ge \frac{c^2 r}{1 + r}|W|. \tag{15}$$

Next, we change time from t to s such that $\mu_s(W) = -W_s$. To do so, let $s(t) = \int_0^t (1 + r(Y_r + y(r))\,\mathrm{d}r$. Since $\mu(W)$ and $\sigma^2(W)$ are both scaled by $\mathrm{d}t/\mathrm{d}s \in (0, 1]$, (15) remains valid

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

after the time change, so $\mu_s(W) = -W_s$ and $\sigma_s^2(W) = O(1/c^2)$. Also, if $|W_s| \le c^{-2\epsilon}$ then $Y_s \le y(s) + c^{-2\epsilon} \le y(0) + c^{-2\epsilon} \le c^\epsilon + c^{-2\epsilon}$, so $s'(t) \le 1 + r(2c^\epsilon + c^{-2\epsilon}) = O(c^\epsilon)$. Since $t^- = O(\log c)$, if $\sup_{s \le s(t^-)} |W_t| \le c^{-2\epsilon}$ then $s(t^-) = O(c^\epsilon \log c) = o(c)$ so it is enough to show that $\sup_{s \le c} |W_s| \le c^{-2\epsilon}$.

We first give an upper bound on W. Next we use Lemma 19, which is proved in [4]. In the notation of Lemma 19, let $x = c^{-2\epsilon}/2$, $X = W - x$, $\Delta_\infty(X) = 1/\gamma n$, $\mu_\star = x$, $\sigma_\star^2 = C/c^2$ for some constant C, $C_\mu = 2x$, and, since $\Delta_\infty(X)\mu_\star/\sigma_\star^2 = c^2 x/C\gamma n = \gamma^2 n x/C\gamma n \le \gamma/C = O(1)$, take C_Δ to be some large enough constant. Then, $\Gamma = \exp(x^2/16\sigma_\star^2) = \exp(c^{2-2\epsilon}/16C) \ge c$ for large enough c, and since $c \to \infty$ by assumption, with probability $1 - o(1)$, $X_s \le x$ or equivalently $W_s \le 2x = c^{-2\epsilon}$ for all $s \le c$. A matching bound for $-W$ is proved in the same way.

For the initial stage, suppose $Y_0 \ge c^\epsilon$ and let $t^\star = \inf\{t : Y_t \le c^\epsilon\}$. If $Y_0 \ge c^\epsilon$ then $Y_0 \to \infty$ so $g(Y_0) = 2\log c - \log r + o(1)$. Since the result is proved for $Y_0 \le c^\epsilon$ it is enough to show that $t^\star = o(1)$. From the drift of Y and Jensen's inequality applied to $\mathbb{E}[Y_t^2]$ we find that $u(t) = \mathbb{E}[Y_t]$ satisfies the differential inequality $u' \le -u - ru^2 \le -ru^2$. Integrating the inequality, $u(t) \le (u(0)^{-1} + rt)^{-1} \le 1/rt$. By Markov's inequality, $Y_t \le c^\epsilon$ or equivalently $t^\star \le t$ with high probability if $u(t) = o(c^\epsilon)$, which is the case if $t = c^{-\epsilon/2}$. Since $c^{-\epsilon/2} = o(1)$, the result is proved. □

7. Non-linear diffusive

Proof of Theorem 4. The first part of the proof is to show convergence to the limiting diffusion, which is done in [7] by convergence of generators; here, we do it using a general result, Lemma 20, that requires convergence of drift and diffusion coefficients and vanishing jump size. The second part of the proof is to show the extinction time is short when the initial value is small, relative to the space and time scale of the limiting diffusion.

For this section let $Y_n(t) = X_n(\sqrt{n}t)/\sqrt{n}$ and let $c_n = \sqrt{n}(r_n - 1)$ as in the statement of the theorem. There are two cases to cover: $Y_n(0) \to y \in (0, \infty)$ and $Y_n(0) \to \infty$.

Case 1: $Y_n(0) \to y \in (0, \infty)$. First we use Lemma 9 in the Appendix, which is a result from [9], to show that if $Y_n(0) \to y$ then, for all but countably many R, $Y_n(\cdot \wedge \tau_n^R) \xrightarrow{(\mathrm{d})} Y(\cdot \wedge \tau^R)$, where τ_n^R and τ^R are the exit times of Y_n and Y from $\left(\frac{1}{R}, R\right)$ as described in Lemma 20. Recall that $Y_n(t)$ has transitions

$$Y_n \to \begin{cases} Y_n + \frac{1}{\sqrt{n}} & \text{at rate} \quad nY + \sqrt{n}Y(c_n - Y) - c_n Y, \\ Y_n - \frac{1}{\sqrt{n}} & \text{at rate} \quad nY, \end{cases}$$

so Y_n has jump size $\frac{1}{\sqrt{n}} = o(1)$ and $\mu(Y_n) = Y(c_n - Y) - c_n Y/\sqrt{n}$, $\sigma^2(Y_n) = 2Y + Y(c_n - Y)/\sqrt{n} - c_n Y/n$. For $|y| \le R$, $\mu(y) \to b(y)$ and $\sigma^2(y) \to a(y)$ uniformly, where $b(y) = y(c_\infty - y)$ and $a(y) = 2y$. Note that b and $\sqrt{a}$ are Lipschitz on compact subsets of $(0, \infty)$. By Lemma 20, the desired convergence holds.

For $y > 0$ define the mapping $T_y(f) = \inf\{t : f(t) \le y\}$ from cÀdlÀg functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ with the topology of uniform convergence on compacts. If $f_i \to f$ and f is continuous then, since $\inf\{f(t) : t \in [0, T_y(f) - \epsilon]\} > y$ for any $\epsilon > 0$, it follows that $\liminf_i T_y(f_i) \ge T_y(f)$. On the other hand, if $y > 0$ then, for any $\epsilon > 0$, $Y(T(y) + \epsilon)$ intersects $[0,y)$ since its diffusion coefficient at y is non-zero. In other words, $\mathbb{Q}(f : \text{if } f_i \to f \text{ then } \limsup_i T_y(f_i) \le T_y(f)) = 1$, where $\mathbb{Q}$ is the law of Y. Combining the two, the discontinuity points of T_y have $\mathbb{Q}$-measure 0.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Let $T_n(y) = T_y(Y_n)$ and $T(y) = T_y(Y)$, and let $T_n = T_n(0)$ and $T = T(0)$. By the continuous mapping theorem and convergence of Y_n to Y, $T_n(y) \xrightarrow{(\mathrm{d})} T(y)$ for $y > 0$. Let $\mathbb{P}_n$ be the law of Y_n. Since $T_n \geq \lim_{y\to 0^+} T_n(y)$ and $T = \lim_{y\to 0^+} T(y)$,

$$\limsup_n \mathbb{P}_n(T_n \leq t) \leq \limsup_{y\to 0^+} \limsup_n \mathbb{P}_n(T_n(y) \leq t) = \lim_{y\to 0^+} \mathbb{Q}(T(y) \leq t) = \mathbb{Q}(T \leq t).$$

To obtain the opposite inequality it is enough to show that for any $\epsilon, t > 0$ there are α, n_0 such that $\mathbb{P}_n(T_n > t \mid Y_n(0) \leq \alpha) \leq \epsilon$ for $n \geq n_0$, since then

$$\begin{aligned}\liminf_n \mathbb{P}_n(T_n \leq t) &\geq \liminf_n \mathbb{P}_n(T_n(\alpha) \leq t) \inf_{x\leq\alpha} \mathbb{P}_n(T_n \leq t \mid Y_n(0) = x)\\ &\geq (1-\epsilon) \lim_{n\to\infty} \mathbb{P}_n(T_n(\alpha) \leq t)\\ &= (1-\epsilon)\mathbb{Q}(T(\alpha) \leq t) \geq (1-\epsilon)\mathbb{Q}(T \leq t).\end{aligned}$$

Returning to the original time scale, X_n is dominated by the linear birth and death process Z with parameter r and $Z_0 = X_n(0)$, so, using $\rho(t)$ as in the proof of Theorem 6, $\mathbb{P}(X_n(\sqrt{n}t) > 0 \mid X_n(0) \leq \sqrt{n}\alpha) \leq 1 - (\rho(\sqrt{n}t))^{\sqrt{n}\alpha}$. If $c_n \leq 0$ we can take $r = 1$ which has $\rho(t) = 1 - 1/(1+t)$, so $1 - (\rho(\sqrt{n}t))^{\sqrt{n}\alpha} \leq \alpha/t$ is at most ϵ if $\alpha \leq \epsilon/t$.

If $c_n > 0$, rewrite $\rho(t)$ from (6) with $\delta = r - 1$ instead of $\gamma = 1 - r$ to obtain

$$\rho(t)^{-1} = \frac{r\mathrm{e}^{\delta t} - 1}{\mathrm{e}^{\delta t} - 1} = 1 + \frac{\delta}{1 - \mathrm{e}^{-\delta t}}.$$

Using $\delta_n = c_n/\sqrt{n}$,

$$(\rho(\sqrt{n}t))^{\sqrt{n}\alpha} = \left(1 + \frac{c_n/\sqrt{n}}{1 - \mathrm{e}^{-c_n t}}\right)^{-\sqrt{n}\alpha} \to \exp\left(-c_\infty \alpha/(1 - \mathrm{e}^{-c_\infty t})\right) \qquad \text{as } n \to \infty,$$

and since for fixed t the limit $\to 1$ uniformly as $\alpha \to 0$, we are done.

Case 2: $Y_n(0) \to \infty$. It remains to show the results of Step 2 are true for $y = \infty$. First we need to make sense of T when $Y(0) = \infty$. Let $T(y, w) = \inf\{t : Y(t) \leq w \mid Y(0) = y\}$. Since Y is continuous, and using the strong Markov property,

$$T(y, 0) \overset{(\mathrm{d})}{=} T(y, w) + T(w, 0), \tag{16}$$

where the last two are independent. In particular, $T(y,0)$ dominates $T(w,0)$ for $y > w$. On the other hand, letting $U = Y - c_\infty$, $\mu(U) = -U(U + c_\infty) \leq -U^2$ so, integrating and using Jensen's inequality, $\mathbb{E}[U(t)] \leq (1/\mathbb{E}[U_0] + t)^{-1} \leq 1/t$. Using Markov's inequality, if $y > w > c_\infty$ then $\mathbb{P}(T(y, w) > t) \leq \mathbb{P}(U(t) > w - c_\infty \mid Y(0) = y) \leq ((w - c_\infty)t)^{-1}$. It follows that $T(y, w) \xrightarrow{(\mathrm{p})} 0$ uniformly over $y \in [w, \infty)$ as $w \to \infty$. Combining with (16), there exists $T(\infty, 0)$ such that $T(y, 0) \xrightarrow{(\mathrm{d})} T(\infty, 0)$ as $y \to \infty$. A similar argument shows that for $T_n(y, 0) = \inf\{t : Y_n(t) = 0 \mid Y_n(0) = y\}$ there is a limit $T_n(\infty, 0)$. By Step 2, $T_n(y, 0) \xrightarrow{(\mathrm{d})} T(y, 0)$ for each $y > 0$, so it follows that $T_n(\infty, 0) \xrightarrow{(\mathrm{d})} T(\infty, 0)$. □

8. Threshold and metastable

In this section we prove Theorem 5. As we've done so far, we'll avoid writing subscript n on everything. So, for example $X_n^\star$ is simply denoted $X_\star$, etc. Since we need some additional

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

decorations later, we move the $\star$ from $\tau_n^\star$ into the subscript, so $\tau_\star = \inf\{t\colon X_t \in \{0, X_\star\}\}$. We begin with some basic theory and estimation of an important function. Let $q_+(j) = rj(1-j/n)$ and $q_-(j) = j$ denote the transition rates of X, and for later use let $q(j) = q_+(j) + q_-(j)$. For integer j, define $h_+(j) = \mathbb{P}(X_{\tau_\star} = X_\star \mid X_0 = j)$ and $h_-(j) = \mathbb{P}(X_{\tau_\star} = 0 \mid X_0 = j) = 1 - h_+(j)$.

By definition, $h_+(X(t\wedge\tau_\star))$ is a martingale, as is $h_-(X(t\wedge\tau_\star))$. Using the generator of the process, it follows that $q_+(j)(h_+(j+1) - h_+(j)) + q_-(j)(h_+(j-1) - h_+(j)) = 0$ and similarly for h_-. Let $\nu(0) = 1$, and for $j \geq 1$ let $\nu(j) = \prod_{i=1}^{j} q_-(i)/q_+(i)$. Using the linear equations for h_+, h_- and the boundary conditions $h_+(0) = 0, h_+(X_\star) = 1, h_-(0) = 1, h_-(X_\star) = 0$, we can solve to find

$$h_+(j) = \frac{\sum_{k=0}^{j-1} \nu(k)}{\sum_{k=0}^{X_\star - 1} \nu(k)}, \qquad h_-(j) = \frac{\sum_{k=j}^{X_\star - 1} \nu(k)}{\sum_{k=0}^{X_\star - 1} \nu(k)}. \tag{17}$$

The solution of the above linear equations for h_+, h_- to obtain (17) is not hard; it can be found, for example, in [8, Example 5.3.9]. We begin by estimating $\nu(k)$. Since it is no more difficult to estimate, and since we will need it later, we estimate the more general $\nu(j,k) = \prod_{i=j+1}^{k} q_-(i)/q_+(i)$, defined for $0 \leq j < k < n$; we recover from it $\nu(k) = \nu(0,k)$. It will be helpful to have both a general upper bound and a precise estimate.

Lemma 4. *Let $V(x) = x(\log r - 1) - (1-x)\log(1-x)$. Then, for integer na and nb with $0 \leq a < b \leq 1$, $\nu(na, nb) \leq \exp(-n(V(b+1/n) - V(a+1/n)))$ and*

$$\nu(na, nb) = \sqrt{\frac{1-a}{1-b}} \exp(-n(V(b) - V(a))E_n(a,b),$$

with $|\log E_n(a,b)| \leq (12n(1-b)^2(b-a))^{-1}$.

Proof. Since $q_+(i)/q_-(i) = r(1-i/n)$,

$$-\log \nu(j,k) = \sum_{i=j+1}^{k} (\log r + \log(1-i/n)) = (k-j)\log r - \sum_{i=j+1}^{k} f(i/n), \tag{18}$$

where $f(x) = -\log(1-x)$ is positive and increasing for $x \in (0,1)$. Since $f(x)$ has the antiderivative $x + (1-x)\log(1-x)$, the upper bound follows. Using a trapezoidal approximation with $k-j$ subintervals of size $1/n$ and writing the approximation as an upper Riemann sum minus a telescoping triangular correction,

$$\int_{j/n}^{k/n} f(x)\,\mathrm{d}x = \frac{1}{n}\sum_{i=j+1}^{k} f(i/n) - \frac{1}{2n}(f(k/n) - f(j/n)) + R_n(j,k), \tag{19}$$

where the error term (see [21] for a simple proof) has the bound

$$|R_n(j,k)| \leq \frac{\max_{x\in[j/n,k/n]} |f''(x)|}{12(k/n - j/n)n^2} = \frac{1}{(1-k/n)^2}\frac{1}{12(k/n - j/n)n^2}.$$

Using the antiderivative of f together with (18) and (19),

$$-\frac{1}{n}\log \nu(na, nb) = V(b) - V(a) + \frac{1}{2n}\log\frac{1-b}{1-a} + R_n(na, nb),$$

and the precise estimate follows. □

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Note that $V(x)$ has $V(0)=0$, $V'(x)=\log(r(1-x))$, and $V''(x)=-1/(1-x)$. In particular, it is concave on [0,1) and has $V'(x_\star)=\log 1=0$, so is increasing and positive on $(0, x_\star)$ and decreasing on $(x_\star, 1)$, with maximum $V_\star=V(x_\star)=\log r+1/r-1>0$, and has $V''(x_\star)=-1/(1-(1-1/r))=-r$. If $\delta_\infty=\lim_n \delta>0$ then $V_\star$ has a positive limit, while if $\delta\to 0$ then $V_\star=\log(1+\delta)+1/(1+\delta)-1=\delta-\delta^2/2-\delta+\delta^2+O(\delta^3)\sim\delta^2/2$.

8.1. Extinction probability

In this section we prove the estimate of extinction probability in Theorem 2. The result in the case $\delta_\infty>0$ belongs to [1, (A2) of Theorem 1], while in the case $\delta\to 0$ it is new. The approach is to estimate $h_-(X_0)$ for h_- as in (17), and the proof works by estimating the values of $\nu(k)$ separately for small and large k.

Lemma 5. *Suppose $\delta X_0\to a_\infty\in(0,\infty)$. Then,*

$$\mathbb{P}(X_{\tau_\star}=0)\to\begin{cases}\mathrm{e}^{-a_\infty} & \text{if}\quad \delta\to 0,\\ (1+\delta_\infty)^{-X_0} & \text{if}\quad \delta\to\delta_\infty>0.\end{cases}$$

If $\delta X_0\to\infty$ then $\mathbb{P}(X_{\tau_\star}=X_\star)\to 1$.

Proof. The quantity of interest is

$$\mathbb{P}(X_{\tau_\star}=0)=h_-(X_0)=\frac{\sum_{k=X_0}^{X_\star-1}\nu(k)}{\sum_{k=0}^{X_\star-1}\nu(k)}.$$

Since $q_-(k)/q_+(k)=(r(1-k/n))^{-1}$, for any $M\ge 1$ and $k\in\{1,\dots,M\}$, $r^{-k}\le\nu(k)\le(r(1-M/n))^{-k}$ and so

$$r^{-j}\frac{1-r^{-(M-j)}}{1-1/r}\le\sum_{k=j}^{M-1}\nu(k)\le r^{-j}\frac{1}{1-(r(1-M/n))^{-1}}. \tag{20}$$

Case 1: $\delta\to 0$. Suppose M is taken large enough that $\delta(M-j)\to\infty$, and small enough that $M=o(\delta n)$. Then $r^{-(M-j)}=(1+\delta)^{-(M-j)}\le\mathrm{e}^{-\delta(M-j)}\to 0$ and $r(1-M/n)=(1+\delta)(1-M/n)=1+\delta+o(\delta)$, so that $1-1/(r(1-M/n))=1-1/(1+\delta+o(\delta))=1-(1-\delta+o(\delta))\sim\delta$. Similarly, $1-1/r\sim\delta$. Since $\delta\to 0$, $r^{-j}=(1+\delta)^{-j}\sim\mathrm{e}^{-\delta j}$, so

$$\sum_{k=j}^{M-1}\nu(k)\sim\mathrm{e}^{-\delta j}/\delta. \tag{21}$$

We have $V'(0)=\log r=\delta-o(\delta)$ and $x_\star=\delta/r=o(1)$, so $V(x)\sim x$ uniformly over $x\in[0,x_\star]$ as $n\to\infty$. Using the upper bound from Lemma 4 with $a=0$ and nb in the range $\{M,\dots,X_\star-1\}$, for large n, $\nu(nb)=\nu(0,nb)\le\exp(-\delta nb/2)$. Summing over nb,

$$\sum_{k=M}^{X_\star-1}\nu(k)\le\frac{\mathrm{e}^{-\delta M/2}}{1-\mathrm{e}^{-\delta/2}}=o(1/\delta),$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

since $\delta M \to \infty$ by assumption and $1 - \mathrm{e}^{-\delta/2} \sim \delta/2$. Noting that $1 - 1/r \sim \delta$ and combining with (21),

$$\sum_{k=j}^{X_\star - 1} \nu(k) \sim \mathrm{e}^{-\delta j}/\delta. \tag{22}$$

Using the values $j = 0$ and $j = X_0$, we conclude that if $\delta M - \delta X_0 \to \infty$ (which also implies $\delta M \to \infty$) and $M = o(\delta n)$, then

$$h_-(X_0) \sim \frac{\mathrm{e}^{-\delta X_0}/\delta}{\mathrm{e}^{-0}/\delta} = \mathrm{e}^{-\delta X_0} \to \mathrm{e}^{-a_\infty}.$$

If $a_\infty < \infty$, since δX_0 has a finite limit and $\sqrt{n}\delta \to \infty$ it is easy to check that $M = \sqrt{n}(\sqrt{n}\delta)^{1/2}$ satisfies the conditions. If $a_\infty = \infty$, since $\mathrm{e}^{-a_\infty} = 0$ and $h_-(X_0)$ decreases with X_0 it is enough to consider the case where $\delta X_0 \to \infty$ arbitrarily slowly; thus, to satisfy the condition $\delta M - \delta X_0 \to \infty$ it is sufficient that $\delta M \to \infty$, for which the above choice of M suffices.

Case 2: $\delta \to \delta_\infty > 0$. Note that $r \to r_\infty = (1 + \delta_\infty) > 1$. Also, the condition $\delta X_0 \to a_\infty$ is equivalent to X_0 eventually being constant if $a_\infty < \infty$, and to $X_0 \to \infty$ if $a_\infty = \infty$. Suppose that $(M - j) \to \infty$ and $M = o(n)$. Then, $r^{-(M-j)} \to 0$ and $1 - M/n \to 1$. From (20) we find

$$\sum_{k=j}^{M-1} \nu(k) \sim r^{-j}/(1 - 1/r_\infty). \tag{23}$$

Since $V(0) = 0$, V is concave and both $x_\star$ and $V(x_\star)$ have a positive limit; for some constant $C_1 > 0$, eventually $V(x) \geq C_1 x$ for $x \in [0, x_\star]$. Using Lemma 2 as before with $a = 0$ and nb in the range $\{M, \ldots, X_\star - 1\}$, since δ is bounded by assumption, $1 - b \geq 1 - x_\star$, which has a positive lower bound, and $nb \geq nM \to \infty$, so uniformly $|\log E_n(a, b)| \to 0$ and for large n and some constant $C_2 > 0$, $\nu(nb) = \nu(0, nb) \leq C_2 \mathrm{e}^{-C_1 nb}$. Again, summing over nb,

$$\sum_{k=M}^{X_\star - 1} \nu(k) \leq C_2 \frac{\mathrm{e}^{-C_1 M}}{1 - \mathrm{e}^{-C_1}} \to 0.$$

Using the values $j = 0$ and $j = X_0$ for constant $X_0 < \infty$ and letting $M = \sqrt{n}$, combining the above with (23) we find

$$h_-(X_0) = \frac{r^{-X_0}/(1 - 1/r_\infty) + o(1)}{1 - 1/r_\infty + o(1)} \to r_\infty^{-X_0}.$$

If $X_0 \to \infty$ we may again assume it does so arbitrarily slowly, in which case $M = \sqrt{n}$ again suffices. Using $j = 0$ and $j = X_0$ as above we find the numerator $\to 0$ while the denominator $\to 1 - 1/r_\infty > 0$, so $h_-(X_0) \to 0$ as desired. □

8.2. Rapid extinction

Next we prove the results on rapid extinction from Theorem 5. Define the probability measures $P^\star$ and P^0 for events E by

$$P^\star(E) = \mathbb{P}(E \mid X_{\tau_\star} = X_\star), \qquad P^0(E) = \mathbb{P}(E \mid X_{\tau_\star} = 0). \tag{24}$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Using the well-known Doob h-transform (which can be found by computing the generator of the conditioned process), we find that with respect to $P^\star$ and P^0 respectively, for $t < \tau_\star$, X is a continuous-time Markov chain with transition rates

$$\begin{aligned} q^\star_+(j) &= q_+(j)\frac{h_+(j+1)}{h_+(j)}, \qquad & q^\star_-(j) &= q_-(j)\frac{h_+(j-1)}{h_+(j)}; \\ q^0_+(j) &= q_+(j)\frac{h_-(j+1)}{h_-(j)}, \qquad & q^0_-(j) &= q_-(j)\frac{h_-(j-1)}{h_-(j)}. \end{aligned} \tag{25}$$

The following lemma is an equivalent formulation of the rapid extinction results of Theorem 5.

Lemma 6. *Suppose $\delta X_0 \to a_\infty \in (0, \infty)$ and let P^0 be as in (24).*

1. *If $\delta \to 0$, then, for $w \geq 0$, $P^0(\tau \leq wX_0) \to \exp(-a_\infty/(\mathrm{e}^{-a_\infty w} - 1))$.*
2. *If $\delta \to \delta_\infty > 0$, then, letting $\gamma_\infty = \delta_\infty/r_\infty$, $P^0(\tau \leq t/r_\infty) \to (1 + \gamma_\infty/(\mathrm{e}^{-\gamma_\infty t} - 1))^{-X_0}$.*

Proof. We first estimate the transition rates (25) with respect to P^0, then approximate by a linear birth and death process as in the proof of Theorems 1 and 2 to obtain the scaling of τ.

$$\frac{h_-(j+1)}{h_-(j)} = \frac{\sum_{k=j+1}^{X_\star - 1} \nu(k)}{\sum_{k=j}^{X_\star - 1} \nu(k)} = 1 - \frac{\nu(j)}{\sum_{k=j}^{X_\star - 1} \nu(k)} = 1 - \frac{1}{\sum_{k=j}^{X_\star - 1} \nu(j,k)}; \tag{26}$$

similarly,

$$\frac{h_-(j-1)}{h_-(j)} = 1 + \frac{1}{\sum_{k=j}^{X_\star - 1} \nu(j-1,k)}. \tag{27}$$

Again, we divide by cases.

Case 1: $\delta \to 0$. Let $c = \sqrt{n}\delta$ so that $c \to \infty$ and $c = o(\sqrt{n})$ by assumption, let $M = \lfloor 2\sqrt{n/c} \rfloor$, so that $M/\sqrt{n} = o(1)$, and let $M' = \lfloor \sqrt{n} \rfloor$. For $j \leq M$, write

$$\sum_{k=j}^{X_\star - 1} \nu(j,k) = \sum_{k=j}^{M'-1} \nu(j,k) + \sum_{k=M'}^{X_\star - 1} \nu(j,k). \tag{28}$$

If $j \leq k < M'$ then $r^{-(k-j)} \leq \nu(j,k) \leq (r(1 - M'/n))^{-(k-j)}$, so, uniformly over $j \leq M$,

$$\frac{1 - r^{-(M'-M)}}{1 - 1/r} \leq \sum_{k=j}^{M'-1} \nu(j,k) \leq \frac{1}{1 - 1/(r(1 - M'/n))}. \tag{29}$$

Thus, the above sum $\sim 1/(1 - 1/r) \sim 1/\delta$ uniformly over $j \leq M$ provided $r^{-(M'-M)} \to 0$ and $M'/n = o(\delta)$. The second point is clear, since $M'/n \leq \frac{1}{\sqrt{n}} = o(\delta)$. To check the first point, since $r^{-(M'-M)} = (1+\delta)^{-(M'-M)} \leq \mathrm{e}^{-\delta(M'-M)}$ it is enough that $\delta(M' - M) \to \infty$. Since $c = \sqrt{n}\delta \to \infty$ and $\delta = O(1)$, $\delta(M' - M) = c - 2\sqrt{c} + O(\delta) \to \infty$. To estimate the second sum on the right-hand side of (28), we note that since $V'(0) = \delta - o(\delta)$ and $x_\star = o(1)$, for large n, $V(x) - V(y) \geq \delta(x - y)/2$ if $0 \leq y < x \leq x_\star$. With $na = j \leq M$ and $M' \leq nb \leq X_\star - 1$,

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

$1-a, 1-b \geq 1-x_\star \to 1$ and $nb-na \geq M'-M \to \infty$, so $\log|E_n(a,b)| \to 0$ uniformly over a and b; thus, for large n and $j \leq M$, $M' \leq k \leq X_\star - 1$, $\nu(j,k) \leq 2\exp(-\delta(k-j)/2)$. Summing over k, the second term in (28) is $\leq 2e^{-\delta(M'-M)/2}/(1-e^{-\delta/2}) \sim 4e^{-\delta(M'-M)/2}/\delta = o(1/\delta)$. Combining the two estimates, it follows that the sum on the left-hand side of (28) $\sim 1/\delta$ uniformly over $j \leq M$.

Since $\nu(j-1,k) = \nu(j-1,j)\nu(j,k)$ and $1/r \leq \nu(j-1,j) \leq 1/(r(1-M/n))$ which $\to 1$ uniformly over $j \leq M$, using (26), (27), and (25) we find that, uniformly over $j \leq M$,

$$q_+^0(j) \sim rj(1-j/n)(1-\delta), \qquad q_-^0(j) \sim j(1+\delta). \tag{30}$$

Let $\tilde{X}$ denote the process with $\tilde{X}_0 = X_0$ and transition rates $\tilde{q}_-(j) = j$ and $\tilde{q}_+(j) = q_+^0(j)(j/q_-(j))$, which are the same as for X with respect to the measure P^0 except multiplied by the factor $j/q_-^0(j)$ at each non-zero j. Then X with respect to P^0 is obtained from $\tilde{X}$ as $X_t = \tilde{X}_{s(t)}$ for $t \leq \tau$, where $s(t)$ is the inverse of the function

$$t(s) = \int_0^s \frac{\tilde{X}_u}{q_-^0(\tilde{X}_u)}\,\mathrm{d}u.$$

From the estimate of (30) we have $\sup_{t\leq\tau_M} s(t)/t \sim 1+\delta \to 1$, so to obtain the desired result for X with respect to P^0 it is enough to show it for $\tilde{X}$. Since $M = o(\sqrt{n})$, $M/\delta n = o(1/\sqrt{n}\delta) = o(1)$. Using $r = 1+\delta$, uniformly over $j \leq M$ we have $\tilde{q}_+(j)/j \sim (1+\delta)(1-o(\delta))(1-\delta)/(1+\delta) = 1-\delta+o(\delta)$. Let γ, γ' be lower and upper bounds on $1-\tilde{q}_+(j)/j$, respectively, and construct linear birth and death processes Z,Z' with parameters $1-\gamma$ and $1-\gamma'$, and initial value X_0, so that $Z_t' \leq X_t \leq Z_t$ for $t \leq \tau_M$. Following the proof of Theorem 2, combining (9) and (10) we have $\mathbb{P}(\tau \neq \tau_M \mid X_0) \leq (1-\gamma)^{M-X_0}$, which $\to 0$ if $X_0 \leq \sqrt{n/c}$ since then $\gamma \sim \delta$ and $\delta(M-X_0) \geq \delta\sqrt{n/c} = \sqrt{c} \to \infty$. As in the proof of Theorem 2, it only remains to check that $(\gamma-\delta)X_0$, $(\gamma'-\delta)X_0 \to 0$. This follows easily from the fact that γ, $\gamma' \sim \delta$.

Case 2: $\delta \to \delta_\infty > 0$. Let $r_\infty = 1+\delta_\infty$. Since δX_0 converges, we may assume X_0 is constant. We follow the same approach as before, only with different M. So, let $M' = \lfloor\sqrt{n}\rfloor$ and let $M \to \infty$ slowly. Then $M'/n \to 0$, $M'-M \to \infty$, and, since $r \to r_\infty > 1$, $r^{-(M'-M)} \to 0$. Thus, from (29) the first sum on the right-hand side of (28) $\sim 1/(1-1/r) \to r_\infty/\delta_\infty$. To estimate the second sum, note that if $j \leq M$ and $M' \leq k \leq X_\star - 1$ then $\nu(j,k) = \nu(j,M')\nu(M',X_\star-1) \leq \nu(M,M') \leq (r(1-M'/n))^{-(M'-M)}$. We may assume that $M = o(\sqrt{n})$; then, for large n, the right-hand side above is $\leq e^{-\delta_\infty\sqrt{n}/2}$. Summing over at most n such terms, the second sum on the right-hand side of (28) is $\leq ne^{-\delta_\infty\sqrt{n}/2} = o(1)$ so, combining the two, the left-hand side of (28) $\to r_\infty/\delta_\infty$ uniformly for $j \leq M$.

Recall that $\nu(j-1,k) = \nu(j-1,j)\nu(j,k)$. Uniformly for $j \leq M$, $\nu(j-1,j) \to 1/r_\infty$ and $(1-j/n) \to 1$. In addition, $1-\delta_\infty/r_\infty = 1/r_\infty$. Thus, uniformly over $j \leq M$,

$$q_+^0(j) \sim r_\infty j(1-\delta_\infty/r_\infty) = j, \qquad q_-^0(j) \sim j(1+\delta_\infty) = r_\infty j.$$

Let $\tilde{X}$ be as before, which has $\tilde{q}_-(j) = j$ and $\tilde{q}_+(j) \sim j/r_\infty$. In this case, $X_t = \tilde{X}_{s(t)}$ with $\sup_{t\leq\tau_m} s(t)/t \sim r_\infty$. Thus, the result for X is obtained from the one for $\tilde{X}$ by changing time by the factor r_∞. For $\tilde{X}$, compare to the linear birth and death process Z with parameter $1/r_\infty$. Following the proof of Theorem 2 it suffices to show that, for fixed $t > 0$, $\mathbb{P}(\tau \neq \tau_M)$, $\mathbb{P}(\tilde{X}_s \neq Z_s$ for some $s \leq t \wedge \tau_M) = o(1)$. Since $\tilde{q}_+(j) \leq \tilde{q}_-(j)$, $\tilde{X}$ is a supermartingale, so the first statement is true provided $M \to \infty$. For $t < \tau_m$ the difference in rates between X and Z, when they take

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

the same value, is at most $\sup_{j\le M} |\tilde{q}_+(j) - j/r_\infty| = o(M)$, so is $o(1)$ if $M \to \infty$ slowly enough. Thus, for constant X_0 and fixed $t > 0$, $\mathbb{P}(X_s \neq Z_s \text{ for some } s \le t \wedge \tau_M \mid X_0) \le 1 - \mathrm{e}^{-o(1)t} = o(1)$, which shows the second statement and completes the proof. □

8.3. Metastability

In this section we prove the metastability results of Theorem 5. In particular, in this section we condition on $A_\star = \{X_{\tau_\star} = X_\star\}$, the event denoted $A_n^\star$ in Theorem 5. On $A_\star$, $\tau_\star$ is the time of the first visit to $X_\star$, and the time to extinction can be broken into three epochs. Define the time of the last visit to $X_\star$ as $\tau_\star^{\mathrm{o}} = \sup\{t\colon X_t = X_\star\}$, setting $\tau_\star^{\mathrm{o}} = -\infty$ if X never reaches $X_\star$. On $A_\star$, the time to extinction $\tau = \inf\{t\colon X_t = 0\}$ is the sum of the *approach time* $\tau_\star$, the *sojourn time* $\tau_\star^{\mathrm{o}} - \tau_\star$, and the *fall time* $\tau - \tau_\star^{\mathrm{o}}$. We proceed as follows:

1. Estimate the expected sojourn time $E_\star^{\mathrm{o}} := \mathbb{E}[\tau_\star^{\mathrm{o}} - \tau_\star \mid X_{\tau_\star} = X_\star]$.
2. Show that the expected approach time and fall time are $o(E_\star^{\mathrm{o}})$.
3. With a coupling argument, show that the rescaled sojourn time $(\tau_\star^{\mathrm{o}} - \tau_\star)/E_\star^{\mathrm{o}}$, conditioned on $X_{\tau_\star} = X_\star$, converges in distribution to exponential with mean 1.

Let us make the formal statements that will be the goal of this section.

Lemma 7. *(Expected sojourn time.)* $E_\star^{\mathrm{o}} \sim \sqrt{\dfrac{2\pi}{n}}\dfrac{r}{\delta^2}\exp\left(n(\log r + 1/r - 1)\right)$.

Lemma 8. *(Approach time.)* $\max_{j\in\{1,\dots,n\}} \mathbb{E}[\tau_\star \mid X_{\tau_\star} = X_\star,\ X_0 = j] = o(E_\star^{\mathrm{o}})$.

Lemma 9. *(Fall time.)* $\mathbb{E}[\tau - \tau_\star^{\mathrm{o}}] = o(E_\star^{\mathrm{o}})$.

Lemma 10. *(Exponential limit. For each $t \ge 0$,* $\mathbb{P}((\tau_\star^{\mathrm{o}} - \tau_\star)/E_\star^{\mathrm{o}} > t) \to \mathrm{e}^{-t}$.

Note that the expected sojourn time and fall time do not depend on X_0; we have emphasized the uniformity of the estimate with respect to X_0 only in Lemma 8. Before proving these results, we use them to prove the rest of Theorem 5.

Proof of metastability results of Theorem 5 The extinction time is the sum of the approach, sojourn, and fall times: $\tau = \tau_\star + (\tau_\star^{\mathrm{o}} - \tau_\star) + (\tau - \tau^{\star\mathrm{o}})$. Combining Lemmas 7, 8, and 9, $\mathbb{E}[\tau \mid X_{\tau_\star} = X_\star] \sim E_\star^{\mathrm{o}}$, which is the desired estimate of the expected time to extinction.

Using Lemmas 8 and 9 and Markov's inequality, conditioned on $X_{\tau_\star} = X_\star$, with probability $1 - o(1)$, $\tau/E_\star^{\mathrm{o}} = (\tau_\star^{\mathrm{o}} - \tau_\star)/E_\star^{\mathrm{o}} + o(1)$. The exponential limit for τ then follows from Lemma 10. □

We begin by deriving some formulas for the expected time to hit $j+1$ or $j-1$, starting from j.

8.3.1. *Crossing times* Let $T_+(j) = \inf\{t\colon X_t = j+1\}$ and $T_-(j) = \inf\{t\colon X_t = j-1\}$, and let $S_+(j) = \mathbb{E}[T_+(j) \mid X_0 = j]$ and $S_-(j) = \mathbb{E}[T_-(j) \mid X_0 = j]$. Define also the conditioned versions, $S_\pm^\star(j) = \mathbb{E}[T_\pm(j) \mid X_0 = j \text{ and } X_{\tau_\star} = X_\star]$ and $S_\pm^0(j) = \mathbb{E}[T_\pm(j) \mid X_0 = j \text{ and } X_{\tau_\star} = 0]$. We will need to estimate the following quantities: (i) $S_+^\star(j)$ for $j < X_\star$, (ii) $S_-(j)$ for $1 \le j \le n$, and (iii) $S_-^0(j)$ for $j < X_\star$.

Quantity (i): For $1 < j < X_\star$ a first step analysis gives

$$S_+^\star(j) = \frac{1}{q_+^\star(j) + q_-^\star(j)} + \frac{q_-^\star(j)}{q_+^\star(j) + q_-^\star(j)}\left(S_+^\star(j-1) + S_+^\star(j)\right),$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

and solving gives $q^\star_+(j)S^\star_+(j) = 1 + q^\star_-(j)S^\star_+(j-1)$. Following [15, Section 5.2] we let $\pi(i,i) = 1$ and $\pi(i,j) = \prod_{k=i}^{j-1} q^\star_+(k)/q^\star_-(k+1)$, and multiply through by $\pi(1,j)$ to obtain $q^\star_+(j)\pi(1,j)S^\star_+(j) = \pi(1,j) + q^\star_+(j-1)\pi(1,j-1)S^\star_+(j-1)$. Then, since $q^\star_+(1)S^+_1 = 1$, we solve to obtain

$$S^\star_+(j) = \frac{1}{q^\star_+(j)\pi(1,j)} \sum_{i=1}^{j} \pi(1,i) = \frac{1}{q^\star_+(j)} \sum_{i=1}^{j} \frac{1}{\pi(i,j)}.$$

Using (25), for $i < j$

$$\begin{aligned}\frac{1}{\pi(i,j)} &= \prod_{k=i}^{j-1} \frac{q_-(k+1)}{q_+(k)} \frac{h_+(k)/h_+(k+1)}{h_+(k+1)/h_+(k)} \\ &= \left(\frac{h_+(i)}{h_+(j)}\right)^2 \frac{q_-(j)}{q_+(i)}\, \nu(i,j-1).\end{aligned}$$

This gives

$$S^\star_+(j) = \frac{q_-(j)}{q^\star_+(j)} \sum_{i=1}^{j} \nu(i,j-1) \frac{h_+(i)^2}{h_+(j)^2} \frac{1}{q_+(i)}. \tag{31}$$

Quantity (ii): For $1 \le j < n$, a first step analysis gives the recursion $q_-(j)S_-(j) = 1 + q_+(j)S_-(j+1)$. Multiplying through by $1/\pi(j,n)$ gives

$$\frac{q_-(j)}{\pi(j,n)} S_-(j) = \frac{1}{\pi(j,n)} + \frac{q_-(j+1)}{\pi(j+1,n)} S_-(j+1);$$

since $S_-(n) = 1/q_-(n)$, the solution is

$$S_-(j) = \frac{\pi(j,n)}{q_-(j)} \sum_{i=j}^{n} \frac{1}{\pi(i,n)} = \frac{1}{q_-(j)} \sum_{i=j}^{n} \pi(j,i).$$

Writing in terms of ν we obtain

$$S_-(j) = \frac{q_+(j)}{q_-(j)} \sum_{i=j}^{n} \nu(i-1,j) \frac{1}{q_-(i)}. \tag{32}$$

Quantity (iii): For $1 \le j < X_\star$ a first step analysis gives $q^0_-(j)S^0_-(j) = 1 + q^0_+(j)S^0_-(j+1)$. In this case it is better to express the solution using the function ν^0, defined like ν except with $q^0_\pm$ in place of $q_\pm$. Since $S^0_-(X_\star - 1) = 1/q^0_-(X_\star - 1)$, following the same approach as above we find that

$$S^0_-(j) = \frac{q^0_+(j)}{q^0_-(j)} \sum_{i=j}^{X_\star - 1} \nu^0(i-1,j) \frac{1}{q^0_-(i)}. \tag{33}$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

8.3.2. *Sojourn time* Here we prove Lemma 7. We use the same basic approach as in [1]; the calculations become more delicate when $\delta \to 0$. Define recursively $\tau_\star(0) = \tau_\star$ and, for $0 < k < K = \min\{k > 0 \colon X_{\tau_\star(k)} = 0\}$, $\tau_\star(k+1) = \inf\{t > \tau_\star(k) \colon X_t \in \{0, X_\star\}$ and $X_s \neq X_\star$ for some $\tau_\star(k) < s < t\}$. For $k = 1, \dots, K$ let $\rho_k = \tau_\star(k) - \tau_\star(k-1)$. Then, conditioned on $X_{\tau_\star} = X_\star$,

$$\tau_\star^0 - \tau_\star = \sum_{k=1}^{K-1} \rho_k. \tag{34}$$

By the strong Markov property, K is geometric with success probability $p_\star = \mathbb{P}(X_{\tau_\star} = 0 \mid X_0 = X_\star)$ and, defining $\tau_\star^+ = \inf\{t \colon X_t = X_\star$ and $X_s \neq X_\star$ for some $0 < s < t\}$, $\mathbb{E}[\rho_k \mid K > k] = \mathbb{E}[\tau_\star^+ \mid X_0 = X_{\tau_\star^+} = X_\star]$. Let $L_\star$ denote this expectation. Applying Wald's equation to (34), $E_\star^0 = L_\star(1/p_\star - 1)$. Recall $V(x) = x(\log r - 1) - (1-x)\log(1-x)$ defined in Lemma 4, and $V_\star := V(x_\star) = \log r + 1/r - 1$. Lemma 7 follows immediately from the following estimates on $p_\star$ and $L_\star$.

Lemma 11. *$p_\star \to 0$ and $p_\star \sim \dfrac{\delta}{2\sqrt{r}} \exp(-nV_\star)$.*

Lemma 12. *$L_\star \sim \dfrac{1}{\sqrt{n}\,\delta} \sqrt{\dfrac{\pi r}{2}}$.*

First we prove Lemma 11.

Proof of Lemma 11. First we show that if $p_\star \sim (\delta/2\sqrt{r}) \exp(-nV_\star)$ then $p_\star \to 0$; then we establish the estimate. Since $r \geq 1$, $V_\star \geq 0$. If $\delta \to 0$ it follows that $p_\star \to 0$. If $\delta \to \delta_\infty > 0$ then $V_\star \to V_\infty = \log r_\infty + 1/r_\infty - 1 > 0$, so $\exp(-nV_\star) \to 0$ and again $p_\star \to 0$.

If the first jump of X is to $X_\star + 1$ then $X_{\tau_\star} = X_\star$. So, conditioning on the first jump,

$$p_\star = \frac{q_-(X_\star)}{q(X_\star)} h_-(X_\star - 1).$$

By the definition of $X_\star$, $q_+(X_\star) \sim q_-(X_\star)$, so if $X_0 = X_\star$ then its first jump is to $X_\star - 1$ with probability $1/2 + o(1)$. Thus, in the notation of (17) it is enough to show that $h_-(X_\star - 1) \sim \frac{\delta}{\sqrt{r}} \exp(-nV_\star)$. We have

$$h_-(X_\star - 1) = \frac{\nu(X_\star - 1)}{\sum_{k=0}^{X_\star - 1} \nu(k)},$$

and we begin by estimating the numerator using Lemma 4 with $a = 0$ and $b = (X_\star - 1)/n$. Notice that $|b - x_\star| \leq 2/n$ and that $x_\star = \delta/r = \delta/(1+\delta)$, so the assumption $\limsup_n \delta < \infty$ implies $\limsup_n b < 1$. Since $V'(x_\star) = 0$, $b < x_\star$, and $x \mapsto |V''(x)|$ is increasing, it follows that $|V(b) - V(x_\star)| \leq \frac{1}{2} V''(x_\star)(2/n)^2 = O(1/n^2)$. Since $nb \geq n(x_\star - 2/n) = n\delta/r - 2 \to \infty$ and $\limsup_n 1 - b > 0$, it follows that $n(1-b)^2(b-a) \to \infty$. Since $1/r = 1 - x_\star$, $1 - b = 1 - x_\star + O(1/n) = (1 - x_\star)(1 + o(1)) \sim 1/r$. Putting it together, $\nu(X_\star - 1) \sim \sqrt{r} \exp(-nV(x_\star))$. The denominator is estimated in the proof of Lemma 5; in both cases,

$$\sum_{k=0}^{X_\star - 1} \nu(k) \sim \frac{1}{1 - 1/r} = r/\delta$$

and the result follows. □

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

In order to estimate $L_\star$ we will need additional information about the function ν. First, extend the domain of ν by defining $\nu(k,j) = 1/\nu(j,k)$ for $0 \le j < k < n$. An equivalent, unifying definition is given by the formula

$$\nu(j,k) = \frac{\prod_{i=j+1}^{n} q_-(i)/q_+(i)}{\prod_{i=k+1}^{n} q_-(i)/q_+(i)}.$$

Say that $f(n,\lambda) \sim g(n,\lambda)$ uniformly over $\lambda \in A$ if $\lim_{n\to\infty} \sup_{\lambda\in A} \left|\log\left(\frac{f(n,\lambda)}{g(n,\lambda)}\right)\right| = 0$.

Lemma 13. *Uniformly over* $|\sigma| \le n^{1/8}$, $\nu(X_\star - \sigma\sqrt{n}, X_\star) \sim \exp(-\sigma^2 r/2)$. *Moreover, for* $1 \le \sigma \le n^{1/8}$,

$$\sum_{0\le j\le n-1\,:\,|j-X_\star|\ge\sigma\sqrt{n}} \nu(j,X_\star) \le (2+o(1))\exp(-\sigma^2 r/2)\frac{\sqrt{n}}{\sigma r}.$$

Proof. Since $r(1-x_\star) = 1$ and $x \mapsto \log(r(1-x))$ is differentiable at $x_\star$, it follows that, uniformly over j such that $|j - X_\star| \le n^{1/4}$, $\log(r(1-j/n)) = O(n^{-3/4})$. Summing at most $n^{1/4}$ terms, $\log \nu(j,X_\star) = O\big(\frac{1}{\sqrt{n}}\big)$. This proves the first statement restricted to $|\sigma| \le n^{-1/4}$.

If $n^{1/4} < |j - X_\star| \le n^{5/8}$, then since $\limsup_n x_\star < 1$, $1 - j/n \sim 1 - X_\star/n$. Since $|j - X_\star| > n^{1/4} \to \infty$, by Lemma 4, uniformly over such j, $E_n(j/n, X_\star/n) \to 1$ if $j < X_\star$ and $E_n(X_\star/n, j/n) \to 1$ if $j \ge X_\star$, so

$$\nu(j,X_\star) \sim \exp(-n(V(X_\star/n) - V(j/n)); \tag{35}$$

note that this expression is valid not only for $j < X_\star$ but also for $j \ge X_\star$ under the extended definition of ν. Since $\limsup_n 1 - x_\star > 0$, $V'''(x) = 1/(1-x)^2$ is bounded on $[0, x_\star + o(1))]$, and recall that $V'(x_\star) = 0$ and $V''(x_\star) = -r$. Thus, if $|\sigma| \le n^{1/8}$, using a second-order Taylor approximation we find that

$$V(x_\star + \sigma/\sqrt{n}) - V(x_\star) = -\sigma^2 r/2n + O(n^{-9/8}). \tag{36}$$

In particular, $V(X_\star/n) - V(x_\star) = O(n^{-9/8})$. Combining this with (35) and (36), the first statement is proved for the remaining values of $|\sigma|$, namely, $(n^{-1/4}, n^{1/8}]$.

Next, for $j < k < X_\star$, since $\log(r(1-x_\star)) = 0$ and $x \mapsto \log(r(1-x))$ is decreasing,

$$-\log \nu(j,k) = \sum_{i=j+1}^{k} \log(r(1-i/n)) \ge (k-j)\log(r(1-k/n))$$

and thus $\nu(j,k) \le (r(1-k/n))^{-(k-j)}$. Fix σ with $1 \le \sigma \le n^{1/8}$ and observe that, for any j, k, ℓ, $\nu(j,\ell) = \nu(j,k)\nu(k,\ell)$. Using this property with $k = X_\star - \sqrt{n}\sigma$ and $\ell = X_\star$ and then bounding the sum by a geometric series,

$$\sum_{j=0}^{X_\star - \sqrt{n}\sigma} \nu(j,X_\star) \le \frac{(1+o(1))\exp(-\sigma^2 r/2)}{1 - (r(1-(X_\star - \sqrt{n}\sigma)/n))^{-1}}.$$

Since $r(1 - X_\star/n) = r(1-x_\star) + O(1/n) = 1 + O(1/n)$, the denominator is

$$1 - (1 + O(1/n) + \sigma r/\sqrt{n})^{-1} = 1 - (1 - (1+o(1))\sigma r/\sqrt{n}) = (1+o(1))n^{-1/2}\sigma r,$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

so

$$\sum_{j=0}^{X_\star-\sqrt{n}\sigma-1} \nu(j, X_\star) \le (1+o(1))\exp(-\sigma^2 r/2)\frac{\sqrt{n}}{\sigma r}.$$

From the other end, if $X_\star < k < j$ then since $\log(r(1-i/n)) \le \log(r(1-x_\star)) = 0$ if $i > X_\star$, we have

$$\log \nu(j,k) = -\log \nu(k,j) = \sum_{i=k+1}^{j} \log(r(1-i/n)) \le (j-k)\log(r(1-k/n))$$

and thus $\nu(j,k) \le (r(1-k/n))^{j-k}$. By an analogous argument we find that

$$\sum_{j=X_\star+\sqrt{n}\sigma}^{n-1} \nu(j, X_\star) \le \frac{(1+o(1))\exp(-\sigma^2 r/2)}{1-r(1-(X_\star+\sqrt{n}\sigma)/n)} \le (1+o(1))\exp(-\sigma^2 r/2)\frac{\sqrt{n}}{\sigma r}.$$

Combining the two estimates completes the proof. □

We are now ready to prove Lemma 12.

Proof of Lemma 12. Let $S^\star_+(j)$, $S_-(j)$ be as in Section 8.3.1. Conditioning on the first step,

$$L_\star = \frac{1}{q_+(X_\star)+q_-(X_\star)}\left(1+q_-(X_\star)S^\star_+(X_\star-1)+q_+(X_\star)S_-(X_\star+1)\right).$$

We have $q_-(X_\star) = X_\star$ and, since $r(1-X_\star/n) = 1+O(1/n)$, $q_+(X_\star) \sim X_\star$, so $L_\star \sim \frac{1}{2}\left(1/X_\star + S^\star_+(X_\star-1) + S_-(X_\star+1)\right)$. Since $h_+(X_\star) = 1$ and $h_+(X_\star-1) \to 1$, $q^\star_+(X_\star-1) \sim q_+(X_\star-1)$. By the definition of $X_\star$, $q_+(X_\star-1) \sim q_-(X_\star-1)$. In addition, by Lemma 13, $\nu(i, X_\star) = \nu(i, X_\star-2)\nu(X_\star-2, X_\star) \sim \nu(i, X_\star-2)$ uniformly over i. Thus, using (31) with $j = X_\star - 1$,

$$S^\star_+(X_\star-1) \sim \sum_{i=1}^{X_\star-1} \nu(i, X_\star)\frac{h_+(i)^2}{q_+(i)}. \tag{37}$$

We will estimate the bulk of the sum in (37), finding that it tends to a, then show that the rest of the sum, as well as $S_-(X_\star+1)$, are negligible in comparison. Recall the notation $c = \sqrt{n}\delta$, noting that $c \to \infty$ by assumption. We then have $X_\star \sim n\delta = \sqrt{n}c$. Since $r(1-x_\star) = 1$, $q_+(i) = ri(1-i/n) = i(1+r(x_\star - i/n))$, so $q_+(i) \sim X_\star$ uniformly over $|i - X_\star| \le C$ provided $C = o(X_\star)$, or equivalently $C = o(\sqrt{n}c)$. By Lemma 5, $h_+(i) \to 1$ uniformly over $i \ge M$ provided $\delta M \to \infty$. If $M = X_\star - C$ with $C = o(\sqrt{n}c)$ then $\delta M \sim \delta X_\star \sim c^2/r \to \infty$. Thus, if we define $\Sigma = \sqrt{c} \wedge n^{1/8}$, then since $\sqrt{nc} = o(\sqrt{n}c)$,

$$\sum_{i=\lfloor X_\star - \Sigma\sqrt{n}\rfloor+1}^{X_\star-1} \nu(i, X_\star)\frac{h_+(i)^2}{q_+(i)} \sim \frac{1}{X_\star}\sum_{i=\lfloor X_\star-\Sigma\sqrt{n}\rfloor+1}^{X_\star-1} \nu(i, X_\star).$$

Using Lemma 13 and the fact that $\Sigma \to \infty$ and $\Sigma \le n^{1/8}$,

$$\sum_{i=\lfloor X_\star-\Sigma\sqrt{n}\rfloor+1}^{X_\star-1} \nu(i, X_\star) \sim \sqrt{n}\int_{-\infty}^{0} \mathrm{e}^{-\sigma^2 r/2}\,\mathrm{d}\sigma = \sqrt{\frac{n\pi}{2r}}.$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Assuming the rest of the sum is negligible in comparison, since $X_\star \sim n\delta/r$ we then have

$$S_+^\star(X_\star - 1) \sim \frac{\sqrt{n}}{X_\star}\frac{\pi}{2r} = \frac{1}{\sqrt{n}\delta}\sqrt{\frac{\pi r}{2}}. \tag{38}$$

So, we now show the rest of the sum in the brackets in (37) is $o(\sqrt{n})$; we may of course ignore the 1. Using the trivial estimate $q_+(i) \geq 1$ and $h_+(i) \leq 1$ for $i \geq 1$, as well as $X_\star \leq n\delta$ and Lemma 13 with $\sigma = n^{1/8}$, if $\sqrt{c} \geq n^{1/8}$ then the rest of the sum is

$$\sum_{i=1}^{\lfloor X_\star - n^{5/8}\rfloor} \nu(i, X_\star - 2)\frac{h_+(i)^2}{q_+(i)}X_\star \leq (2 + o(1))\mathrm{e}^{-n^{1/4}r/2}\frac{n^{3/8}}{r}\frac{n\delta}{r} = O(1) = o(\sqrt{n}).$$

If $\sqrt{c} < n^{1/8}$ we treat the remainder in two parts, beginning with $i \leq 1/\delta = \sqrt{n}/c$. X is dominated by the linear birth and death process Z with $Z_0 = X_0$ and parameter r, and using (5), since 0 is absorbing for Z, $\mathbb{P}(Z_t = 0 \text{ for some } t > 0 \mid Z_0) = \lim_{t\to\infty} \rho(t)^{Z_0} = \min(1, 1/r)^{Z_0} = r^{-Z_0}$ since $r > 1$ by assumption in this section. If $Z_t = i$ for $i > 0$, then with probability $p_i > 0$, Z hits 0 before it returns to i. Thus, Z visits any $i > 0$ at most a geometric(p_i) number of times, which is almost surely finite and implies that a.s., $\lim_{t\to\infty} Z_t \in \{0, \infty\}$. Therefore, $\mathbb{P}(\lim_{t\to\infty} Z_t = \infty \mid Z_0) = 1 - r^{-Z_0} = 1 - (1+\delta)^{-Z_0} \leq \delta Z_0$. The last inequality follows from $(1+\delta)^{-Z_0} \geq e^{-\delta Z_0} \geq 1 - \delta Z_0$, where we used the estimate $1 + u \leq \mathrm{e}^u$ for $u \in \mathbb{R}$ which follows from the convexity of $u \mapsto \mathrm{e}^u$. Since Z dominates X,

$$h_+(i) \leq \mathbb{P}(\lim_{t\to\infty} Z_t = \infty \mid Z_0 = i) \leq \delta i. \tag{39}$$

Since $h_+(i) \leq 1$, $h_+(i)^2 \leq h_+(i)$, and since $X_\star \leq n\delta$ and $q_+(i) \geq i$ for $i < X_\star$, $h_+(i)^2 X_\star/q_+(i) \leq \delta^2 n = c^2$. Using Lemma 13 with $\sigma = (X_\star - \lfloor\sqrt{n}/c\rfloor)/\sqrt{n} \sim c - 1/c = c - o(1)$,

$$\sum_{i=1}^{\lfloor\sqrt{n}/c\rfloor} \nu(i, X_\star)\frac{h_+(i)^2}{q_+(i)}X_\star \leq (2 + o(1))\mathrm{e}^{-\sigma^2 r/2}\frac{\sqrt{n}}{\sigma r}c^2 \sim 2\mathrm{e}^{-(c-o(1))^2 r/2}\frac{c}{r}\sqrt{n} = o(\sqrt{n}),$$

since $c \to \infty$ and thus $(2c/r)\mathrm{e}^{-(c-o(1))^2 r/2} \to 0$ as $n \to \infty$. Using again the trivial estimate $h_+(i) \leq 1$ and $q_+(i) \geq i$, and also $\nu(i, X_\star) \leq \nu(X_\star - \sqrt{nc}, X_\star) \sim \mathrm{e}^{-cr/2}$ for $i \leq X_\star - \sqrt{nc}$,

$$\begin{aligned}\sum_{i=\lfloor\sqrt{n}/c\rfloor+1}^{\lfloor X_\star - \sqrt{nc}\rfloor} \nu(i, X_\star)\frac{h_+(i)^2}{q_+(i)}X_\star &\leq (1 + o(1))\mathrm{e}^{-cr/2}\sqrt{nc}\sum_{i=\lfloor\sqrt{n}/c\rfloor+1}^{\lfloor X_\star - \sqrt{nc}\rfloor}\frac{1}{i},\\ &\sim \sqrt{nc}\mathrm{e}^{-cr/2}(\log(\sqrt{nc}) - \log(\sqrt{n}/c))\\ &\sim \sqrt{n}(2c\log(c)\mathrm{e}^{-cr/2}) = o(\sqrt{n}).\end{aligned}$$

This completes the estimation of the sum from (37).

For $S_-(X_\star + 1)$, using (32) with $j = X_\star + 1$ and simplifying as before,

$$S_-(X_\star + 1) \sim \sum_{i=X_\star+1}^{n} \nu(i-1, X_\star)\frac{1}{q_-(i)}.$$

Since this case is similar to the one before, we just give an outline. Breaking up the sum in the same way, the bulk of the sum is estimated in the same way as before and gives the same result.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

To bound the remainder, it suffices to note that $q_-(i) = i \geq X_\star$ for $i \geq X_\star$, then use Lemma 13 directly, noting that $2\mathrm{e}^{-\Sigma^2 r/2}/(\Sigma r) \to 0$. Since $L_\star$ scales like the average of the two values, the result follows from (38). □

8.3.3. *Approach time* We now prove Lemma 8. For this and for Lemma 8, we first derive a more concrete lower bound on the expected sojourn time $E^{\mathrm{o}}_\star$, using the formula $E^{\mathrm{o}}_\star \sim \sqrt{2\pi/n}(r/\delta^2)\exp(nV_\star)$ of Lemma 7, where $V_\star = \log r + 1/r - 1$.

If $\delta \to 0$ then $V_\star = \log(1+\delta) + 1/(1+\delta) - 1 = \delta - \delta^2/2 + 1 - \delta + \delta^2 - 1 + O(\delta^3) = \delta^2/2 + O(\delta^3) \sim \delta^2/2$ and, since $c = \sqrt{n}\delta$ and $r \to 1$, $\sqrt{2\pi/n}(r/\delta^2) = \sqrt{2\pi}(r/c\delta)$ and $nV_\star \sim c^2/2$. If $\delta \to \delta_\infty > 0$ then $V_\star \to V_\infty = \log r_\infty + 1/r_\infty - 1 > 0$. Thus in either case,

$$E^{\mathrm{o}}_\star \geq \begin{cases} (1-o(1))\dfrac{\sqrt{2\pi}}{\delta}\,\dfrac{1}{c}\exp((1-o(1))c^2/2) & \text{if } \delta \to 0, \\ \sqrt{2\pi/n}(r_\infty/\delta_\infty^2)\mathrm{e}^{(1-o(1))V_\infty n} & \text{if } \delta \to \delta_\infty > 0. \end{cases} \tag{40}$$

Using (40), the following result implies Lemma 8, since $c^2 \log c = o((1/c)\exp((1-o(1))c^2/2))$ in the case $\delta \to 0$ and $n\log n = o(\exp((1-o(1))V_\infty n))$ in the case $\delta \to \delta_\infty > 0$.

Lemma 14.

$$\max_{j\in\{1,\dots,n\}} \mathbb{E}[\tau_\star \mid X_{\tau_\star} = X_\star,\ X_0 = j] = \begin{cases} O((1/\delta)c^2\log(c)) & \text{if } \delta \to 0 \\ O(n\log(n)) & \text{if } \delta \to \delta_\infty > 0. \end{cases}$$

Proof. Using the natural coupling, by Lemma 1 it is enough to consider the initial values $X_0 = 1$ and $X_0 = n$; we begin with $X_0 = n$. We break up the travel time to $X_\star$ into three checkpoints: $2nx_\star$, $nx_\star + \sqrt{n}$, and $X_\star$.

First checkpoint: If $x > x_\star$ then $\mu(x - x_\star) = \mu(x) = x(r(1-x)-1) = rx(x_\star - x) \leq -r(x - x_\star)^2 \leq -(x-x_\star)^2$. The differential equation $y' = -y^2$ has solution flow $\phi(t, y) = 1/(1/y + t)$, so letting $\tau_1 = \inf\{t\colon x_t \leq 2x_\star\}$ and defining the continued process $\tilde{x}$ by $\tilde{x}_t = \phi(t - t\wedge\tau_1, x_{t\wedge\tau_1})$ we have $\mu_t(\tilde{x}) \leq -\tilde{x}_t^2$ for all $t \geq 0$. Taking expectations and using Jensen's inequality,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\tilde{x}_t] \leq -\mathbb{E}[\tilde{x}_t^2] \leq -(\mathbb{E}[\tilde{x}_t])^2,$$

which gives $\mathbb{E}[\tilde{x}_t] \leq \phi(t, x_0) = 1/(1/x_0 + t)$. Since $x_0 \leq 1$, Markov's inequality then gives

$$\mathbb{P}(\tau_1 > t) \leq \mathbb{P}(\tilde{x}_t > 2x_\star) \leq \frac{1}{2x_\star}\frac{1}{1+t} = \frac{r}{2\delta(1+t)}.$$

Letting $t = 1/\delta$, the above is at most $1/2$. Using the Markov property and iterating, $\mathbb{P}(\tau_1 > k/\delta) \leq 2^{-k}$, so $\mathbb{E}[\tau_1] \leq (1/\delta)\sum_{k\geq 0}\mathbb{P}(\tau_1 > k/\delta) \leq 2/\delta$.

Second checkpoint: To get from $2x_\star$ to $x_\star + \frac{1}{\sqrt{n}}$, let $\tau_2 = \inf\left\{t\colon x_t \leq x_\star + \frac{1}{\sqrt{n}}\right\}$ and note that if $x > x_\star$ then $\mu(x - x_\star) = -rx(x - x_\star) \leq -rx_\star(x - x_\star) = -\delta(x - x_\star)$. Thus, $\xi_t = \mathrm{e}^{\delta(t\wedge\tau_2)}(x_{t\wedge\tau_2} - x_\star)$ is a supermartingale and if $x_0 \leq 2x_\star$ then $\xi_0 \leq x_\star$ and $\mathbb{P}(\tau_2 > t \mid x_0 \leq 2x_\star) \leq \mathbb{P}(\xi_t \geq \mathrm{e}^{\delta t}/\sqrt{n}) \leq \mathrm{e}^{-\delta}\sqrt{n}x_\star \leq \mathrm{e}^{-\delta t}\sqrt{n}\delta = \mathrm{e}^{-\delta t}c$. Thus, $\mathbb{E}[\tau_2 \mid x_0 \leq 2x_\star] = \int_0^\infty \mathbb{P}(\tau_2 > t \mid x_0 \leq 2x_\star)\,\mathrm{d}t \leq c/\delta$.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Third checkpoint: Finally we estimate $\tau_3 = \inf\{t\colon X_t = X_\star\}$, assuming $X_0 \le nx_\star + \sqrt{n}$. With $S_-(j)$ as in (32),

$$\mathbb{E}[\tau_3 \mid X_0 \le nx_\star + \sqrt{n}] \le \sum_{j=X_\star+1}^{X_\star+\sqrt{n}+1} S_-(j). \tag{41}$$

Since $h_+(i) = h_+(j) = 1$, $q_+(j) \le q_-(j) = q_-^\star(j)$, and $q_-(i) = i \ge X_\star$ for $i, j > X_\star$,

$$S_-(j) \le \frac{1}{X_\star} \sum_{i=j}^{n} \nu(i-1, j).$$

If $X_\star < j \le X_\star + \sqrt{n} + 1$ then, using Lemma 13 with $\sigma \le 1 + 1/\sqrt{n}$,

$$\nu(i-1, j) = \frac{\nu(i-1, X_\star)}{\nu(j, X_\star)} \le (1 + o(1))\mathrm{e}^{r/2}\nu(i-1, X_\star).$$

Using Lemma 13 again and approximating the sum by a Gaussian integral, we obtain

$$S_-(j) \le \frac{1}{X_\star}(1 + o(1))e^{r/2}\sqrt{\frac{n\pi}{2r}} = (1 + o(1))\sqrt{\frac{\pi}{2r}}\mathrm{e}^{r/2}\frac{\sqrt{n}}{n\delta} = O(1/\sqrt{n}\delta),$$

since r is bounded by assumption. Summing over $\sqrt{n} + 1$ terms of the same size and using (41), we find $\mathbb{E}[\tau_3 \mid X_0 \le nx_\star + \sqrt{n}] = O(\sqrt{n}/\sqrt{n}\delta) = O(1/\delta)$. In all three cases the expected travel time is $O(c/\delta)$, which satisfies the stated estimates; for the case $\delta \to \delta_\infty > 0$, note that $c/\delta = \sqrt{n}$.

Next, we consider the case $X_0 = 1$. From (31), and since $q_-(j) \le q_+(j) \le q_+^\star(j)$ and $q_+(i) \ge i$ for $i, j < X_\star$,

$$S_+^\star(j) \le \sum_{i=1}^{j} \nu(i, j-1)\frac{h_+(i)^2}{h_+(j)^2}\,\frac{1}{i}.$$

Let s_{ij} denote the above summands. Then,

$$\mathbb{E}[\tau_\star \mid X_0 = 1,\ X_{\tau_\star} = X_\star] = \sum_{j=1}^{X_\star-1} S_+^\star(j) \le \sum_{j=1}^{X_\star-1}\sum_{i=1}^{j} s_{ij}.$$

If $i \le j < X_\star$ then $\nu(i, j-1) \le r$ and $h_+(i)/h_+(j) \le 1$, which we use below. In order to obtain good enough estimates, we need to be a bit more precise. We treat the cases $\delta \to 0$ and $\delta \to \delta_\infty > 0$ separately.

Case 1: $\delta \to 0$. Let $c = \sqrt{n}\delta$, so $c \to \infty$ and $c = o(\sqrt{n})$. We treat the sum in three parts: (i) $1 \le i \le j \le 1/\delta$, (ii) $1 \le i \le 1/\delta < j \le X_\star - 1$, and (iii) $1/\delta < i \le j < X_\star$.

Part (i): From (39), $h_+(i) \le \delta i$, and since the denominator $\sim 1/(1 - 1/r)$,

$$h_+(j) = \frac{\sum_{k=0}^{X_0-1}\nu(k)}{\sum_{k=0}^{X_\star-1}\nu(k)} \ge \frac{(1 - r^{-j})/(1 - 1/r)}{(1 + o(1)/(1 - 1/r)} = (1 - o(1))(1 - r^{-j}).$$

Since $\delta \to 0$, $(1 + \delta)^{-j} = ((1 + \delta)^{1/\delta})^{-\delta j} \to \mathrm{e}^{-\delta j}$ uniformly over $\delta j \le 1$. Since $\mathrm{e}^{-x} \le 1 - (1 - 1/\mathrm{e})x$ for $x \in [0, 1]$, if $\delta j \le 1$ then

$$h_+(j) \ge (1 - o(1))(1 - \mathrm{e}^{-\delta j}) \ge (1 - o(1))(1 - 1/\mathrm{e})\delta j, \tag{42}$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

which is at least $\delta j/2$ for large n, since $1-1/\mathrm{e}>1/2$. Thus, $s_{ij}\le r(i^2/(j/2)^2)(1/i)=4ri/j^2$, so

$$\sum_{j=1}^{\lfloor 1/\delta\rfloor}\sum_{i=1}^{j}s_{ij}\le\sum_{j=1}^{\lfloor 1/\delta\rfloor}\frac{4r}{j^2}\sum_{i=1}^{j}i\le\sum_{j=1}^{\lfloor 1/\delta\rfloor}2r\le 2r/\delta.$$

Part (ii): Since $q_-(k)/q_+(k)\ge 1/r$, $\nu(0,i)\ge r^{-i}=(1+\delta)^{-i}\ge \mathrm{e}^{-\delta i}$ for each i. Since $q_-(k)\le q_+(k)$ for $k<X_\star$ and $i\le\lfloor 1/\delta\rfloor$, $\nu(i,j-1)\le\nu(\lfloor 1/\delta\rfloor,j-1)$, so

$$\nu(i,j-1)=\frac{\nu(0,j-1)}{\nu(0,i)}\le \mathrm{e}^{\delta i}\nu(0,j-1).$$

Thus, if $i\le 1/\delta<j<X_\star$ then

$$\nu(i,j-1)=\frac{\nu(0,j-1)}{\nu(0,i)}\le \mathrm{e}^{1}\nu(0,j-1).$$

Since $j>\lfloor 1/\delta\rfloor$ and $j\mapsto h_+(j)$ is non-decreasing, using (42), $h_+(j)\ge(1-o(1))(1-1/\mathrm{e})\delta\lfloor 1/\delta\rfloor$ is at least $1/2$ for large n, since $\delta\to 0$ implies $\delta\lfloor 1/\delta\rfloor\to 1$. Again using $h_+(i)\le\delta i$ and combining,

$$\sum_{j=\lfloor 1/\delta\rfloor+1}^{X_\star-1}\sum_{i=1}^{\lfloor 1/\delta\rfloor}s_{ij}\le\sum_{j=\lfloor 1/\delta\rfloor+1}^{X_\star-1}\mathrm{e}^{1}\nu(0,j-1)\sum_{i=1}^{\lfloor 1/\delta\rfloor}\frac{(\delta i)^2}{1/4}\frac{1}{i}.$$

We easily estimate

$$\sum_{i=1}^{\lfloor 1/\delta\rfloor}\frac{(\delta i)^2}{1/4}\frac{1}{i}\le 4\delta^2\sum_{i=1}^{\lfloor 1/\delta\rfloor}i\le 4\delta^2\frac{(1/\delta)^2}{2}=2.$$

Using (22),

$$\sum_{j=\lfloor 1/\delta\rfloor+1}^{X_\star-1}\mathrm{e}^{1}\nu(0,j-1)\le(1+o(1))/\delta.$$

Combining the two, the sum is at most $(2+o(1))/\delta$.

Part (iii): This part is the easiest; we simply use $\nu(i,j-1)\le r$, $h_+(i)/h_+(j)\le 1$, and $1/q_+(i)\le 1/i$ and, treating the sum as a right-endpoint Riemann sum,

$$\sum_{j=\lfloor 1/\delta\rfloor+1}^{X_\star-1}\sum_{i=\lfloor 1/\delta\rfloor+1}^{j}\frac{1}{i}\le\sum_{j=\lfloor 1/\delta\rfloor+1}^{X_\star-1}r(\log(j)-\log(1/\delta)).$$

We can combine the logs as $\log(\delta j)$, which is increasing in j. Treating the sum as a left-endpoint Riemann sum of the function $\log(x)$ with interval widths δ, and noting that $\delta X_\star\le n\delta^2=c^2$, the sum is at most $\frac{1}{\delta}(\delta X_\star(\log(\delta X_\star)-1)-\delta(1/\delta)(\log(\delta(1/\delta))-1)\le\frac{1}{\delta}(c^2(\log(c^2)-1)+1)$. Combining all three parts, we find $\mathbb{E}[\tau_\star\mid X_0=1,\ X_{\tau_\star}=X_\star]\le\frac{1}{\delta}(c^2\log(c^2)-c^2+O(1))\le\frac{1}{\delta}2c^2\log(c)$ for large n, since $c\to\infty$.

Case 2: $\delta\to\delta_\infty>0$. Since $1/\delta=O(1)$ in this case, the whole sum can be treated as in part (iii) above. Since $s_{ij}\le r/i$,

$$\sum_{j=1}^{X_\star-1}\sum_{i=1}^{j}s_{ij}\le\sum_{j=1}^{X_\star-1}(1+\log j).$$

Treating the sum as a left-endpoint Riemann sum, it is at most $X_\star\log(X_\star)-1\log(1)\le n\delta\log(n\delta)=O(n\log(n))$. □

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

8.3.4. *Fall time* Here we show that the time to hit zero after the last visit to $X_\star$ is small compared to the sojourn time. The following is an equivalent formulation of Lemma 9.

Lemma 15. $\mathbb{E}[\tau \mid X_0 = X_\star, X_{\tau_\star} = 0] = o(E^{\mathrm{o}}_\star)$.

Proof. Let $L^0_\star$ denote the above expectation. With $S^0_-(j)$ as in (33),

$$L^0_\star = \sum_{j=1}^{X_\star} S^0_-(j).$$

Since we condition on $X_{\tau_\star} = 0$, the initial jump off $X_\star$ is to $X_\star - 1$ with rate $q_+(X_\star) + q_-(X_\star)$, which we denote $q^0_-(X_\star)$, after which we use the rates $q^0_\pm$ given by (25). Thus, $S_{X_\star} = 1/q^0_-(X_\star)$. For $j \le X_\star - 1$, $q^0_+(j) \le q_+(j)$ and $q^0_-(j) \ge q_-(j)$, so $q^0_+(j)/q^0_-(j) \le q_+(j)/q_-(j) \le r$. Moreover, $q^0_-(i) \ge q_-(i) = i$, so

$$S^0_-(j) \le r \sum_{i=j}^{X_\star - 1} \nu^0(i-1, j)\, \frac{1}{i}.$$

Let s_{ij} denote the summands. Summing over j and exchanging the order of summation,

$$L^0_\star = \frac{1}{q^0_-(X_\star)} + r \sum_{i=1}^{X_\star - 1} \sum_{j=1}^{i} s_{ij}.$$

The first term is at most $1/X_\star$, which is clearly $o(E^{\mathrm{o}}_\star)$. To estimate the sum we need more information about ν^0, so we first estimate the ratios

$$\frac{q^0_+(j)}{q^0_-(j)} = \frac{q_+(j)}{q_-(j)}\, \frac{h_-(j+1)/h_-(j)}{h_-(j-1)/h_-(j)} \tag{43}$$

of the conditioned rates given by (25). To do so we use the formulas (26) and (27). Since $\nu(j-1, k) \le \nu(j, k)$ for $j < X_\star$, from (27),

$$\frac{h_-(j-1)}{h_-(j)} \ge 1 + \frac{1}{\sum_{k=j}^{X_\star - 1} \nu(j, k)}, \tag{44}$$

which simplifies some calculations. Define $\sigma_j = (X_\star - j)/\sqrt{n}$ and similarly for σ_k, and let $\Sigma = n^{1/8} \wedge \sqrt{c}$.

Estimation for $\sigma_j \le \Sigma$: By Lemma 13, uniformly over $0 \le \sigma_k \le \sigma_j \le n^{1/8}$, $\nu(j, k) = \nu(j, X_\star)/\nu(k, X_\star) \sim \mathrm{e}^{(-\sigma_j^2 + \sigma_k^2)r/2}$, and so

$$\sum_{k=j}^{X_\star - 1} \nu(j, k) \sim \sqrt{n} \int_0^{\sigma_j} \mathrm{e}^{-(\sigma_j - \sigma_k)(\sigma_j + \sigma_k)r/2}\, \mathrm{d}\sigma_k.$$

Changing variables to $u = \sigma_j - \sigma_k$, $\sigma_j + \sigma_k = 2\sigma_j - u$ and the integral becomes

$$\int_0^{\sigma_j} \mathrm{e}^{-u(2\sigma_j - u)r/2}\, \mathrm{d}u \le \int_0^{\sigma_j} \mathrm{e}^{-u\sigma_j r/2}\, \mathrm{d}u \le \frac{2}{\sigma_j r}.$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Letting $b_j = (1-\epsilon_n)\sigma_j r/\sqrt{n}$ with $\epsilon_n \to 0$ sufficiently slowly, and using (26) and (44), it follows that

$$\frac{h_-(j-1)}{h_-(j)} \geq 1 + b_j/2, \qquad \frac{h_-(j+1)}{h_-(j)} \leq 1 - b_j/2.$$

Since $\sigma_j \leq n^{1/8}$, $b_j = o(1)$ so $h_-(j+1)/h_-(j-1) \leq 1 - b_j + o(b_j)$. On the other hand,

$$\begin{aligned}\frac{q_+(j)}{q_-(j)} &= r(1-j/n) = (1 + r(x_\star - j/n)) \\ &= 1 + r\sigma_j/\sqrt{n} + O(1/n) = 1 + (1+o(1))b_j + O(1/n). \end{aligned} \tag{45}$$

Using (43) and the above estimates,

$$\frac{q_+^0(j)}{q_-^0(j)} \leq (1 + (1+o(1))b_j + O(1/n))(1 - b_j + O(b_j^2)) = 1 - b_j^2 + o(b_j^2) + O(1/n). \tag{46}$$

Estimation for $\sigma_j \geq \Sigma$: Recall the upper bound from Lemma 4,

$$\nu(j,k) \leq \exp\big(-n(V((k+1)/n) - V((j+1)/n)\big).$$

Using the fact that V is non-decreasing and V' is non-increasing on $[0, x_\star]$, $n(V((k+1)/n) - V((j+1)/n)) \geq ((k-j) \wedge \sqrt{n})V'((j+1+\sqrt{n})/n)$. With this bound,

$$\sum_{k=j}^{X_\star - 1} \nu(j,k) \leq \frac{1}{1 - \mathrm{e}^{-V'((j+1+\sqrt{n})/n)}} + (X_\star - j)\mathrm{e}^{-\sqrt{n}V'(j+1+\sqrt{n}/n)}, \tag{47}$$

the first at most $\sqrt{n}$ terms forming a partial geometric series, and the last at most $X_\star - j$ terms each contributing at most a constant. Since $V'(x) = \log(r(1-x))$ and $r(1-x) = 1 + r(x_\star - x)$,

$$\begin{aligned}\mathrm{e}^{V'((j+1+\sqrt{n})/n)} &= 1 + r(x_\star - (j+1+\sqrt{n})/n) \\ &= 1 + r\big((X_\star - j)/n - \frac{1}{\sqrt{n}} + O(1/n)\big) \\ &\geq 1 + r(\sigma_j - 2)/\sqrt{n}\end{aligned}$$

for large n. This easily gives the bound $\sqrt{n}/(r(\sigma_j - 2))$ on the first term on the right-hand side of (47). To bound the second term, note that $\sigma_j \leq \sqrt{n}$ and $r \geq 1$, and that $1 + x \geq \mathrm{e}^{x/2}$ for $x \in [0,1]$, so $1 + r(\sigma_j - 2)/\sqrt{n} \geq \mathrm{e}^{(\sigma_j - 2)/2\sqrt{n}}$. Using this on the second term on the right-hand side of (47) and combining the two estimates, for large n,

$$\sum_{k=j}^{X_\star - 1} \nu(j,k) \leq \frac{\sqrt{n}}{r(\sigma_j - 2)} + \sqrt{n}\sigma_j \mathrm{e}^{-(\sigma_j - 2)/2}.$$

Since $x\mathrm{e}^{-x/2} \to 0$ faster than $\frac{1}{x}$ as $x \to \infty$, using $b_j = (1-\epsilon_n)r\sigma_j/\sqrt{n}$ with $\epsilon_n \to 0$ slowly enough, since $\Sigma \to \infty$ it follows that, uniformly over $\sigma_j \geq \Sigma$, $\sum_{k=j}^{X_\star - 1} \nu(j,k) \leq 1/b_j$, and so

$$\frac{h_-(j+1)}{h_-(j-1)} \leq \frac{1 - b_j}{1 + b_j}.$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Since $\sigma_j \geq \Sigma \to \infty$, $b_j = \omega\left(\frac{1}{n}\right)$, and using (45), $q_+(j)/q_-(j) = 1 + (1+o(1))b_j$. Using (43), uniformly over $j \leq X_\star - \Sigma\sqrt{n}$,

$$\frac{q_+^0(j)}{q_-^0(j)} \leq (1+(1+o(1))b_j)\frac{1-b_j}{1+b_j} = 1 - b_j + o(b_j). \tag{48}$$

Case 1: $\delta \to 0$. Since $\nu^0(i, i-1) \to 1$ uniformly over i, we can work with $\nu^0(i,j)$ instead of $\nu^0(i-1,j)$. Again, we break the sum into parts; the decomposition is similar to the one in the second half of the proof of Lemma 14, except that the third part has been further subdivided into three parts, for a total of five: (i) $1 \leq j \leq i \leq 1/\delta$, (ii) $1 \leq j \leq 1/\delta < i \leq X_\star - \Sigma\sqrt{n}$, (iii) $1/\delta < j \leq i \leq X_\star - \Sigma\sqrt{n}$, (iv) $1/\delta < j \leq X_\star - \Sigma\sqrt{n} < i \leq X_\star - 1$, and (v) $X_\star - \Sigma\sqrt{n} < j \leq i \leq X_\star - 1$. Note that $\sigma_i \geq \Sigma$ in parts (i)–(iii), and $\sigma_j \geq \Sigma$ in parts (i)–(iv).

Part (i): Note that if $j = o(X_\star)$, which is the case if $j \leq 1/\delta$, then $\sigma_j \sim \sqrt{n x_\star}$ and $b_j \sim \delta$, so $\nu^0(i,j) \leq (1-\delta+o(\delta))^{i-j} \leq \mathrm{e}^{-(1+o(1))\delta(i-j)}$, and treating it as a partial geometric sum,

$$\sum_{j=1}^{i} \nu^0(i,j) \leq \frac{1-\mathrm{e}^{-(1+o(1))\delta i}}{1-\mathrm{e}^{-(1+o(1))\delta}} \leq \frac{(1+o(1))\delta i}{\delta} = (1+o(1))i.$$

Thus,

$$\sum_{i=1}^{\lfloor 1/\delta \rfloor} \sum_{j=1}^{i} s_{ij} \leq \sum_{i=1}^{\lfloor 1/\delta \rfloor} \frac{1}{i}(1+o(1))i \leq (1+o(1))/\delta.$$

Part (ii): If $i > 1/\delta$ then, from the above,

$$\sum_{j=1}^{\lfloor 1/\delta \rfloor} \nu^0(i,j) \leq \frac{1}{1-\mathrm{e}^{-(1+o(1))\delta}} \leq (1+o(1))/\delta.$$

Noting that $X_\star \leq n\delta = \sqrt{n}c$ and $1/\delta = \sqrt{n}/c$,

$$\sum_{i=\lfloor 1/\delta \rfloor+1}^{X_\star - \Sigma\sqrt{n}} \sum_{j=1}^{\lfloor 1/\delta \rfloor} s_{ij} \leq \frac{1+o(1)}{\delta}(\log(\sqrt{n}c) - \log(\sqrt{n}/c)) = \frac{2+o(1)}{\delta}\log(c).$$

Part (iii): Since $i \geq j$, $b_j = (1-\epsilon_n)r(X_\star - j)/n \leq (1-\epsilon_n)r(X_\star - i)/n = b_i$, $\nu^0(i,j) \leq \mathrm{e}^{-(1-o(1))b_i(i-j)}$, and $\sum_{j \leq i} \nu^0(i,j) \leq 1/(1-\mathrm{e}^{-(1-o(1))b_i}) \sim 1/b_i = n/(X_\star - i)$ uniformly over i, since $b_i \leq \delta$ and $\delta \to 0$. This gives

$$\sum_{i=\lfloor 1/\delta \rfloor+1}^{\lfloor X_\star - \Sigma\sqrt{n} \rfloor} \sum_{j=\lfloor 1/\delta \rfloor+1}^{i} s_{ij} \leq \sum_{i=\lfloor 1/\delta \rfloor+1}^{\lfloor X_\star - \Sigma\sqrt{n} \rfloor} \frac{n}{i(X_\star - i)} \leq \frac{n}{X_\star} \sum_{i=\lfloor 1/\delta \rfloor+1}^{\lfloor X_\star - \Sigma\sqrt{n} \rfloor} \left(\frac{1}{i} + \frac{1}{X_\star - i}\right).$$

Treating the sums as Riemann sums and noting that $n/X_\star \leq n/(n\delta - 1) \sim \frac{1}{\delta}$, $\delta X_\star \leq n\delta^2 = c^2$ and $X_\star/(\Sigma\sqrt{n}) \leq \sqrt{n}\delta/\Sigma = o(\sqrt{c})$, this is at most

$$\frac{1+o(1)}{\delta}\left(\log(X_\star) - \log(1/\delta) + \log(X_\star) - \log(\Sigma\sqrt{n})\right) \leq \frac{1+o(1)}{\delta}(c^2 + o(\sqrt{c})).$$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Part (iv): Writing as a product and using (46) on the first term, then proceeding as in part (iii) on the sum, for $i \ge X_\star - \Sigma\sqrt{n} \ge j$,

$$\sum_{j \le X_\star - \Sigma\sqrt{n}} \nu^0(i,j) = \nu^0(i, X_\star - \Sigma\sqrt{n}) \sum_{j \le X_\star - \Sigma\sqrt{n}} \nu^0(X_\star - \Sigma\sqrt{n}, j)$$
$$\le \mathrm{e}^{n^{5/8} O(1/n)} / (1 - \mathrm{e}^{-(1-o(1))\Sigma/\sqrt{n}}) = (1+o(1))\sqrt{n}/\Sigma.$$

Since $\Sigma = o(c) = o(X_\star/\sqrt{n})$, $1/i \sim 1/X_\star$ and

$$\sum_{i=X_\star - \Sigma\sqrt{n}+1}^{X_\star - 1} \sum_{j=\lfloor 1/\delta \rfloor + 1}^{X_\star - \Sigma\sqrt{n}} s_{ij} \le (1+o(1)) \frac{\sqrt{n}}{\Sigma} \sum_{i=X_\star - \Sigma\sqrt{n}+1}^{X_\star - 1} \frac{1}{i}$$
$$\le (1+o(1)) \frac{\sqrt{n}}{\Sigma} \frac{\Sigma\sqrt{n}}{X_\star} \sim \frac{n}{X_\star} \le \frac{1}{\delta}.$$

Part (v): Using (46) as in part (iv), $\nu^0(i,j) \le \mathrm{e}^{\Sigma\sqrt{n} O(1/n)} = 1 + o(1)$. Again, $1/i \sim 1/X_\star$. Since there are at most $\Sigma^2 n$ terms in the sum and $\Sigma \le \sqrt{c}$, it is bounded by

$$(1+o(1)) \frac{\Sigma^2 n}{X_\star} = (1+o(1)) \frac{c}{\delta}.$$

In all five parts, the sum is $O(c^2/\delta)$; referring to (40), this is $o(E^{\mathrm{o}}_\star)$.

Case 2: $\delta \to \delta_\infty > 0$. Since, by (48), $q^-_+(j)/q^0_-(j) \le 1$ for $j \le X_\star - \Sigma\sqrt{n}$ and, by (46), $q^0_+(j)/q^0_-(j) \le 1 + O(1/n) \le \mathrm{e}^{O(1/n)}$ for $X_\star - \Sigma\sqrt{n} \le j < X_\star$, and since $\Sigma \le n^{1/8}$, $\nu^0(i,j) \le \mathrm{e}^{n^{5/8} O(1/n)} = 1 + o(1)$ for $j \le i < X_\star$. Thus,

$$\sum_{i=1}^{X_\star - 1} \sum_{j=1}^{i} s_{ij} \le (1+o(1)) \sum_{i=1}^{X_\star - 1} \frac{1}{i}\, i \sim X_\star,$$

and $X_\star \le n\delta = o(E^{\mathrm{o}}_\star)$, again by (40). □

8.3.5. *Exponential limit* Here we prove Lemma 10. By the strong Markov property, this is equivalent to showing that $\tau^{\mathrm{o}}_\star / E^{\mathrm{o}}_\star$ converges in distribution to exponential with mean 1, assuming $X_0 = X_\star$. Let Φ denote the natural coupling, so that for each $j \in \{0, \dots, N\}$, $((\Phi(j,t))_{t \ge 0}$ is a copy of the logistic process with initial value j, and by Lemma 1, $\Phi(i,t) \le \Phi(j,t)$ for all t if $i \le j$; let $\tau_\star(j) = \inf\{t > 0 \colon \Phi(j,t) \in \{0, X_\star\}\}$ and $\tau^{\mathrm{o}}_\star(j) = \sup\{t > 0 \colon \Phi(j,t) = X_\star\}$. We give a sufficient condition for $(\tau^{\mathrm{o}}_\star - \tau_\star)/E^{\mathrm{o}}_\star$ to have an exponential limit.

Lemma 16. *Let $\overline{E}^{\mathrm{o}}_\star = \mathbb{E}[\tau^{\mathrm{o}}_\star(n)]$ and assume that (i) $\mathbb{P}(\tau^{\mathrm{o}}_\star(j) = \tau^{\mathrm{o}}_\star(n) \mid \Phi(j, \tau_\star(j)) = X_\star) = 1 - o(1)$ uniformly over $j \in \{1, \dots, n\}$, and (ii) uniformly over $j \in \{1, \dots, n\}$ and $t > 0$, $\mathbb{P}(\tau^{\mathrm{o}}_\star(n) > t\overline{E}^{\mathrm{o}}_\star \mid \Phi(j, \tau_\star(j)) = X_\star) \ge \mathbb{P}(\tau^{\mathrm{o}}_\star(n) > t\overline{E}^{\mathrm{o}}_\star) - o(1)$. Then, for each $t > 0$, $\mathbb{P}(\tau^{\mathrm{o}}_\star(X_\star) > tE^{\mathrm{o}}_\star) \to \mathrm{e}^{-t}$.*

Proof. By definition, $\Phi(X_\star, \tau^{\mathrm{o}}_\star(X_\star)) = X_\star$, so using assumption (i) it is enough to show that $\mathbb{P}(\tau^{\mathrm{o}}_\star(n) > tE^{\mathrm{o}}_\star) \to \mathrm{e}^{-t}$ for $t > 0$. By the definition of $\tau^{\mathrm{o}}_\star(n)$, $\overline{E}^{\mathrm{o}}_\star = \mathbb{E}[\tau^{\mathrm{o}}_\star \mid X_0 = n]$, so using the Markov property and then Lemma 8 we find that $\overline{E}^{\mathrm{o}}_\star = E^{\mathrm{o}}_\star + \mathbb{E}[\tau_\star \mid X_0 = n] \sim E^{\mathrm{o}}_\star$. Thus, it is enough to show that $p_n(t) := \mathbb{P}(\tau^{\mathrm{o}}_\star(n) > t\overline{E}^{\mathrm{o}}_\star) \to \mathrm{e}^{-t}$ for $t > 0$.

Using the natural coupling of Section 3, if $j \le n$ then, for any $t > 0$, $\Phi(j,t) \le \Phi(n,t)$, which implies $\tau^{\mathrm{o}}_\star(j) \le \tau^{\mathrm{o}}_\star(n)$. Conditioning on the value of $\Phi(n,t)$ and using the Markov

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

property, it follows that $p_n(t+s) = p_n(t) \sum_j \mathbb{P}(\tau_\star^{\mathrm{o}}(j) > s\overline{E}_\star^{\mathrm{o}})\mathbb{P}(\Phi(n, t\overline{E}_\star^{\mathrm{o}}) = j) \le p_n(t)p_n(s)$, i.e. $t \mapsto p_n(t)$ is submultiplicative for each n. Given $t, s > 0$, conditioning on $\Phi(n, t\overline{E}_\star^{\mathrm{o}})$, using the Markov property, then using assumption (i) followed by assumption (ii) and the law of total probability,

$$\begin{aligned}
&\mathbb{P}(\tau_\star^{\mathrm{o}}(n) > (t+s)\overline{E}_\star^{\mathrm{o}} \mid \tau_\star^{\mathrm{o}}(n) > t\overline{E}_\star^{\mathrm{o}}) \\
&\qquad = \sum_j \mathbb{P}(\tau_\star^{\mathrm{o}}(j) > s\overline{E}_\star^{\mathrm{o}}) \mid \Phi(j, \tau_\star(j)) = X_\star)\mathbb{P}(\Phi(n, t\overline{E}_\star^{\mathrm{o}}) = j) \\
&\qquad \ge \sum_j \mathbb{P}(\tau_\star^{\mathrm{o}}(n) > s\overline{E}_\star^{\mathrm{o}}) \mid \Phi(j, \tau_\star(j)) = X_\star)\mathbb{P}(\Phi(n, t\overline{E}_\star^{\mathrm{o}}) = j) - o(1) \\
&\qquad \ge \sum_j \mathbb{P}(\tau_\star^{\mathrm{o}}(n) > s\overline{E}_\star^{\mathrm{o}}))\mathbb{P}(\Phi(n, t\overline{E}_\star^{\mathrm{o}}) = j) - o(1) \\
&\qquad = p_n(s) - o(1)
\end{aligned}$$

uniformly over t and s. Since the above left-hand side is just $p_n(t+s)/p_n(t)$, we obtain $p_n(t+s) \ge p_n(t)p_n(s) - o(1)$. Combining with $p_n(t+s) \le p_n(t)p_n(s)$, it follows easily that for rational t, $p_n(t) = p_n(1)^t + o(1)$, and since $t \mapsto p_n(t)$ is non-increasing, the same holds for real $t > 0$ by rational approximation. Thus, it remains only to show that $p_n(1) \to \frac{1}{\mathrm{e}}$ as $n \to \infty$.

Let $T(n) = \tau_\star^{\mathrm{o}}(n)/\overline{E}_\star^{\mathrm{o}}$, so that $\mathbb{E}[T(n)] = 1$ for each n. By Markov's inequality, $p_n(2) \le \frac{1}{2}$, and $p_n(2j) \le p_n(2)^j = 2^{-j}$ for integer $j \ge 1$, so $\mathbb{E}[\, T(n)\mathbf{1}(T(n) > 2k)\,] \le \sum_{j \ge k} 2\mathbb{P}(T_\star^{\mathrm{o}} > 2j) \le 2^{-(k-2)}$, which $\to 0$ as $k \to \infty$ uniformly in n. Since $\mathbb{P}(T(n) > t) = \mathbb{P}(T(n) > 1)^t + o(1)$ for each t, an easy approximation argument using the monotonicity of $t \mapsto \mathbb{P}(T(n) > t)$ then shows that $\mathbb{E}[\, T(n)\,] = \int_0^\infty \mathbb{P}(T(n) > 1)^t \, \mathrm{d}t + o(1)$. Since $\mathbb{E}[\, T(n)\,] = 1$ for all n, it follows that $\mathbb{P}(T(n) > 1) \to \frac{1}{e}$ as $n \to \infty$, as desired. □

It remains to show that assumptions (i) and (ii) of Lemma 16 are satisfied. Let $\tau_{\mathrm{c}}(j) = \inf\{t\colon \Phi(j, t) = \Phi(n, t)\}$ denote the coupling time of the two trajectories, which is a.s. finite since both eventually hit 0. We begin by extracting a further sufficient condition, which we then prove.

Lemma 17. *Let* $\tau_{\mathrm{c}}^\star(j) = \inf\{t \ge \tau_{\mathrm{c}}(j)\colon \Phi(j, t) = X_\star\}$. *Suppose that*

$$\min_{j \in \{1, \dots, n\}} \mathbb{P}(\tau_{\mathrm{c}}^\star(j) < \infty \mid \Phi(j, \tau_\star(j)) = X_\star) \to 1. \tag{49}$$

Then assumptions (i) and (ii) of Lemma 16 are satisfied.

Proof. Assumption (i): Since $\Phi(j, t) = \Phi(n, t)$ for all $t \ge \tau_{\mathrm{c}}(j)$ (see Lemma 3.1), the event $\tau_\star^{\mathrm{o}}(j) = \tau_\star^{\mathrm{o}}(n)$ is equivalent to the event that $\Phi(j, t) = \Phi(n, t) = X_\star$ for some $t > 0$, which in turn is equivalent to the event $\tau_{\mathrm{c}}^\star(j) < \infty$, and assumption (i) follows directly from (49).

Assumption (ii): First note that, since $\mathbb{P}(\Phi(X_\star, \tau_\star(X_\star)) = X_\star) = 1$, it follows from the above that

$$\mathbb{P}(\tau_\star^{\mathrm{o}}(X_\star) = \tau_\star^{\mathrm{o}}(n)) = 1 - o(1). \tag{50}$$

Next, using the strong Markov property, $\tau_\star^{\mathrm{o}}(n)$ conditioned on $\tau_{\mathrm{c}}^\star(j) < \infty$ is equal in distribution to $\tau_{\mathrm{c}}^\star(j)$ conditioned on $\tau_{\mathrm{c}}^\star(j) < \infty$, plus an independent copy of $\tau_\star^{\mathrm{o}}(X_\star)$. In particular, $\tau_\star^{\mathrm{o}}(n)$, conditioned on $\tau_{\mathrm{c}}^\star(j) < \infty$, dominates $\tau_\star^{\mathrm{o}}(X_\star)$ (with no conditioning). Since $\tau_{\mathrm{c}}^\star(j) < \infty$

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

implies $\Phi(j, \tau_\star(j)) = X_\star$, using (49), then the above observation, then (50), it follows that, uniformly over j,

$$\begin{aligned}
&\mathbb{P}(\tau_\star^{\mathrm{o}}(n) > t\overline{E}_\star^{\mathrm{o}} \mid \Phi(j, \tau_\star(j)) = X_\star) \\
&\qquad = \mathbb{P}(\tau_\star^{\mathrm{o}}(n) > t\overline{E}_\star^{\mathrm{o}} \mid \tau_{\mathrm{c}}^\star(j) < \infty)\mathbb{P}(\tau_{\mathrm{c}}^\star(j) < \infty \mid \Phi(j, \tau_\star(j)) = X_\star) \\
&\qquad = \mathbb{P}(\tau_\star^{\mathrm{o}}(n) > t\overline{E}_\star^{\mathrm{o}} \mid \tau_{\mathrm{c}}^\star(j) < \infty)(1 - o(1)) \\
&\qquad \geq \mathbb{P}(\tau_\star^{\mathrm{o}}(X_\star) > t\overline{E}_\star^{\mathrm{o}})(1 - o(1)) \\
&\qquad = (\mathbb{P}(\tau_\star^{\mathrm{o}}(n) > t\overline{E}_\star^{\mathrm{o}}) - o(1))(1 - o(1)) \\
&\qquad = \mathbb{P}(\tau_\star^{\mathrm{o}}(n) > t\overline{E}_\star^{\mathrm{o}}) - o(1).
\end{aligned}$$

□

Finally, we prove the hypothesis of Lemma 17. To do so we show that, within a short time after the paths started from j and from n reach $X_\star$, they meet (if they have not met already), and then with probability $1 - o(1)$ their common trajectory hits $X_\star$ at least once more before going to 0.

Lemma 18. *As $n \to \infty$, $\min_{j\in\{1,\dots,n\}} \mathbb{P}(\tau_{\mathrm{c}}^\star(j) < \infty \mid \Phi(j, \tau_\star(j)) = X_\star) \to 1$.*

Proof. Let X^j, X^n denote the processes $(\Phi(j, t))_{t\geq 0}, (\Phi(n, t))_{t\geq 0}$. As we will see, when $X^j, X^n > (1+\epsilon)X_\star/2$, the drift tends to push them together. Let $\tau_b(j) = \inf\{t > \tau_\star(j) \vee \tau_\star(n) \colon \max(|X_t^j - nx_\star|, |X_t^n - nx_\star|) \geq nx_\star/4\}$. If $\tau_{\mathrm{c}}(j) < \tau_\star(j) \vee \tau_\star(n)$ then $X_t^j = X_t^n = X_\star$ with $t = \tau_\star(j) \vee \tau_\star(n)$, which implies $\tau_{\mathrm{c}}^\star(j) < \infty$. On the other hand, if $\tau_\star(j) \vee \tau_\star(n) < \tau_{\mathrm{c}}(j) < \tau_b(j)$ then $X^j_{\tau_{\mathrm{c}}(j)} = X^n_{\tau_{\mathrm{c}}(j)} \geq 3X_\star/4$. Using the strong Markov property and the fact that X^j and X^n remain together once they meet, on the latter event it follows from Lemma 8.1 that $\tau_{\mathrm{c}}^\star(j) < \infty$ with probability $1 - o(1)$ uniformly over j. Thus, it is enough to show that

$$\max_j \mathbb{P}(\tau_b(j) \wedge \tau_{\mathrm{c}}(j) < \tau_b(j) \mid \Phi(j, \tau_\star(j)) = X_\star) \to 1. \tag{51}$$

We begin with a lower bound on $\tau_b(j)$, that ensures both X^j and X^n remain fairly close to $X_\star$ for a while after they hit it. Then, we estimate the drift and diffusivity of $X^j - X^n$ assuming that both are at least $3X_\star/4$, and with the help of the lower bound, deduce (51).

Lower bound on $\tau_b(j)$: Let $W = X - nx_\star$. Then $\mu(W) = \mu(X) = X(r(1 - X/n) - 1) = rX(x_\star - X/n) = -rXW/n$ and $\sigma^2(W) = \sigma^2(X) = X(r(1 - X/n) + 1) \leq (1 + r)X$. Since W jumps by ± 1, if $|W| \geq 1$ then $\mu(|W|) = \mathrm{sgn}(W)\mu(W)$ and $\sigma^2(|W|) = \sigma^2(W)$. Suppose $nx_\star/8 \leq |W| \leq nx_\star/4$, noting that $1 \leq nx_\star/8$ for large n. Since $3nx_\star/4 \leq X \leq 5nx_\star/4$, $\sigma^2(|W|) = O(nx_\star) = O(n\delta)$, $\mu(|W|) \leq -(3rx_\star/4)(nx_\star/8) = -3n\delta^2/32r$, and $|\mu(|W|)| \leq (5rx_\star/4)nx_\star/4 = O(n\delta^2)$. We shall use Lemma 19 (with apologies for overloading notation). In the notation of Lemma 19, let $X = |W| - nx_\star/8$, $x = nx_\star/8 = n\delta/8r$, $\mu_\star = 3n\delta^2/32r = 3c^2/32r$, $\sigma_\star^2 = Cn\delta$, and $C_{\mu_\star} = Cn\delta^2$ for some $C > 0$ and $C_\Delta = \frac{1}{2}$. Since $\Delta_\infty(X) = 1$, $\Delta_\infty(X)\mu_\star/\sigma_\star^2 = 3\delta/32Cr$ is at most $\frac{1}{2}$ if $C > 0$ is chosen large enough. Then, $\Gamma = \exp(\Omega(c^2))$ and $x/16C_{\mu_\star} = \Omega(1/\delta)$, so $\mathbb{P}(\sup_{t\leq(1/\delta)\exp(\Omega(c^2))} |W_t| > nx_\star/4 \,\big|\, |W_0| \leq nx_\star/8) = o(1)$. Applying this bound to X^j, X^n from time $\tau_\star(j)$, respectively $\tau_\star(n)$, we find that

$$\mathbb{P}(\tau_b(j) \leq (1/\delta)\exp(\Omega(c^2)) \mid \Phi(j, \tau_\star(j)) = X_\star) = o(1) \tag{52}$$

uniformly over $j \in \{1, \dots, n\}$.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Upper bound on $\tau_b(j) \wedge \tau_c(j)$: Let $F(x) = x(r(1-x)-1) = rx(x_\star - x)$ and $G(x) = x(r(1-x)+1) \geq x$, so that $\mu(X) = nF(X/n)$ and $\sigma^2(X) = nG(X/n) \geq X$. We have $F'(x) = r(x_\star - 2x)$, so if $3x_\star/4 \leq x \leq 5x_\star/4$ then $F'(x) \in [F'(5x_\star/4), F'(3x_\star/4)] = [-3rx_\star/2, -rx_\star/2] = [-3\delta/2, -\delta/2]$. If $X^j, X^n \geq 3nx_\star/4$ and $X^j \neq X^n$ then, letting $U = X^j - X^n$, by the mean value theorem, $\mu(U)/(U) = n(F(X^j/n) - F(X^n/n))/(X^j - X^n) \in [-3\delta/2, -\delta/2]$ and, since X^j and X^n evolve independently until they collide, $\sigma^2(U) = n(G(X^j/n) + G(X^n/n)) \geq 3nx_\star/2 \geq n\delta/r$. Since U jumps by ± 1 and takes values in $\mathbb{Z}$, if $U \neq 0$ then $\mu(|U|) = \mathrm{sgn}(U)\mu(U)$ and $\sigma^2(|U|) = \sigma^2(U)$, so letting $V = |U|$, the above implies that, conditional on $\Phi(j, \tau_\star(j)) = X_\star$, $\mu_t(V) \in [-(3\delta/2)V_t, -(\delta/2)V_t]$ and $\sigma_t^2(V) \geq n\delta/r$ for all $\tau_\star(j) \vee \tau_\star(n) \leq t < \tau_c(j) \wedge \tau_b(j)$. If this interval is empty, then $\tau_c(j) \wedge \tau_b(j) < \tau_\star(j) \vee \tau_\star(n) \leq \tau_b(j)$ so there is nothing to show. Otherwise, since on this time interval $\mathrm{sgn}(U_t)$ is fixed, then, given $\mathrm{sgn}(U_{\tau_\star(j)})$, on the same time interval V is a Markov chain with state space a subset of $\mathbb{Z}$. Let $u = \lfloor \sqrt{n} \rfloor$, $t_0 = \tau_\star(j)$, and $t_1 = \tau_b(j) \wedge \inf\{t > t_0 \colon V_t = u\}$, and define recursively $t_i = \tau_b(j) \wedge \inf\{t > t_{i-1} \colon V_t \in \{0, u, 2u\} \setminus V_{t_{i-1}}\}$. Let $\rho_i = t_i - t_{i-1}$ for $i = 1, \ldots, N = \min\{i \colon V_{t_i} = 0 \text{ or } t_i = \tau_b(j)\} = \min\{i \colon t_i = \tau_c(j) \wedge \tau_b(j)\}$. By the a priori bound, we may assume that $|V_{t_0}| \leq nx_\star/2 = n\delta/2r$. Then, $\xi_t = \mathrm{e}^{\delta(t_0 + t \wedge \rho_1)/2} V_{t_0 + t \wedge \rho_1}$ is a supermartingale with $\xi_0 \leq n\delta/2r$, so

$$\mathbb{P}(\rho_1 > t) \leq \mathbb{P}(\xi_t > \mathrm{e}^{\delta t/2} u) \leq \mathrm{e}^{-2\delta t} \frac{n\delta/2r}{\sqrt{n} - 1} \sim \mathrm{e}^{-2\delta t} c/2r.$$

This probability is $O(1/c) = o(1)$ if $t = (1/\delta)\log c$. Using a similar estimate with $\xi_0 = 2u$, $\mathbb{P}(\rho_i > t \mid V_{\rho_{i-1}} = 2u) \leq \mathrm{e}^{-2\delta t} \frac{2u}{u} \leq 2\mathrm{e}^{-2\delta t}$ and, integrating over t, $\mathbb{E}[\rho_i \mid V_{\rho_{i-1}} = 2u] \leq \frac{1}{\delta}$. To estimate ρ_2, we note that, for $\alpha > 0$, $\mu_t(\alpha V^2) = 2\alpha V_t \mu(V_t) + \alpha^2 \sigma_t^2(V) \geq -3\alpha\delta V_t^2 + \alpha^2 n\delta/r$ so, for $t_1 \leq t < t_2$, since $V_t \leq 2u \leq 2\sqrt{n}$, choosing $\alpha = 13r$ we have $\mu_t(\alpha V^2) \geq \alpha n\delta$. Thus, $V^2_{t_1 + t \wedge \rho_2} - n\delta(t \wedge \rho_2)$ is a submartingale. Since $V_{t_1} = u$ and $V_{t_2} \leq 2u$, using optional stopping,

$$\mathbb{E}[\rho_2] \leq \frac{1}{n\delta}(\mathbb{E}[V_{t_2}^2 - V_{t_1}^2]) \leq \frac{1}{n\delta}(4u^2 - u^2) \leq \frac{3n}{n\delta} = \frac{3}{\delta}.$$

By the Markov property, the same estimate holds for $\mathbb{E}[\rho_i \mid V_{t_{i-1}} = u]$. Using simply that $\mu_t(V) \leq 0$ and optional stopping, $\mathbb{P}(V_{t_i} = 2u \mid V_{t_{i-1}} = u) \leq 1/2$. Summarizing, on the time interval $[\tau_\star(j) \vee \tau_\star(n), \tau_c(j) \wedge \tau_b(j)]$, V hits u, then goes to $2u$ and back to u at most geometric$(1/2)$ number of times before either $V_{t_i} = 0$ or $t_i = \tau_b(j)$. As shown above, ρ_1, the time to first hit u, is at most $\frac{1}{\delta}\log c$ with probability $1 - o(1)$, and the expected time to go from u to either 0 or to $2u$ and back to u is at most $(4/\delta)$. Using Wald's lemma and $\mathbb{E}[\text{geometric}(1/2) = 2]$, $\sum_{i=2}^{N-1} \rho_i \leq 8/\delta$. Using Markov's inequality, the sum is at most $\frac{1}{\delta}\log c$ with probability $1 - o(1)$, so combining with the estimate on ρ_1, we find that, uniformly over $j \in \{1, \ldots, n\}$, $\mathbb{P}(\tau_b(j) \wedge \tau_c(j) - \tau_\star(j) \vee \tau_\star(n) > (2/\delta)\log c \mid \Phi(j, \tau_\star(j)) = X_\star) = o(1)$. From Lemma 14, if $\delta \to 0$ then $\mathbb{E}[\tau_\star(j)]$, $\mathbb{E}[\tau_\star(n)] = O(c^2 \log(c)/\delta)$ uniformly over $j \in \{1, \ldots, n\}$, so using Markov's inequality, with probability $1 - o(1)$ uniformly in j, $\tau_\star(j) \vee \tau_\star(n) = o(\mathrm{e}^{\epsilon c^2}/\delta)$ for any fixed $\epsilon > 0$. Summing the two and combining with (52), we obtain (51). □

Appendix A. Stochastic calculus

We recall a useful probability estimate and diffusion limit result, stated in the context of semimartingales. We give here a very brief list of definitions, enough for the acquainted reader to understand the context for this paper; for an overview of the theory, see [13].

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Recall that a semimartingale is an optional process X that can be written $X = X_0 + M + A$, where M is a local martingale and A has finite variation. It is special if A can be taken to be predictable, in which case we write $X = X_0 + X^{\mathrm{m}} + X^{\mathrm{p}}$, where X^{m} is the martingale part and X^{p} is the (predictable) compensator. A sufficient condition for X to be special is if it has bounded jumps, i.e. if the process of jumps $\Delta X_t = X_t - X_{t^-}$ satisfies $|\Delta X| \le \gamma$ for some non-random $\gamma < \infty$. If so, let $\Delta_\infty(X)$ denote the least such γ. In this case, a fortiori X^{m} is locally square-integrable, i.e. the predictable quadratic variation $\langle X \rangle$ exists.

A process is *quasi-left continuous* if $\Delta X_T = 0$ a.s. on $\{T < \infty\}$ for any predictable time T. Feller processes, which include continuous-time Markov chains, are quasi-left continuous. As noted in [12], if X is special and X^{m} is locally square-integrable then X is quasi-left continuous if and only if both $\langle X^{\mathrm{m}} \rangle$ and X^{p} are continuous. This motivates the following definition (not found in other references).

Definition 1. Let X be a special supermartingale with X^{m} locally square-integrable. Then X is quasi-absolutely continuous if both X^{p} and $\langle X^{\mathrm{m}} \rangle$ are absolutely continuous. In this case define the drift $\mu(X)$ and diffusivity $\sigma^2(X)$ by

$$\mu_t(X) = \frac{\mathrm{d}}{\mathrm{d}t} X_t^{\mathrm{p}}, \qquad \sigma_t^2(X) = \frac{\mathrm{d}}{\mathrm{d}t} \langle X^{\mathrm{m}} \rangle_t. \tag{53}$$

Any right-continuous continuous-time Markov chain X on a finite state space $S \subset \mathbb{R}$ has finite variation so is a supermartingale. Index the possible transitions by $i \in \{1, \ldots, m\}$ for some m, with $q_i : S \to \mathbb{R}_+$ the rates and $\Delta_i : S \to S - S$ the jumps. Writing X as a sum of jumps and using the standard linear and quadratic martingales for Poisson processes, it is easy to show that X is quasi-absolutely continuous and has

$$\mu_t(X) = \sum_{i=1}^{m} q_i(X_t)\Delta_i(X_t), \qquad \sigma_t^2(X) = \sum_{i=1}^{m} q_i(X_t)(\Delta_i(X_t))^2. \tag{54}$$

Our first result gives a strong (exponential in μ/σ^2) lower bound on the escape time from a barrier with negative drift. It is proved in [4].

Lemma 19. *(Drift barrier.) Fix $x > 0$ and let X be a quasi-absolutely continuous supermartingale on $\mathbb{R}$ with jump size $\Delta_\infty(X) \le x/2$. Suppose there are positive reals $\mu_\star$, $\sigma_\star^2$, $C_{\mu_\star}$, and C_Δ with $\max\left\{\Delta_\infty(X)\mu_\star/\sigma_\star^2, \frac{1}{2}\right\} \le C_\Delta$ so that if $0 < X_t < x$ then $\mu_t(X) \le -\mu_\star$, $|\mu_t(X)| \le C_{\mu_\star}$, and $\sigma_t^2(X) \le \sigma_\star^2$. Let $\Gamma = \exp\left(\mu_\star x/(32 C_\Delta \sigma_\star^2)\right)$. Then we have*

$$P\left(\sup_{t \le \lfloor \Gamma \rfloor x/16 C_{\mu_\star}} X_t \ge x \mid X_0 \le x/2 \right) \le 4/\Gamma.$$

The next result gives a diffusion limit, assuming the drift and diffusivity converge while the jump size tends to 0. It follows from [9, Theorem 4.1, Chapter 7], and from the Lipschitz existence and uniqueness condition for SDEs, if (i) in the proof of the former result we let τ_n^R be the exit time of X^n from $\left(\frac{1}{R}, R\right)$ instead of $(-R, R)$, as described below, and (ii) we allow that the limiting diffusion Y may be defined only on the interval $[0, \zeta)$ where $\zeta := \lim_{R \to \infty} \tau^R$, with τ^R as defined below.

Lemma 20. *(Diffusion limit) Let X^n be a sequence of quasi-absolutely continuous semimartingales with drift and diffusivity given by functions μ_n, σ_n^2, and suppose $a : (0, \infty) \to \mathbb{R}_+$ and $b : (0, \infty) \to \mathbb{R}$ are such that $\sqrt{a}$ and b are Lipschitz on compact subsets of $(0, \infty)$.*

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press

Suppose the largest jump in X^n tends to 0 as $n \to \infty$. Also assume that, for each $R > 0$, as $n \to \infty$, $\sup_{|x| \le R} |\mu_n(x) - b(x)|$, $|\sigma_n^2(x) - a(x)| \to 0$. Suppose $X^n(0) \to x \in \mathbb{R}$ and let $\tau_n^R = \inf\{t\colon X^n(t) \notin (\frac{1}{R}, R)$ or $X^n(t^-) \notin (\frac{1}{R}, R)\}$. Then, for all but countably many R, $X^n(\cdot \wedge \tau_n^R)$ converges in distribution to $X(\cdot \wedge \tau^R)$, where X solves the initial value problem $x_0 = x$, $\mathrm{d}x = b(x)\,\mathrm{d}t + \sqrt{a(x)}\,\mathrm{d}B$, and $\tau^R = \inf\{t\colon X(t) \notin (\frac{1}{R}, R)\}$.

References

[1] Andersson, H. and Djehiche, B. (1998). A threshold limit theorem for the stochastic logistic epidemic. *J. Appl. Prob.* **35**, 662–670.
[2] Athreya, K. B. and Ney, P. E. (1972). *Branching Processes*. Springer, Berlin.
[3] Barbour, A. D., Chigansky, P. and Klebaner, F. (2015). On the emergence of random initial conditions in fluid limits. *J. Appl. Prob.* **53**, 1193–1205.
[4] Basak, A., Durrett, R. and Foxall, E. (2018). Diffusion limit for the partner model at the critical value. *Electron. J. Prob.* **23**, 102.
[5] Brightwell, G., House, T., and Luczak, M. (2018). Extinction times in the subcritical stochastic SIS logistic epidemic. *J. Math. Biol.* **77**, 455–493.
[6] Doering, C. R., Sargsyan, K. V. and Sander, L. M. (2005). Extinction times for birth–death processes: Exact results, continuum asymptotics, and the failure of the Fokker–Planck approximation. *Multiscale Model Simul.* **3**, 283–299.
[7] Dolgoarshinnykh, R. G. and Lalley, S. P. (2006). Critical scaling for the SIS stochastic epidemic. *J. Appl. Prob.* **43**, 892–898.
[8] Durrett, R. (2010). *Probability: Theory and Examples*, 4th edn. Cambridge University Press.
[9] Ethier, S. N., and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence.* John Wiley, New York.
[10] Feller, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheoretica* **5**, 11–40.
[11] Feller, W. (2015). *Selected Papers*, Vol. **I**. Springer, New York.
[12] Foxall, E. (2018). The naming game on the complete graph. *Electron. J. Prob.* **23**, 126.
[13] Jacod, J. and Shiryaev, A. N. (2002). *Limit Theorems for Stochastic Processes.* Springer, New York.
[14] Kallenberg, O. (1997). *Foundations of Modern Probability.* Springer, New York.
[15] Keilson, J. (1979). *Markov Chain Models: Rarity and Exponentiality* (Appl. Math. Sci. 28). Springer, New York.
[16] Kryscio, R. J. and LefÉvre, C. (1989). On the extinction of the S-I-S stochastic logistic epidemic. *J. Appl. Prob.* **26**, 685–694.
[17] Kurtz, T. G. (1978). Strong approximation theorems for density-dependent Markov chains. *Stoch. Process. Appl.* **6**, 223–240.
[18] Nasell, I. (1996). The quasi-stationary distribution of the closed endemic SIS model. *Adv. Appl. Prob.* **28**, 895–932.
[19] Nasell, I. (1999). On the quasi-stationary distribution of the stochastic logistic epidemic. *Math. Biosci.* **156**, 21–40.
[20] Nasell, I. (2011). *Extinction and Quasi-Stationarity in the Stochastic Logistic SIS Model.* Springer, New York.
[21] Talvila, E. and Wiersma, M. (2012). Simple derivation of basic quadrature formulas. arXiv:1202.0249.

https://doi.org/10.1017/jpr.2020.112 Published online by Cambridge University Press