MULLER'S RATCHET IN A NEAR-CRITICAL REGIME: TOURNAMENT VERSUS FITNESS PROPORTIONAL SELECTION.

JAN LUKAS IGELBRINK, ADRIÁN GONZÁLEZ CASANOVA, CHARLINE SMADI, AND ANTON WAKOLBINGER

ABSTRACT. Muller's ratchet, in its prototype version, models a haploid, asexual population whose size N is constant over the generations. Slightly deleterious mutations are acquired along the lineages at a constant rate, and individuals carrying less mutations have a selective advantage. The classical variant considers *fitness proportional* selection, but other fitness schemes are conceivable as well. Inspired by the work of Etheridge et al. [EPW09] we propose a parameter scaling which fits well to the "near-critical" regime that was in the focus of [EPW09] (and in which the mutation-selection ratio diverges logarithmically as $N \to \infty$). Using a Moran model, we investigate the "rule of thumb" given in [EPW09] for the click rate of the "classical ratchet" by putting it into the context of new results on the long-time evolution of the size of the best class of the ratchet with (binary) tournament selection, which (other than that of the classical ratchet) follows an autonomous dynamics up to the time of its extinction.

In [GSW23] it was discovered that the tournament ratchet has a hierarchy of dual processes which can be constructed on top of an Ancestral Selection graph with a Poisson decoration. For a regime in which the mutation/selection-ratio remains bounded away from 1, this was used in [GSW23] to reveal the asymptotics of the click rates as well as that of the type frequency profile between clicks. We will describe how these ideas can be extended to the near-critical regime in which the mutation-selection ratio of the tournament ratchet converges to 1 as $N \to \infty$.

CONTENTS

1. INTRODUCTION

Muller's ratchet is a prototype model in population genetics. Originally it was conceived to explain the ubiquity of sexual reproduction among eukaryotes despite its many costs [Mul64, Fel74]. In its bare bones version, Muller's ratchet models a haploid, asexual population whose size N is constant over the generations. The neutral part of the random reproduction is given by a Wright-Fisher or a Moran dynamics. Slightly deleterious mutations are acquired along the lineages at a rate m, and individuals carrying less mutations have a selective advantage. The classical variant of Muller's ratchet considers *fitness proportional* selection, where the selective advantage of an individual carrying κ deleterious mutations over a contemporanean that carries a larger number κ' of deleterious mutations is $\frac{s}{N}(\kappa' - \kappa)$. Since the mutation mechanism is assumed to be unidirectional, every once in a while the type with the currently smallest number of mutations κ will disappear from the population. As Herbert Muller puts it in his pioneering paper [Mul64], "*an irreversible ratchet mechanism exists in the non-recombining species . . . that prevents selection, even if intensified, from reducing the mutational loads below the lightest . . . , whereas, contrariwise, 'drift', and what might be called 'selective noise' must allow occasional slips of the lightest loads in the direction of increased weight.*" Every slip of the lightest loads is what we call a *click of the ratchet*. The question "How often does the ratchet click?" was asked by Etheridge, Pfaffelhuber and one of the present authors in [EPW09], and there it was found

that

$$\gamma := \frac{m}{\mathfrak{s} \log(Nm)} \tag{1}$$

is "*an important factor in determining the rate of the ratchet*". Specifically, under the assumption $1 \ll Nm \ll N$, [EPW09] states the following *Rule of Thumb*:

(RT) *The rate of the ratchet is of the order $N^{\gamma-1}m^\gamma$ for $\gamma \in (\frac{1}{2}, 1)$, whereas it is exponentially slow in $(Nm)^{1-\gamma}$ for $\gamma < \frac{1}{2}$.*

For a polynomial mutation rate $m = N^{-\beta}$, $0 < \beta < 1$, the condition that γ remains constant (or at least bounded away from 0 and ∞) as $N \to \infty$ amounts to the requirement that the *mutation-selection ratio*

$$\theta := \frac{m}{\mathfrak{s}}$$

is of the order $\log N$ as $N \to \infty$.
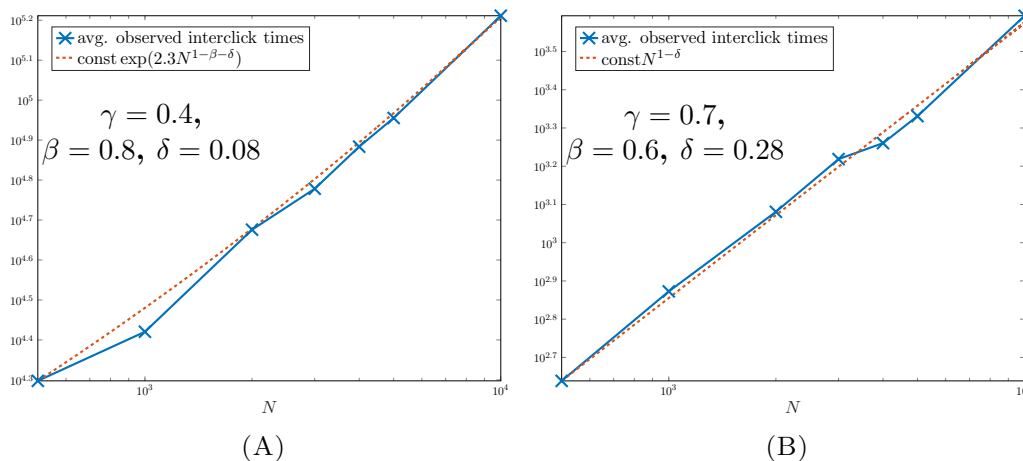

(A)   (B)

FIGURE 1. This is an illustration of the *Rule of Thumb* (RT) predicting the order of magnitude of the interclick times of the classical ratchet. Each data point was obtained by pooling the first 4 interclick times from 100 simulations of the (classical) ratchet for the corresponding parameter configuration (N, β, δ) in the (β, δ)-scaling (2). In the exponential regime, (RT) predicts for the interclick times an order of magnitude $\exp(c\, N^{1-\beta-\delta})$. We see from panel (A) that the constant c is difficult to estimate from simulations up to $N = 10^4$, but $c = 2.3$ as chosen there seems to give a reasonable fit. For the polynomial regime, (RT) predicts the order of $N^{1-\delta}$, which fits very well in the situation of panel (B).

We will thus focus on a family of parameter scalings which we call the (β, δ)-*scaling of the classical ratchet*:

$$m = N^{-\beta}, \qquad \theta = \delta \log N. \tag{2}$$

This means that we consider *moderate mutation-selection*, with m and $\mathfrak{s}$ scaling with the same power $-\beta$ in N. Still, the mutation-selection ratio θ diverges logarithmically in N, and (together with β) the factor δ in front of $\log N$ turns out to be critical for the click rate. In the (β, δ)-scaling, (1) takes the form

$$\gamma = \frac{\delta}{1 - \beta}.$$

The condition $0 < \gamma < 1$ from (RT) restricts the pair (β, δ) to the triangle

$$\Delta := \{(\beta, \delta) : 0 < \beta, \, 0 < \delta < 1 - \beta\}.$$

The *polynomial* and the *exponential regime* predicted by (RT) thus correspond to

$$\mathcal{P} := \{\tfrac{1}{2} < \gamma < 1\} = \{(\beta, \delta) \in \Delta : \tfrac{1}{2}(1 - \beta) < \delta < 1 - \beta\}, \tag{3}$$

$$\mathcal{E} := \{0 < \gamma < \tfrac{1}{2}\} = \{(\beta, \delta) \in \Delta : 0 < \delta < \tfrac{1}{2}(1 - \beta)\}, \tag{4}$$

and the predictions for the orders of magnitude of the expected interclick times take the form

$$N(Nm)^{-\gamma} = N^{1-\delta} = Ne^{-\theta} \qquad \text{for } \gamma \in (\tfrac{1}{2}, 1), \tag{5}$$

$$\exp\left(\text{const}(Nm)^{1-\gamma}\right) = \exp\left(\text{const}\, N^{1-\beta-\delta}\right) \quad \text{for } \gamma \in (0, \tfrac{1}{2}). \tag{6}$$

In view of the predicted transition from polynomial to exponential click rates we refer to $\mathcal{P} \cup \mathcal{E}$ as a *near-critical regime*. See Figure 1 for an illustration of (RT) via simulations.

As observed by John Haigh ([Hai78]), in the deterministic limit ($N \to \infty$ and $m, \mathfrak{s}$ not depending on N) the type frequency profile in equilibrium becomes Poisson with parameter θ. Consequently, $Ne^{-\theta}$ is asymptotically equal to the size of the best class in equilibrium, and the rule (5) goes along with Haigh's prediction that the rate of the ratchet should be proportional to the inverse of the size of the best class (unless the ratchet is not exponentially slowed down by selection).

The evidence for (RT) that is given in [EPW09] is based on a diffusion approximation for the evolution of the relative size X_0 of the *best class* (which consists of the individuals that carry the least amount of mutations in the current population). Because of the fitness proportional selection, the drift coefficient in this diffusion approximation contains the first moment M of the type frequency configuration $(X_0, X_1, \ldots)$. In order to obtain an approximate autonomous dynamics for X_0, the empirical first moment M has to be predicted based on X_0. A classical way to do this uses the so-called *Poisson profile approximation*, which we will explain in some detail in Section 3.

In the present paper we will consider, apart from the classical one, a variant of Muller's ratchet in which fitness proportional selection is replaced by *(binary) tournament selection*. This kind of selection has been studied in the context of evolutionary computation ([BT96, BFM00]) and has found attention also in the biological literature [PBB+15]. In the ratchet's context this means that selective advantage of an individual carrying κ deleterious mutations over a contemporanean that carries a larger number κ' of deleterious mutations is constant (say $\frac{s}{N}$ for some $s = s_N > 0$), irrespective of the value of the difference $\kappa' - \kappa$. For the Moran version of the tournament ratchet, which was introduced in

[GSW23] and whose definition we recall in Section 2, this means that "pairwise selective fights" are always won by the fitter individual.

Other than in the classical ratchet, the size of the (m, s)-tournament ratchet's best class follows an autonomous dynamics. As we will see in Section 3, this dynamics is *equal* to that of the Poisson profile approximation of the size of the classical $(m, \mathfrak{s})$-ratchet's best class, provided that

$$\rho := \frac{m}{s} = 1 - \exp(-m/\mathfrak{s}) = 1 - e^{-\theta}. \tag{7}$$

In view of (7) we define, in analogy to (2), the (β, δ)-scaling for the tournament ratchet as

$$m = N^{-\beta}, \qquad \rho = \frac{m}{s} = 1 - N^{-\delta}. \tag{8}$$

We now state the main result of the present paper. Besides extending [GSW23, Theorem 2.3] on the asymptotic rates of clicks from the subcritical to the near-critical regime, this result also gains importance with regard to [EPW09]'s Rule (RT), in the light of the correspondence (7) between the tournament and the classical ratchet. Recall definitions (3) and (4).

Main result (MR). *In the (β, δ)-scaling of the tournament ratchet, as $N \to \infty$, the expected time between clicks is*

$$\asymp N^{\frac{1+\beta}{2}} \qquad \text{if } (\beta, \delta) \in \mathcal{P}, \tag{9}$$

$$\asymp \exp\left(N^{1-\beta-2\delta}\right) \quad \text{if } (\beta, \delta) \in \mathcal{E}. \tag{10}$$

Here and below, $\asymp$ stands for logarithmic equivalence, i.e. $a_N \asymp b_N$ means $\log a_N \sim \log b_N$, or equivalently $\frac{\log a_N}{\log b_N} \to 1$. See Figure 2 for an illustration of (MR).

While both (RT) and (MR) state the same boundary ($\gamma = \frac{1}{2}$) between the polynomial and the exponential regime, the exponents differ between (5) and (9) as well as between (6) and (10). Specifically, in the polynomial regime $\mathcal{P}$ the exponent $\frac{1+\beta}{2}$ for the tournament ratchet is larger than the exponent $1 - \delta$ for the classical ratchet.

Here is an explanation for the polynomial regime. The centers of attraction of the equilibrium profile weights of the best and the second best class differ by the factor $N^{\frac{\delta}{2}}$ for the tournament ratchet (see Sec. 4.3), while they are given by the Poisson weights $e^{-\theta}$ and $\theta e^{-\theta}$ for the classical ratchet and hence for the latter differ only by the factor $\theta = \delta \log N$ (and thus have the same polynomial order $N^{1-\delta}$). This latter factor is only logarithmic in N; therefore, when starting the "new best class" at the time of a click in its "old" center of attraction, the tournament ratchet has a longer way to go than the classical ratchet. The exponent $\frac{1+\beta}{2}$ in (9) will be obtained by a Green function analysis in the proof of Theorem 3.2. This analysis will also explain the exponent $1 - \delta$ in (5), which corresponds to Haigh's prediction, saying that "the interclick times are of the order of the size of the best class". An intuitive explanation for the appearance of the exponent $1 - \beta - 2\delta$ in (10) will be given at the end of Section 3.2.

The exponential rate in (6) has been obtained in [EPW09] by a rescaling of the diffusion approximation of the so called *relaxed Poisson profile approximation*. We will explain the
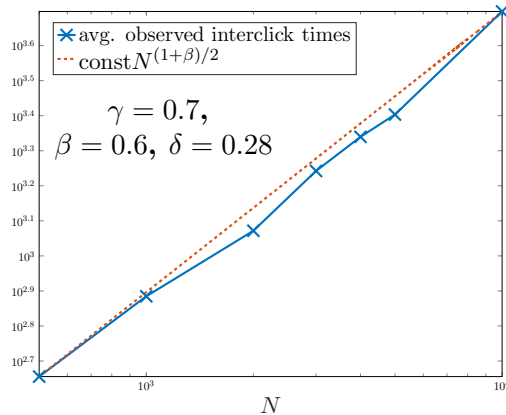
FIGURE 2. This is an illustration of the prediction of the *Main Result* (MR), which predicts the order of magnitude of the interclick times of the tournament ratchet. Each data point was obtained by pooling the first 4 interclick times from 100 simulations of the tournament ratchet for the corresponding value of N. Here, $(\beta, \delta) = (0.6, 0.28)$, which belongs to the polynomial regime $\mathcal{P}$. Even for N as small as 10^3 the asymptotics predicted by (MR) are already observable.

idea of this approximation, and define it in the context of the Moran model, in Section 3.4. There we will also give an intuitive explanation why the exponent $1 - \beta - \delta$ appears in (6).

Similar as [EPW09], the papers [PSW12, AP13, MPV20] used a diffusion approximation for the classical ratchet and modications thereof. Metzger and Eule [ME13] consider, as a proxy to the classical ratchet, a two type Moran model with selective advantage s of class 0 over class 1 and mutation rate m from class 0 to class 1. Their formula (8) corresponds to our formula 7 but their approximations for the classical ratchet concentrate on a regime in which θ remains bounded (see the discussion around [ME13, (23)]), whereas we focus here on a regime in which θ diverges logarithmically with N.

In [GSW23] it was discovered that the tournament ratchet has a dual which consists of a hierarchy of competing logistic processes. The main results of [GSW23] (on the click rate of the tournament ratchet and its type frequency profile between clicks) were obtained for the so-called subcritical regime (see Sec. 2.2) and were proved there via duality, with the help of recent results on logistic processes (see in particular [Lam05, CCM16]). This "backward in time" view, which comes on top of an Ancestral Selection Graph decorated with mutation events, will be briefly explained in Section 4.1, since it opens a route for proving the above stated result (MR) and for analysing the type frequency profile of the tournament ratchet also in the near-critical regime. In Section 4.4 we will sketch central ingredients of a proof that the type frequency profile of the tournament ratchet "between clicks" is (at least in the exponential regime of the (β, δ)-scaling) asymptotically close to the deterministic equilibrium profile, some of whose properties will be explored in Sec. 4.3.

2. Muller's ratchet as a Moran process with mutation and selection

2.1. **Model and basic concepts.** In the Moran version of Muller's ratchet, neutral resampling within any ordered pair of individuals happens at rate $\frac{1}{2N}$, and mutation from κ to $\kappa + 1$ takes place at rate m/N along each individual lineage. Selective reproduction for an individual of type κ happens at rate $\frac{1}{N} \sum_j \Phi(\kappa' - \kappa)$, where the sum is taken over all those individuals j whose type κ' is larger (and therefore "worse") than κ. For the classical case of *proportional selection*, one has $\Phi(\kappa' - \kappa) = \mathbf{s}(\kappa' - \kappa)$, while for the case of *(binary) tournament selection* one has $\Phi(\kappa' - \kappa) = \mathbf{s}$. In the sequel we will refer to these two Moran variants of Muller's ratchet briefly as the *classical ratchet* and the *tournament ratchet*. Both models have $(N, m, \mathbf{s})$ as their parameter triple, and in both models a crucial role is played by the *mutation-selection ratio* $\frac{m}{\mathbf{s}}$. In this section we reserve the symbol $\mathbf{s}$ for the selection parameter. Later, this will be specified as different parameters s and $\mathfrak{s}$ for the tournament and the classical ratchet, respectively. The following definition gives the rates for the type frequencies of the two ratchets.

Definition 2.1.

a) *Writing N_κ for the current number of individuals of type κ, the jump rates are specified as follows:*

  - *Resampling: for $\kappa \neq \kappa'$,*
    $(N_\kappa, N_{\kappa'})$ jumps to $(N_\kappa + 1, N_{\kappa'} - 1)$ at rate $\frac{1}{2N} N_\kappa N_{\kappa'}$

  - *Mutation: for κ,*
    $(N_\kappa, N_{\kappa+1})$ jumps to $(N_\kappa - 1, N_{\kappa+1} + 1)$ at rate $m N_\kappa$

  - *Selection: for $\kappa < \kappa'$,*

  $(N_\kappa, N_{\kappa'})$ jumps to $(N_\kappa + 1, N_{\kappa'} - 1)$ at rate $\begin{cases} \frac{\mathbf{s}}{N} N_\kappa N_{\kappa'}(\kappa' - \kappa) & \text{for the classical ratchet} \\ \frac{\mathbf{s}}{N} N_\kappa N_{\kappa'} & \text{for the tournament ratchet} \end{cases}$

b) *The* currently best type *is*

$$K^*(t) := \min\left\{ \kappa \in \mathbb{N}_0 : N_\kappa(t) > 0 \right\}.$$

c) *The* click times of the ratchet *are the jump times of K^*, i.e. the times at which the currently best type is lost from the population. The* type frequency profile seen from the currently best type *has the (random) weights*

$$X_k(t) := \frac{1}{N} N_{K^*(t)+k}(t), \qquad k = 0, 1, 2 \dots \tag{11}$$

In the limit $N \to \infty$ with $m, \mathbf{s}$ remaining constant, the stationary (non-random) type frequency profile $(p_k)_{k \in \mathbb{N}_0}$ is given by the mutation-selection equilibrium conditions

$$m(p_{k-1} - p_k) = \mathbf{s}\, p_k \left(\sum_{k' \in \mathbb{N}_0} \Phi(k' - k) \right), \quad k = 0, 1, 2 \dots, \tag{12}$$

where we put $p_{-1} := 0$, $\Phi(0) = 0$ and $\Phi(-d) = -\Phi(d)$ for $d \in \mathbb{N}$.

For the classical ratchet, (12) turns into

$$m(p_{k-1} - p_k) = \mathbf{s}\, p_k(\mu - k), \quad k = 0, 1, 2\ldots, \tag{13}$$

where $\mu := \sum_{k'} k p_k$ is the first moment of the profile. As already noticed by John Haigh ([Hai78]), this is solved by the *Poisson weights* with first moment $\mu = \frac{m}{\mathbf{s}}$. Indeed, this is the unique solution of (13) under the condition $p_0 > 0$.

For the tournament ratchet, (12) turns into

$$m\, (p_k - p_{k-1}) = \mathbf{s}\, p_k \left(\sum_{k' \in \mathbb{N}_0} p_{k'} \left(\mathbf{1}_{\{k' > k\}} - \mathbf{1}_{\{k' < k\}} \right) \right), \quad k = 0, 1, 2\ldots \tag{14}$$

Here the condition $p_0 > 0$ leads to the requirement $m < \mathbf{s}$ and yields $p_0 = 1 - \frac{m}{\mathbf{s}}$. Various properties of the solution $(p_{k'})$ of (14) are stated in [GSW23] Theorem 2.4. The r.h.s. of (14) equals

$$\mathbf{s}\, p_k \left(1 - p_k - 2 \sum_{k'=0}^{k-1} p_{k'} \right), \quad k = 0, 1, 2\ldots \tag{15}$$

A formal analogy between (13) and (14) results because (15) is close to $2\mathbf{s}\, p_k(\frac{1}{2} - g(k))$, where g is the cumulative distribution function of $(p_{k'})$. In this sense the role played by the profile's first moment in (13) is taken by the profile's median in (14).

2.2. The subcritical regime of the tournament ratchet.

We now report briefly on the main results of the recent paper [GSW23]. The parameters of the tournament ratchet will be denoted by (m, s) and its mutation-selection ratio by $\rho := \frac{m}{s}$. In [GSW23], as $N \to \infty$, the mutation-selection ratio $\rho = \frac{m}{s}$ is kept constant and smaller than 1, and it is assumed that $m \to 0$ and $mN \to \infty$. (For technical reasons, mN is assumed to be of larger order of $\log\log N$, which keeps the regime slightly away from that of weak mutation, in which mN would be of order one as $N \to \infty$.) We will refer to this regime as the *subcritical regime* of the tournament ratchet. The main results of [GSW23] are

Theorem 2.2. *In the subcritical regime the click rate of the tournament ratchet on the $\frac{1}{m}$-timescale is, as $N \to \infty$, logarithmically equivalent to*

$$e^{2Nm\left(\frac{1}{\rho} - 1 + \log\rho\right)}. \tag{16}$$

Theorem 2.3. *In the subcritical regime and for N large, the empirical type frequency profile at generic time points between clicks of the tournament ratchet is with high probability close to the mutation-selection equilibrium system (14) (p_k) with $p_0 = 1 - \rho$.*

These two theorems are proved in [GSW23] via a hierarchical duality. We will explain this approach in Section 4, and give an outlook how it can be extended to the near-critical regime of the tournament ratchet described in Section 1.

3. A SYNOPSIS OF THE CLASSICAL AND THE TOURNAMENT RATCHET

3.1. The dynamics of the best classes.

For $k = 0, 1, \ldots$ let $Y_k^{\mathrm{C}}(t) = N_{K*+k}^{\mathrm{C}}(t)$ and $Y_k^{\mathrm{T}}(t) = N_{K*+k}^{\mathrm{T}}(t)$ be the sizes of the $(k+1)^{\mathrm{st}}$-best class of the classical and the tournament ratchet, where $(N_\kappa^{\mathrm{C}})_{\kappa \in \mathbb{N}_0}$ and $(N_\kappa^{\mathrm{T}})_{\kappa \in \mathbb{N}_0}$ follow the dynamics specified in Definition 2.1. Here we assume that the mutation rate m is equal for both ratchets, but the selection coefficients are different:

$$\mathbf{s} = \begin{cases} \frac{m}{\theta} =: \mathfrak{s} & \text{for the classical ratchet} \\ \frac{m}{\rho} =: s & \text{for the tournament ratchet.} \end{cases}$$

The jump rates from n to $n-1$ are given for both Y_0^{C} and Y_0^{T} by

$$n \left(\frac{1}{2} \left(1 - \frac{n}{N} \right) + m \right), \tag{17}$$

but the jump rates from n to $n+1$ are different: those of Y_0^{T} are

$$n \left(\frac{1}{2} \left(1 - \frac{n}{N} \right) + s \left(1 - \frac{n}{N} \right) \right), \tag{18}$$

while those of Y_0^{C} are

$$n \left(\frac{1}{2} \left(1 - \frac{n}{N} \right) + \mathfrak{s} \sum_{k=1}^{\infty} k X_k \right). \tag{19}$$

where $(X_k(t))_{k \in \mathbb{N}_0}$ is the type frequency profile as defined in (11), with (N_κ^{C}) in place of (N_κ). Writing

$$M(t) := \sum_{k=1}^{\infty} k X_k(t)$$

for the first moment of the type frequency profile (X_k), (19) takes the form

$$n \left(\frac{1}{2} \left(1 - \frac{n}{N} \right) + m \frac{M}{\theta} \right). \tag{20}$$

An inspection of the jump rates in Definition 2.1 reveals that for each $k \in \mathbb{N}$ the process $(Y_0^{\mathrm{T}}, \ldots, Y_k^{\mathrm{T}})$ obeys an autonomous dynamics; for $k = 0$ this is evident from (17) and (18). For later reference we note here that (Y_0^T, Y_1^T) has, as $N \to \infty$, the center of attraction

$$(\mathfrak{a}, \mathfrak{b}) \sim (N(1 - \rho), N\sqrt{1 - \rho}) \tag{21}$$

provided $Nm \to \infty$ and $\rho \to 1$.

In contrast to the tournament ratchet, the rates (20) depend not only on the size of the best class but also on the profile $(X_k(t))_{k \geq 0}$ (via its first moment $M(t)$). There are various ways to predict $M(t)$ on the basis of $Y^{\mathrm{C}}(t)$, and thereby to replace (20) by a rate which is autonomous; one of these will be described in the next subsection. As conjectured already by John Haigh [Hai78], such a strategy should work in a regime in which the

expected time to extinction of Y_0^C typically takes long compared to the relaxation time of the noiseless classical ratchet. The latter is $\frac{\log\theta}{\mathfrak{s}}$ (see Remark 4.3 [EPW09]), whereas the time to extinction of a neutral Moran(N)-process starting in $Ne^{-\theta}$ is of the order $Ne^{-\theta}$. Hence this amounts to the parameter regime

$$Ne^{-\theta} \gg \frac{1}{\mathfrak{s}},$$

which in the (β, δ)-scaling (2) just means that $(\beta, \delta) \in \Delta$.

3.2. The Poisson profile approximation for the classical ratchet.
A first idea is to think of the profile $(X_k)_{k\geq 1}$ as (nearly) proportional to the *Poisson profile*

$$\pi_k = e^{-\theta}\frac{\theta^k}{k!}, \quad k \geq 1,$$

and as the mass $\pi_0 - X_0$ being distributed proportionally upon this profile. This leads to the so-called *Poisson profile approximation* of $(X_k)_{k\geq 1}$ based on X_0, given by

$$\Pi(X_0) := \left(X_0, \frac{1 - X_0}{1 - \pi_0}(\pi_1, \pi_2, \dots) \right). \tag{22}$$

(cf [EPW09, (2.5)]). The first moment of $\Pi(X_0)$ is

$$M(X_0) := (1 - X_0)\frac{\theta}{1 - \pi_0}, \tag{23}$$

in accordance with [EPW09, (5.3a)]. Plugging this into (20) in place of M leads to the following *Poisson profile approximation* of the upward jump rates (20):

$$n\left(\frac{1}{2}\left(1 - \frac{n}{N}\right) + \frac{m}{1 - e^{-\theta}}\left(1 - \frac{n}{N}\right) \right). \tag{24}$$

We denote the birth-and death-process on $\mathbb{N}_0$ with downward jump rates (17) and upward jump rates (24) by Y_{PPA}; this process can be seen as an approximation of Y_0^C.

3.3. A correspondence between the classical and the tournament ratchet.
A crucial observation is that the upward jump rates (24) and (18) are equal if and only if $m = s(1 - e^{-\theta})$, which is equivalent to the "dictionary" (7).

Remark 3.1.   a) With (7), the jump rates (17) and (18) of the size of the best class of the (m, s)-tournament ratchet are *equal* to the jump rates (17) and (24) of the Poisson profile approximation for the size of the best class of the classical $(m, \mathfrak{s})$-ratchet. Thus the "dictionary" (7) gives a 1-1 correspondence between the (β, δ)-scaling for the tournament ratchet defined in (8) and the (β, δ)-scaling for the classical ratchet defined in (2), with the dynamics of the size of the best class being preserved under this correspondence.

b) With $\widehat{\pi}_0 := \frac{\pi_0}{1-\pi_0}$, (24) takes the form

$$n\left(\frac{1}{2}\left(1-\frac{n}{N}\right) + m + \frac{m}{N}\left(N\widehat{\pi}_0 - n(1-\widehat{\pi}_0)\right)\right). \tag{25}$$

Hence in the (β,δ)-scaling (2), (25) is asymptotically equal (as $N \to \infty$) to

$$n\left(\frac{1}{2}\left(1-\frac{n}{N}\right) + m - \mathfrak{c}_0\left(n-\mathfrak{a}\right)\right) \tag{26}$$

with

$$\mathfrak{a} := N^{1-\delta}, \quad \mathfrak{c}_0 := N^{-1-\beta}. \tag{27}$$

Then, as long as $n \ll N$, the dynamics (17)&(26) is close to that of a logistic branching process with standard binary branching term, and parameters $\mathfrak{a}$ and $\mathfrak{c}_0$ given by (27). For $(\beta,\delta) \in \Delta$ the center of attraction $\mathfrak{a}$ becomes large and the "mean reversion" coefficient $\mathfrak{c}_0$ becomes small as $N \to \infty$. Here is a quick intuitive argument why $\gamma = \frac{1}{2}$ marks the boundary between the exponential and the polynomial regime for the extinction time of both Y^C and Y^T (which, as we have seen, have the dynamics (17)&(26) in common). For $\varepsilon\mathfrak{a} < n < (1-\varepsilon)\mathfrak{a}$, we roughly see a slightly supercritical binary branching Galton-Watson process with supercriticality of the order $N^{-1-\beta}N^{1-\delta}$. When starting with the order of $\mathfrak{a} = N^{1-\delta}$ individuals, the probability for "long time survival" is high if and only if $N^{1-\delta}N^{-\beta-\delta}$ is large, i.e. $\gamma < \frac{1}{2}$. More specifically, for $\gamma < 1/2$ the extinction probability "in one quick go" is

$$\asymp \left(1 - N^{-\beta-\delta}\right)^{N^{1-\delta}} \sim e^{-N^{1-\beta-2\delta}}, \tag{28}$$

which explains (10).

3.4. The relaxed Poisson profile approximation for the classical ratchet.

The Poisson profile approximation of the classical ratchet (which we discussed in Section 3.2) predicts that immediately before the time of a click the profile should (at least concerning its first moment) have features of

$$\Pi(0) = \tfrac{1}{1-\pi_0}\left(0, \pi_1, \pi_2, \dots\right).$$

A simple "shift by 1" would turn $\Pi(0)$ immediately after the click into

$$\widetilde{\pi} := \tfrac{1}{1-\pi_0}\left(\pi_1, \pi_2, \dots\right).$$

However, depending on the parameter constellation, a "relaxation" towards the equilibrium profile π may have started already before the click. Not least to provide a systematic framework for previous approaches ([SCS93, GC00] to the approximation of the size of the ratchet's best class, [EPW09] suggested to replace the PPA $\Pi(x_0)$ by the so-called

relaxed Poisson profile approximation $R_A(x_0)$. This is based on the ratchet's deterministic mutation-selection dynamics

$$\mathrm{d}x_k(t) = \left(\mathfrak{s} \sum_\ell x_\ell(\ell - k) + m(x_{k-1}(t) - x_k(t)) \right) \mathrm{d}t, \qquad k = 0, 1, \dots \tag{29}$$

(with $x_{-1} \equiv 0$) and on its *relaxation time*

$$\tau := \frac{\log \theta}{\mathfrak{s}}$$

which turns out to be the time for which $x_0(\tau) = \frac{e}{e-1}\pi_0 \approx 1.6\pi_0$ provided $x(0) = \widetilde{\pi}$, see [EPW09, Remark 4.3]. For $A > 0$ and $x_0 \in (0, 1)$, $R_A(x_0)$ is defined as the profile which has evolved over time $A\tau$ according to the dynamics (29), starting from some $\Pi(x^*)$, where $x^* \in (0, 1)$ is chosen such that the first component of $R_A(x_0)$ equals x_0. It is proved in [EPW09, Proposition 4.4] that the first moment of $R_A(x_0)$ is

$$\widetilde{M}_A(x_0) := \theta + \frac{\theta^{1-A}}{\exp(\theta^{1-A}) - 1} \left(1 - \frac{x_0}{\pi_0} \right).$$

For the choice $A = 0$ this renders the PPA prediction (22), and for the choice $A = 1$ we obtain

$$\widetilde{M}_1(x_0) = \theta + \frac{1}{e-1} \left(1 - \frac{x_0}{\pi_0} \right) = \theta + 0.58 \left(1 - \frac{x_0}{\pi_0} \right), \tag{30}$$

which is [EPW09, (5.3b)]. See Figure 3 for a comparison of the quality of the PPA $\tilde{M}_0$ and the RPPA(1) $\tilde{M}_1$ prediction of M in the light of simulations. Let us note that, according to the RPPA rule, (30) is applicable only for those x_0 that appear as first component of the result of the deterministic dynamics started from some $\Pi(x^*)$ and run over time τ. The maximal such x_0 results for $x^* := 1$, which according to [EPW09, (4.11)] leads to $x_0 = \pi_0 \cdot e$ and to $\widetilde{M}_1(x_0) = \theta - 1$.

Plugging (30) into (20) gives the RPPA(1) upward jump rates

$$n \left(\frac{1}{2} \left(1 - \frac{n}{N} \right) + m - \mathfrak{c}_1 \left(n - \mathfrak{a} \right) \right). \tag{31}$$

with

$$\mathfrak{a} := N\pi_0, \quad \mathfrak{c}_1 := \frac{0.58 \, m}{N\pi_0\theta}. \tag{32}$$

Like those of PPA, the downward jump rates of RPPA(1) are given by (17), which is (31) without the term $\mathfrak{c}_1 (n - \mathfrak{a})$. In the (β, δ)-scaling (2), (32) turns into

$$\mathfrak{a} = N^{1-\delta}, \quad \mathfrak{c}_1 = \frac{0.58}{\delta} \frac{N^{-\beta+\delta-1}}{\log N}.$$

We thus see a close analogy between the dynamics (17)&(26) and (17)&(31). However the mean reversion coefficients $\mathfrak{c}_1$ and $\mathfrak{c}_0$ differ by the polynomial order N^δ. Replacing $-\beta$ by $-\beta + \delta$ turns the r.h.s. of (28) from $\exp(-N^{1-\beta-2\delta})$ into $\exp(-N^{1-\beta-\delta})$. This shows that the boundary between the polynomial and the exponential regime for the extinction time

of RPPA(1) is marked by $\gamma = 1$, and makes it clear also in the Moran framework that the clicking rate predicted by (6) is based on RPPA(1) (rather than on PPA).
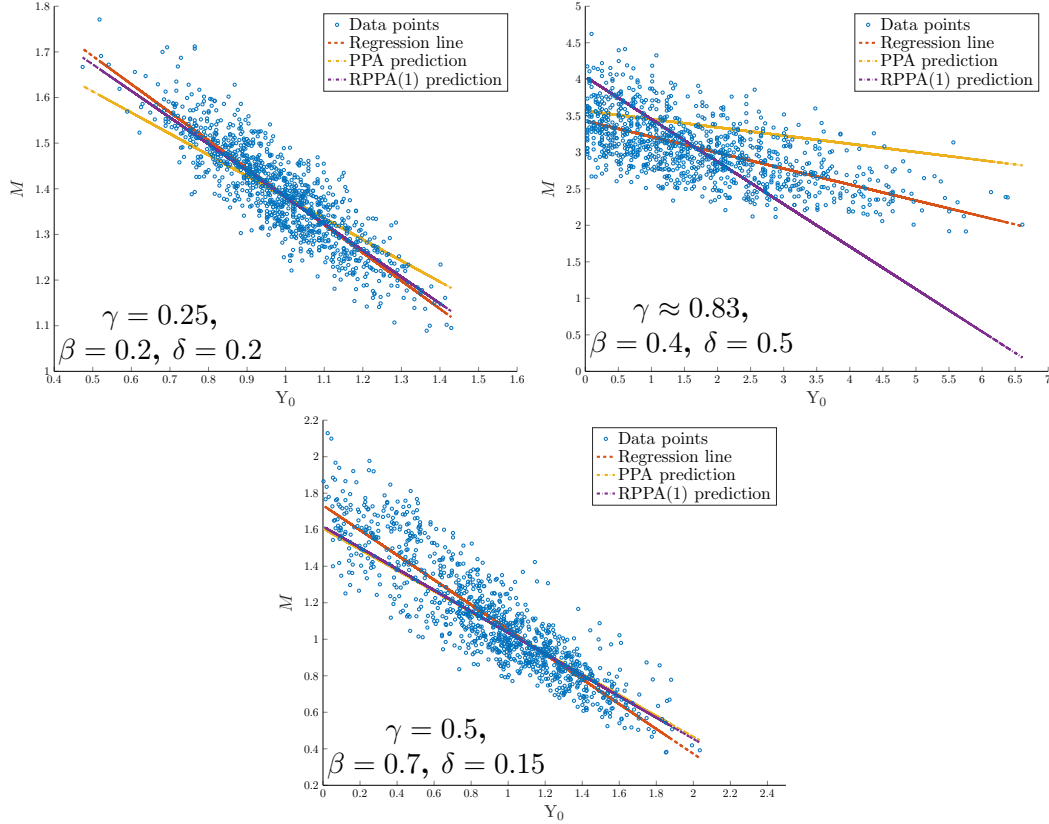

FIGURE 3. This figure is inspired by [EPW09, Figure 6]. It illustrates the quality of the PPA prediction (22) and of the RPPA(1) prediction (30) of the profile's first moment M based on the size Y_0 of the best class. For $N = 1.000$ we took snapshots of (Y_0, M) at times $t = 10k$, $k = 1, 2, \ldots$. The plots diplay the estimated regression line with the predictions given by PPA and RPPA(1). The fit of PPA is worse than that of RPPA(1) for small γ and becomes better for larger γ. See also [EPW09] Fig. 6 for analogous simulations based on the Wright-Fisher model.

3.5. **On the expected time to extinction of the best class in the (β, δ)-scaling.** In this subsection we focus on the birth-and-death process $Y := Y_0^{\mathrm{T}}$ with jump rates (17) and (18), and denote its extinction time by T_0. Let us emphasise again that this process has the same dynamics as the process Y_{PPA} defined in Section 3.2, provided the mutation rates are equal and the selection coefficients are translated through the "dictionary" (7).

The proof of the following theorem, which will be given in Section 5, is based on a careful analysis of the Green function associated with the dynamics (17)&(18).

Theorem 3.2. *Consider the (β, δ)-scaling (8). Then, as $N \to \infty$,*

a) *For $(\beta, \delta) \in \mathcal{P}$,*

$$\mathbb{E}_{N^\alpha}[T_0] \quad \sim \quad 2\left(\frac{1+\beta}{2} - \alpha\right)(\log N)\, N^\alpha, \qquad \alpha < \frac{1+\beta}{2}, \tag{33}$$

$$\mathbb{E}_{N^\alpha}[T_0] \quad \sim \quad \frac{\pi^{3/2}}{2} N^{\frac{1+\beta}{2}}, \qquad\qquad \alpha \geq \frac{1+\beta}{2}. \tag{34}$$

b) *For $(\beta, \delta) \in \mathcal{E}$ and every sequence $n_0 = n_0^{(N)}$ such that $n_0 \geq N^{1-\delta}(1 - N^{-\delta})^{-1}$,*

$$\mathbb{E}_{n_0}[T_0] \quad \asymp \quad \exp\left(N^{1-\beta-2\delta}\right), \tag{35}$$

where $\asymp$ denotes logarithmic equivalence (as defined after (10)).

Remark 3.3.   a) While (35) cannot be deduced directly from Theorem 2.2, the asymptotics of (16) relates nicely to that of (35). Indeed, the exponent in (35) equals $Nm(1-\rho)^2$, while the exponent in (16) is $2Nmh(\rho)$ with

$$h(\rho) = \frac{1}{\rho} - 1 + \log\rho \sim -\frac{1}{2}(1-\rho)^2 \quad \text{as } \rho \to 1.$$

b) In the light of Remark 3.1.a), Theorem 3.2 is relevant not only for the tournament ratchet, but also for the Poisson profile approximation of the classical ratchet. Prominent starting values for Y are

   – with regard to the tournament ratchet: $n_0^T := N^{1-\delta/2}$, which according to (21) is the asymptotic center of attraction of the size of its second best class,

   – with regard to the classical ratchet: $n_0^C := N\pi_1 = N\theta e^{-\theta}$, which in the (β, δ)-scaling equals $N^{1-\delta}\delta \log N$. Figure 4 illustrates that this asymptotics of the starting value can indeed be seen in simulations of the classical ratchet. The starting value n_0^T is used in Figure 6, and the starting value n_0^C is used in Figures 5 and 7.

   For $(\beta, \delta) \in \mathcal{P}$ we have

$$1 - \delta < \frac{1+\beta}{2} < 1 - \frac{\delta}{2}.$$

Hence Theorem 3.2.a) gives

$$\mathbb{E}_{n_0^T}[T_0] \asymp N^{\frac{1+\beta}{2}} \qquad \text{and} \qquad \mathbb{E}_{n_0^C}[T_0] \asymp N^{1-\delta},$$
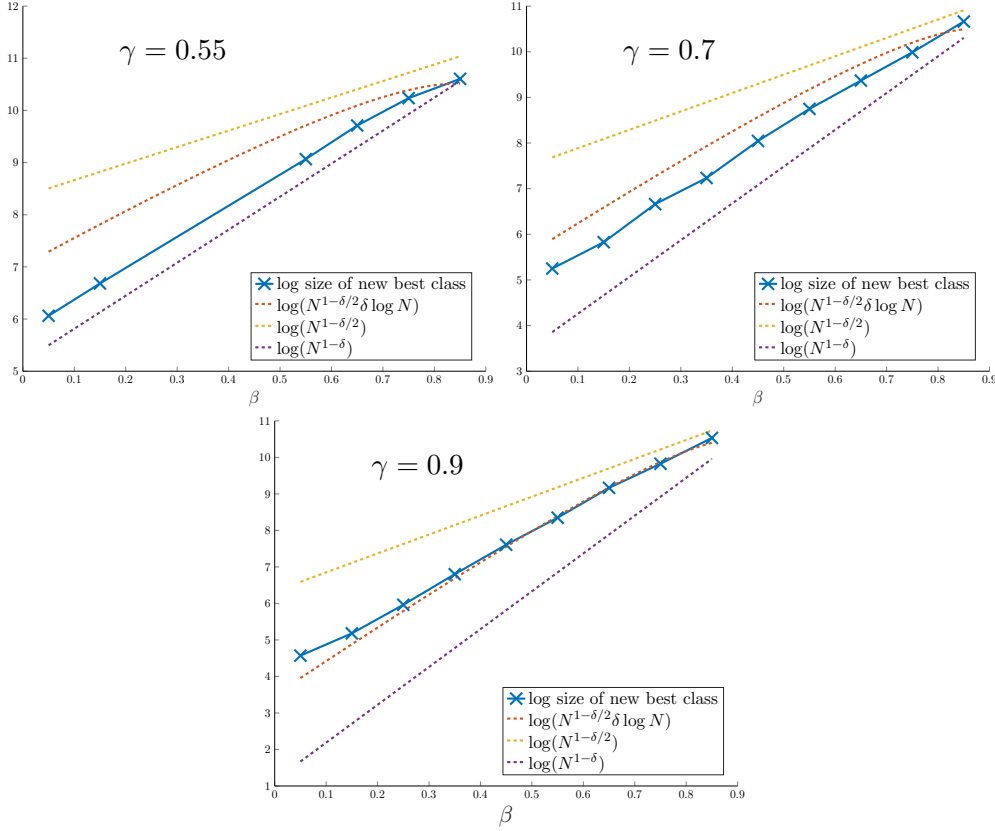
which meets (9) and (5).

FIGURE 4. For $N = 100.000$ we compare the size of the "new" best class of the classical ratchet immediately after after a click (as observed by simulations) with $N\pi_0 = N^{1-\delta}$, $N\pi_1 = N^{1-\delta}\delta \log N$ and $N^{1-\delta/2}$, which are the centers of attraction of the best and the second best class of the classical ratchet and the center of attraction of the best class of the fancy ratchet (cf. Remark 3.3. b)). For various fixed values of γ, we consider (the logarithms of) these quantities as functions of β (recall that $\delta = \gamma/(1-\beta)$). Each data point was obtained by pooling the first 4 interclick times from 20 simulations of the classical ratchet for the corresponding parameter configuration. Depending on the choice of γ the size of the new best class seems to lie between $N\pi_0$ and $N\pi_1$, but way below $N^{1-\delta/2}$.
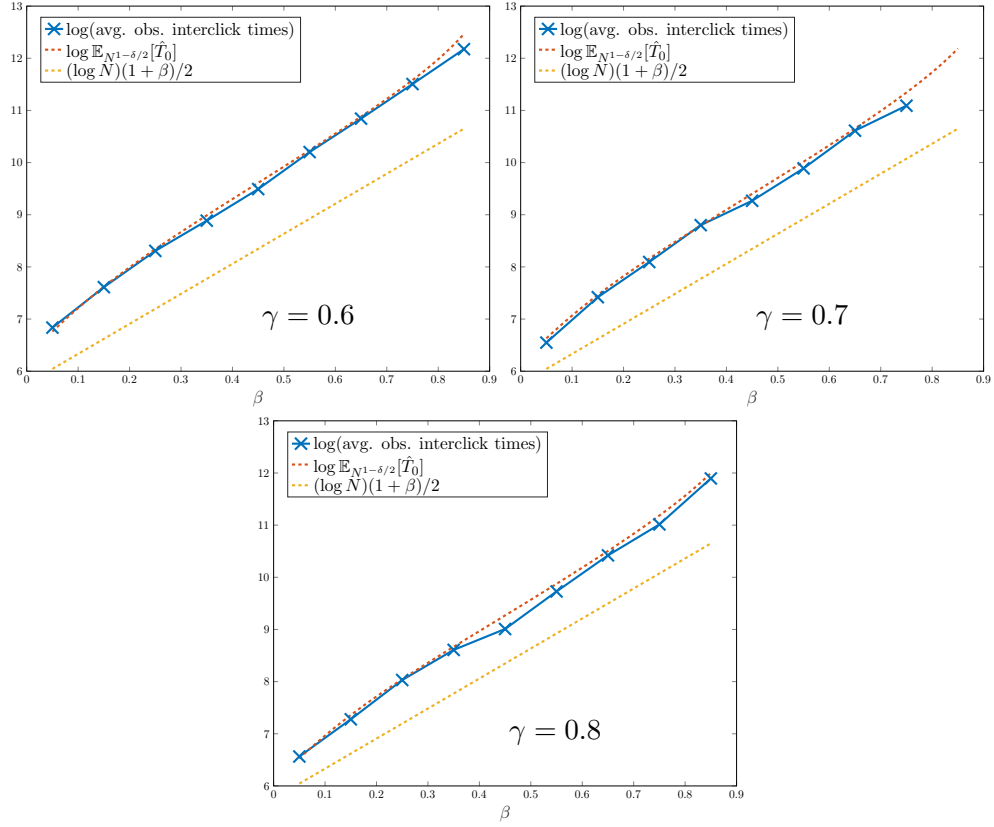
FIGURE 5. For fixed population size $N = 100.000$ the prediction for the expected interclick time of the tournament based on i) a numerical calculation of the Green function and ii) Theorem 3.2 are compared with simulations. Each data point was obtained by pooling the first 4 interclick times from 20 simulations of the tournament ratchet for the corresponding parameter configuration. Each plot shows this for one fixed value of γ with varying β.
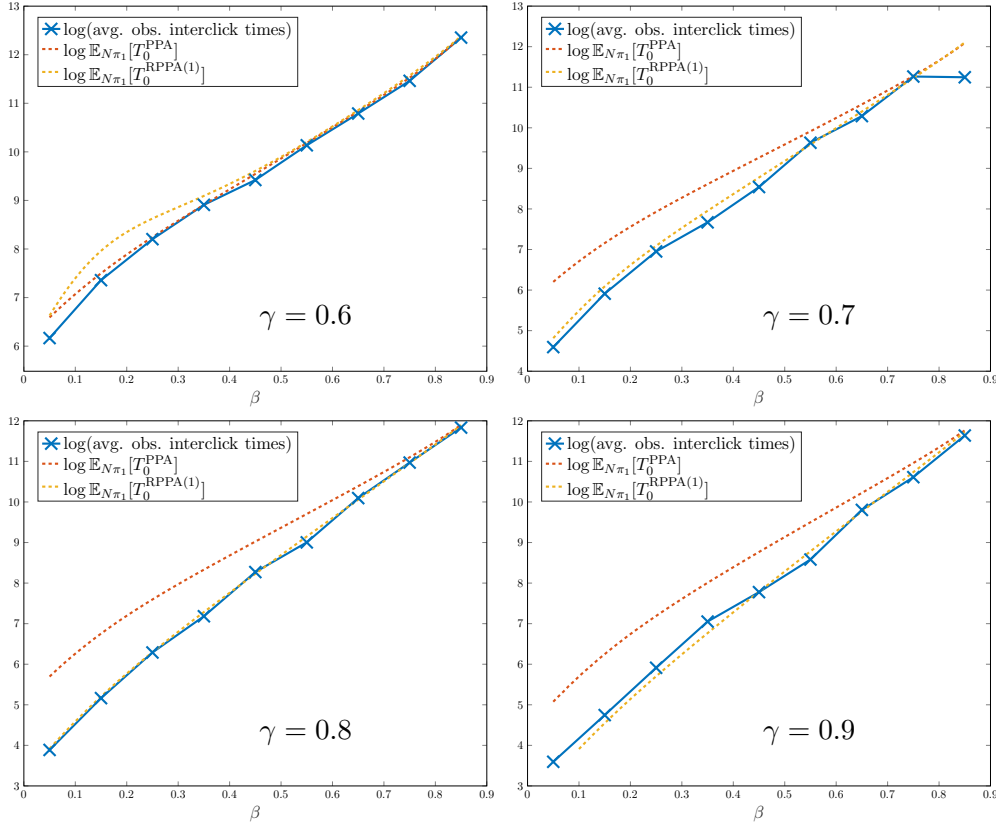
FIGURE 6. For fixed population size $N = 100.000$ the PPA as well as the RPPA(1) prediction for the expected interclick time of the classical ratchet are checked via simulations for various values of β and γ. Each data point was obtained by pooling the first 4 interclick times from 20 simulations of the classical ratchet for the corresponding parameter configuration. In agreement with the empirical findings on the size of the new best class (as reported in Figure 4) we use $N\pi_1$ as a starting value and compute the expectations of the extinction times of Y_{PPA} and of $Y_{\mathrm{RPPA}(1)}$ by a numerical evaluation of their Green functions. Each plot shows this for one fixed value of γ for varying β. In most of the observed scenarios the observed interclick times are approximated better by RPPA(1) than by PPA.

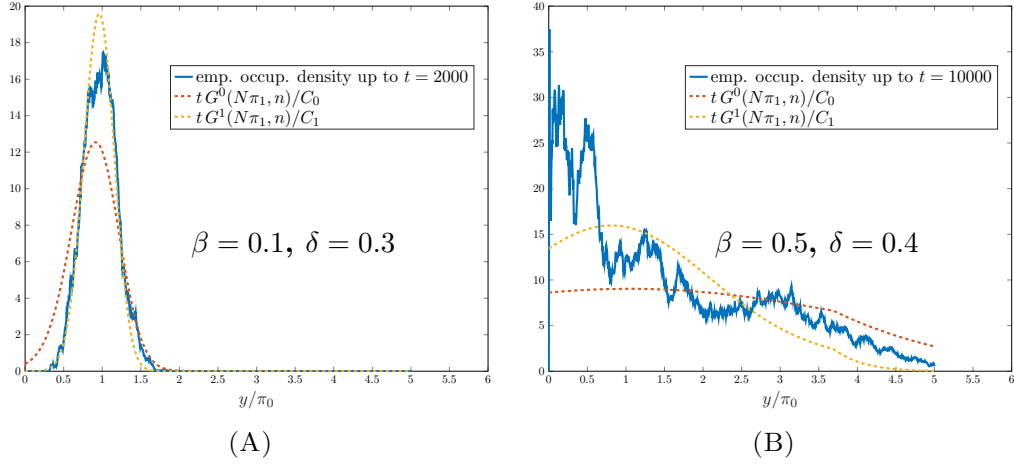(A)   (B)

FIGURE 7. This figure is inspired by [EPW09, Figure 5]. The empirical occupation density of the size of the best class in a simulation of the classical ratchet is compared to the (numerically computed) Green functions $G^0(N\pi_1, \cdot)$ (of Y_{PPA}) and $G^1(N\pi_1, \cdot)$ (of $Y_{\mathrm{RPPA}(1)}$). The norming constants are $C_i := \sum_{n \geq 1} G^i(N\pi_1, n)$, $i = 0, 1$. The choice of the starting value $N\pi_1$ is in agreement with the findings reported in Figure 4. Panels (A) and (B) feature the exponential and the polynomial regime, respectively, with $\gamma = 0.33$ in panel (A) and $\gamma = 0.8$ in panel (B). In panel (A) the population size is $N = 2000$ and simulations were run up to (generation) time $t = 2.000$. In panel (B) the population size is $N = 10.000$ and simulations were run up to $t = 10.000$. In (A) no click time was observed till time $t = 2000$, and in (B) 13 click times occurred till time $t = 10000$. Panel (B) may give the impression of a certain clustering of the process of click times. This may go along with a conjecture stated by Mariani, Pardoux and Velleret in [MPV20] on the transitory regime where the "relaxation time" starts to become larger than the interclick times. Here is a quote from the introduction of [MPV20]: *"If [the relaxation time] t_R is of the same order as [the expected click time] t_C or larger, we a priori can not exclude that trains of short interdependent intervals could alter this observed distribution of interval length. But already if t_R is of the same order as t_C, there shall still be long realizations of inter-click intervals after which we can say that the dependence in the past is forgotten."*

4. The tournament ratchet near criticality

4.1. **A backward-in-time view.** We briefly review some essentials of the graphical approach developed in [GSW23]. This kind of approach dates back to the pioneering work of Krone and Neuhauser [KN97, NK97] and has been intensely used in mathematical population genetics since then; for some recent related work see for instance [GCS18, CS20, CHS22]). For fixed parameters N, m, s this approach gives a representation both of the Moran tournament ratchet as well as of its hierarchical dual, and it is instrumental for [GSW23]'s proof of Theorems 2.2 and 2.3. Some of the ideas of this proof are summarized in Figure 1 of [GSW23].

The line counting process of the Ancestral Selection Graph (ASG) is a birth-and death process on $\mathbb{N}$ with upward jump rate $sn(N - n)$ and downward jump rate $\frac{n(n-1)}{2N}$. The lineages are decorated by a Poisson point process with intensity m; the number of points on a (potential ancestral) lineage is the *(mutational) load* accumulated along that lineage. The *load of a potential ancestor $\mathcal{I}'$* of an individual $\mathcal{I}$ is the mimimum of the loads of the lineages connecting $\mathcal{I}$ and $\mathcal{I}'$. If all the individuals at time 0 are assumed to be of equal type, then the type of an individual sampled at time t is the minimum of the mutational loads carried by the ancestral lineages that lead back from this individual to time 0 (cf. [GSW23, Remark 4.2]).

We now consider times t, u with $\frac{1}{s} \ll t \ll u$. For $k \geq 0$, the *minimum+k load ASG (back from the total population at time u)*, denoted by $\bar{\mathcal{A}}_k^u =: \bar{\mathcal{A}}_k$, consists at any time $t' \leq u$ of all those individuals living at time t' that are a load $(k + k_{\min})$-potential ancestor of some individual living at time u, where $k_{\min}$ is the minimum of all the loads of lineages connecting times t' and u. We say that t' is a *backward click time along $\bar{\mathcal{A}}_0$* if t' is the time of a jump of $k_{\min}$.

For an individual $\mathcal{I}$ sampled at time t, and $k \geq 0$, the graph made up by its load k potential ancestors is the *load k ASG of $\mathcal{I}$*, denoted by $\mathcal{A}_k = (\mathcal{A}_k(t - r))_{r \geq 0}$. We write $A_k(t-r) := \#\mathcal{A}_k(t-r)$, and $\bar{A}_k(t-r) := \#\bar{\mathcal{A}}_k(t-r)$, $r \geq 0$, for the line counting processes of $\mathcal{A}_k$ and $\bar{\mathcal{A}}_k$.

Key insights of [GSW23] are that (in the parameter regime considered there)

(i) the process of backward click times along $\bar{\mathcal{A}}_0$ is asymptotically close to (and a fortiori has asymptotically the same rate as) the process of click times of the tournament ratchet, and

(ii) the type of an individual sampled at time t results with high probability as $N \to \infty$ from its potential ancestry within $\bar{\mathcal{A}}_0$.

[GSW23, Lemma 5.1] tells that both $A_0(t-r)$ and $\bar{A}_0(t-r)$, $r \geq 0$, are birth-and-death processes whose jump rates are given by

$$
\begin{aligned}
ns\left(1 - \frac{n}{N}\right) \quad &\text{from } n \text{ to } n+1, \\
n\left(m + \frac{n-1}{2N}\right) \quad &\text{from } n \text{ to } n-1.
\end{aligned}
\tag{36}
$$

More generally, for any $k \in \mathbb{N}_0$, both the processes $(A_0, \ldots, A_k)$ and $(\bar{A}_0, \ldots, \bar{A}_k)$ have the Markovian multivariate birth-and death dynamics specified in [GSW23] Lemma 5.1. Let $(Z_0, Z_1, \ldots)$ be a process whose first $k+1$ components have the just mentioned dynmaics. It is not difficult to check that, asymptotically as $N \to \infty$, the process (Z_0, Z_1) has center of attraction

$$(a, b) \sim \left(2Ns(1-\rho), 2Ns\sqrt{(1-\rho)} \right), \tag{37}$$

provided that $m \to 0$, $Nm \to \infty$ and $\rho \to 1$. Notably, this is (21) up to the factor $2s$. In the (β, δ)-scaling (37) turns into

$$(a, b) \sim (2N^{1-\beta-\delta}, 2N^{1-\beta-\delta/2}).$$

Starting the process $(Z_0, Z_1, \ldots)$ from $(N, 0, 0, \ldots)$ leads to a process which is similar in structure to $(Y_0, Y_1, \ldots)$ started in $(N, 0, 0, \ldots)$: every once in a while, the minimal index $\hat{k}$ for which Z_k is different from 0 jumps to $\hat{k}+1$. This process of jump times has the same distribution as the process of backward click times mentioned above, and in the subcritical regime was shown to be asymptotically Poisson as $N \to \infty$. For two pairs (β, δ) in the near-critical regime, Figure 8 shows simulations of the asymptotics of backward interclick times compared with that of the (forward) interclick times.

4.2. **A basic duality relation.** For fixed N, let $Y_0(t)$ be the number of type 0 individuals at time t. Assume that at time 0 all individuals are of type 0. The event $\{Y_0(t) > 0\}$ then occurs if and only if among the individuals living at time 0 there is at least one load zero potential ancestor of some individual living at time t. Consequently, for Z_0 being a birth-and-death process with jump rates (36), we have

$$\mathbb{P}[Y_0(t) > 0 \mid Y_0(0) = N] = \mathbb{P}[Z_0(t) > 0 \mid Z_0(0) = N]. \tag{38}$$

(More general than (38), [GSW23, Remark 2.5] says that the process Z_0 is in hypergeometric duality with the process Y_0/N; see also [JK14] for more details on this notion.) Such duality relations are a powerful tool to study ancestries in population genetics models [Möh99, PP13, BCH18].

Let us write τ for the first click time of the tournament ratchet, and T for the extinction time of Z_0. Then (38) translates into

$$\mathbb{P}[\tau > t \mid Y_0(0) = N] = \mathbb{P}[T > t \mid Z_0(0) = N]. \tag{39}$$

Integrating (39) with respect to t from 0 to ∞ gives

$$\mathbb{E}[\tau \mid Y_0(0) = N] = \mathbb{E}[T \mid Z_0(0) = N]. \tag{40}$$

In the next subsection we will make use of the analogue of (40) that is based on the graphical approach described in Sec. 4.1.

In [GSW23] the click rate of the tournament ratchet was obtained via the process $(Z_0, Z_1, \ldots)$, defined at the end of Sec. 4.1. As soon as Z_0 gets extinct, the old Z_k becomes the new Z_{k-1}, and in particular the old Z_1 becomes the new Z_0. The following result can this be seen as a "dual analogue" of Theorem 3.2. Its proof (which we will not include in the present paper) follows similar lines as that of Theorem 3.2.

Theorem 4.1. *Denote the extinction time of Z_0 by $\widehat{T}_0$, and consider the (β, δ)-scaling. Then, as $N \to \infty$,*

a) For $(\beta, \delta) \in \mathcal{P}$,

$$\mathbb{E}_{N^\alpha}\left[\widehat{T}_0\right] \quad \sim \quad \left(\tfrac{1-\beta}{2} - \alpha\right)(\log N) N^{\alpha+\beta}, \qquad \alpha < \tfrac{1-\beta}{2}, \tag{41}$$

$$\mathbb{E}_{N^\alpha}\left[\widehat{T}_0\right] \quad \sim \quad \frac{\pi^{3/2}}{2} N^{\frac{1+\beta}{2}}, \qquad\qquad \alpha \geq \tfrac{1-\beta}{2}. \tag{42}$$

b) For $(\beta, \delta) \in \mathcal{E}$ and every sequence $n_0 = n_0^{(N)}$ such that $n_0 \geq 4N^{1-\beta-\delta}$,

$$\mathbb{E}_{n_0}\left[\widehat{T}_0\right] \quad \asymp \quad \exp\left(N^{1-\beta-2\delta}\right). \tag{43}$$
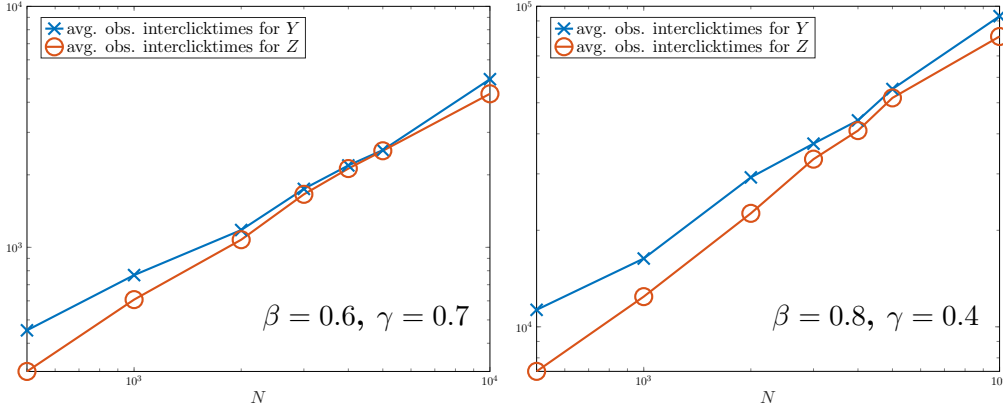

FIGURE 8. For two parameter constellations in the near-critical regime, this displays the outcomes of a simulation comparing (on logarithmic scales for N up to 10^4) a long-term average of observed (forward) interclick times with a long term average of observed backwar interclick times. The left and the right panel belong to the polynomial and to the exponential regime, respectively. The backward and the forward rates seem to have (virtually) the same asymptotics. This still leaves the conjecture that the click time processes should be asymptotically Poisson in the exponential regime, while they exhibit clusters in the polynomial regime.

4.3. The equilibrium profile in the deterministic limit.

For $0 < \rho < 1$ we denote by $\mathbf{p}^{(\rho)} = (p_k^{(\rho)})_{k \in \mathbb{N}_0}$ the solution of the recursion

$$p_1 = 0, \qquad p_0 = 1 - \rho, \qquad \rho(p_k - p_{k-1}) = p_k\left(1 - p_k - 2\sum_{k'=0}^{k-1} p_{k'}\right), \; k \geq 1. \tag{44}$$

For $0 < m < s$ and $\rho := \frac{m}{s}$, $\mathbf{p}^{(\rho)}$ is the solution of (14) (with $\mathbf{s} := s$), and hence is the equilibrium frequency profile of the noiseless (m, s)-tournament ratchet. As already observed at the end of Section 2, the recursion (44) implies

$$\rho(p_k - p_{k-1}) = p_k \left(\sum_{k'>k} p_{k'} - \sum_{k'<k} p_{k'} \right), \ k \geq 1. \tag{45}$$

Thus $p_k - p_{k-1}$ is positive as long as k is strictly smaller than the median of $\mathbf{p}^{(\rho)}$, and it becomes negative as soon as k is strictly larger that the median. This immediately leads to the following

Remark 4.2. The mode and the median of $\mathbf{p}^{(\rho)}$ have a distance of no more than 1.

It has been proved in [GSW23] that the tails $\sum_{\ell>k} p_\ell^{(\rho)}$, $k \geq 0$, are given by the k-fold iterations $\mathfrak{G}_\rho^{(k)}(u) := \mathfrak{G}_\rho(\mathfrak{G}_\rho(\dots(u)))$ of the function

$$\mathfrak{G}_\rho(u) = \frac{1}{2}\left(1 + \rho - \sqrt{(1+\rho)^2 - 4\rho u}\right), \quad 0 \leq u \leq 1. \tag{46}$$

More precisely, [GSW23, Theorem 2.4 d)] tells that

$$\sum_{\ell>k} p_\ell^{(\rho)} = \mathfrak{G}_\rho^{(k)}(\rho), \quad k \in \mathbb{N}_0. \tag{47}$$

With regard to the (β, δ)-scaling (8) we put for $N \in \mathbb{N}$ and $\delta \in (0, 1)$

$$p_\ell^{(N,\delta)} := p_\ell^{(1-N^{-\delta})}, \quad \ell \in \mathbb{N}_0. \tag{48}$$

The following result will be proved in Section 5.

Proposition 4.3. *For any fixed $k \geq 0$ we have the asymptotics as $N \to \infty$*

$$p_k^{(N,\delta)} \sim N^{-\delta/2^k}, \tag{49}$$

$$\sum_{\ell<k} p_\ell^{(N,\delta)} \sim N^{-\delta/2^{k-1}}. \tag{50}$$

The relation (50) invites to approximate the median of $\mathbf{p}^{(N,\delta)}$ by solving the equation

$$\tfrac{1}{2} = 1 - N^{-\delta/2^k}.$$

That this strategy works relies on a refinement of Proposition 4.3 which is provided by

Proposition 4.4.   *a) The asymptotics (49) and (50) are valid up to k of order $\log \log N$.*
b) The following asymptotics holds true for the median of the type frequency profile:

$$\mathrm{median}(\mathbf{p}^{(N,\delta)}) \sim \frac{\log(\log(N^\delta)/\log(2))}{\log(2)} \qquad \text{as } N \to \infty. \tag{51}$$

A proof of this proposition is given in Section 5.
Plugging (51) into (49) (which is allowed because of Proposition 4.4 part a) and recalling Remark 4.2 we arrive at

Corollary 4.5. $\max_k p_k^{(N,\delta)} \to \frac{1}{4}$ *as* $N \to \infty$.

Remark 4.6. Here is a quick argument based on the recursion (44) which shows that the weight of the median of $\mathbf{p}^{(\rho)}$ is roughly $\frac{1}{4}$ as soon as ρ is close to 1. For $\nu_k := \frac{1}{\rho}p_k$ we obtain for all $K \in \mathbb{N}$, dividing (45) by ρ^2 and summing over k between 1 and K,

$$
\begin{aligned}
0 &= \sum_{k=1}^{K}(\nu_k - \nu_{k-1}) - \sum_{k=1}^{K}\left[\nu_k\left(\sum_{j>k}\nu_j - \sum_{j<k}\nu_j\right)\right] \\
&= \nu_K - \nu_0 - \sum_{k=1}^{K}\sum_{j=0}^{\infty}\nu_k\nu_j\left[\mathbf{1}_{\{j>k\}} - \mathbf{1}_{\{j<k\}}\right] \\
&= \nu_K - \nu_0 + \nu_0\sum_{k=1}^{K}\nu_k - \sum_{k=1}^{K}\sum_{j=K+1}^{\infty}\nu_k\nu_j.
\end{aligned}
$$

For ρ close to 1 and K such that $\sum_{k=0}^{K}\nu_k \approx \frac{1}{2\rho}$ this gives (since then $\nu_0 = \frac{1-\rho}{\rho}$ becomes small)

$$
0 \approx \nu_K - \nu_0 + \nu_0\left(\tfrac{1}{2} - \nu_0\right) - \tfrac{1}{4} \approx \nu_K - \tfrac{1}{4}.
$$

4.4. The type frequency profile in the near-critical exponential regime.

In this section we give evidence that, in the near-critical exponential regime (4), the type frequency profile between clicks is with high probability close to (p_k) given by (44). We will start by briefly reviewing how this was proved in [GSW23] for the subcritical regime. Resuming the notation of Sec. 4.1 we consider for $k \in \mathbb{N}_0$ the event

$$\mathscr{E}_k := \{\text{there is a coalescence event between } \mathcal{A}_k \text{ and } \bar{\mathcal{A}}_0$$

$$\text{but no coalescence event between } \mathcal{A}_j \text{ and } \bar{\mathcal{A}}_0 \text{ for } j < k\}.$$

In [GSW23] the following was proved for the subcritical regime:

(H) *The probability that the randomly sampled individual $\mathcal{I}$ has type k is asymptotically equal to* $\mathbb{P}(\mathscr{E}_k)$.

We conjecture that (H) is valid also in the exponential regime of the (β, δ)-scaling, but defer the proof of this to future work. The link between (H) and the prediction (48) for the empirical type frequency profile is established by the following

Claim 4.7. *For $(\beta, \delta) \in \mathcal{E}$ and for each $k \in \mathbb{N}_0$,*

$$(\mathbb{P}(\mathscr{E}_0), \ldots, \mathbb{P}(\mathscr{E}_k)) \sim (p_0^{(N,\delta)}, \ldots, p_k^{(N,\delta)}) \quad \text{as } N \to \infty. \tag{52}$$

In the rest of the subsection we will outline the main mathematical insights that support this claim. Let us first consider the case $k = 0$. In the regime $\mathcal{E}$, the quantity $a = 2N^{1-\beta-\delta}$ is the typical (order of) size which the process $\bar{\mathcal{A}}_0$ attains (and keeps) "between clicks". As long as $\mathcal{A}_0$ and $\bar{\mathcal{A}}$ are disjoint, the rate of coalescence events between $\mathcal{A}_0$ and $\bar{\mathcal{A}}$ on the N^β-timescale is

$$N^\beta\frac{1}{2N}A_0(t-r)\cdot\bar{A}(t-r) \sim A_0(t-r)N^{-\delta}. \tag{53}$$

The event $\mathscr{E}_0$ will thus with high probability go along with the event

$$\widetilde{\mathscr{E}_0} := \{A_0 \text{ reaches an order of magnitude beyond } N^\delta\}.$$

Notably, the condition $(\beta, \delta) \in \mathcal{E}$, which is equivalent to $1 - \beta - \delta > \delta$, ensures that the center of attraction of A_0 is above N^δ. However, since A_0 starts in 1 and then is only slightly supercritical, the probability that A_0 makes it up to its center of attraction is small for large N. Indeed, in view of its dynamics (36), $(A_0(t-r))_{r\geq 0}$ behaves for sufficiently small r similar to a binary Galton-Watson process. That this approximation is good enough for our purposes goes along with the following

Lemma 4.8. *Let $H = (H_r)_{r\geq 0}$ be the binary Galton-Watson process with splitting rate $b := N^{-\beta}$ and death rate $d := N^{-\beta}(1 - N^{-\delta})$. Put $R := KN^{\beta+\delta}\log N$ with $K < \frac{1-\delta}{2}$. Then*

$$\frac{1}{2N} \int_0^R \mathbb{E}[H_v(H_v - 1)|H \text{ survives}] \, dv \sim \frac{1}{2N} \frac{2b}{b-d} e^{2(b-d)R} = \frac{N^{-\beta}}{N^{1-\beta-\delta}} e^{2N^{-\beta-\delta}R} = o(1).$$

Proof. This follows from the fact that the second moment of H_r conditioned on survival is

$$\mathbb{E}[H_r^2|H \text{ survives}] = e^{2(b-d)r} + \frac{b+d}{b-d}\left(e^{2(b-d)r} - e^{(b-d)r}\right).$$

$\square$

As a consequence of this lemma we obtain

Corollary 4.9. *The processes $(A_0(t - r))_{r\geq 0}$ and $(H_r)_{r\geq 0}$ can be coupled such that, as $N \to \infty$, they coincide with high probability up to times $r \ll N^{\beta+\delta}\log N$.*

With H as in Lemma 4.8, and with

$$p_H(r) := \mathbb{P}(H_r > 0)$$

we have

$$\mathbb{E}[H_r|H_r > 0] = \frac{1}{p_H(r)} \exp(N^{-\beta-\delta}r). \tag{54}$$

Let us now choose a time r_0 such that

$$N^{\beta+\delta}\log\log N \gg r_0 \gg N^{\beta+\delta}.$$

We claim that for this r_0 we have

$$N^\delta \log N \gg \frac{1}{p_H(r_0)} \exp(N^{-\beta-\delta}r_0) \gg N^\delta \tag{55}$$

and

$$p_H(r_0) \sim N^{-\delta}. \tag{56}$$

For checking (56) we recall that for a binary Galton-Watson process (G_τ) with individual birth rate 1 and individual death rate $1 - N^{-\delta}$ the generating function $F(\zeta, \tau) := \mathbb{E}[\zeta^{G_\tau}]$ fulfills

$$F(0, \tau) = \frac{N^{-\delta} - 1 + e^{-\tau N^{-\delta}}\left(1 - N^{-\delta}\right)}{-1 + e^{-\tau N^{-\delta}}\left(1 - N^{-\delta}\right)}.$$

Putting

$$\tau_0 := r_0 N^{-\beta} \tag{57}$$

we obtain

$$p_H(r_0) = \mathbb{P}(G_{\tau_0} > 0) = 1 - F(0, \tau_0),$$

which meets (56) since $\tau_0 \gg N^\delta$. Finally, (55) is immediate from (56). This proves Claim 4.7 for the case $k = 0$. Let us now turn to $k > 0$, and abbreviate

$$\widetilde{p}_k := \mathbb{P}(\mathscr{E}_k).$$

In order to argue that $(\widetilde{p}_k)$ satisfies, at least for sufficiently small k, approximately the recursion (44), we proceed similarly as in the proof of [GSW23] Proposition 11.1, and consider a standard Yule tree $\mathscr{Y}$ with Poisson decoration at rate $\rho = 1 - N^{-\delta}$. For τ_0 as in (57) we have

$$N^\delta \log \log N \gg \tau_0 \gg N^\delta.$$

Let $\mathscr{Y}^{(\tau_0)}$ be that part of $\mathscr{Y}$ which is below time τ_0 (i.e. the tree $\mathscr{Y}$ cut at height τ_0). Let $L^{(\tau_0)}$ be the minimum of the Poisson loads carried by the branches of $\mathscr{Y}^{(\tau_0)}$. By a similar argument that lead to (54) and (55) one can see that, conditional on $\{L^{(\tau_0)} = k\}$, the number of branches of $\mathscr{Y}^{(\tau_0)}$ that carry the Poisson load k is of an order (slightly) larger than N^δ. Thus the same argument we used based on (53) for the case $k = 0$ shows, now for any fixed $k \geq 0$, that conditional on $\{L^{(\tau_0)} = k\}$ the event $\mathscr{E}_k$ occurs with high probability, thus implying

$$\widetilde{p}_k \sim \mathbb{P}(L^{(\tau_0)} = k) \quad \text{as } N \to \infty. \tag{58}$$

We write $(\mathfrak{p}_k)_{k \in \mathbb{N}_0} := (\mathfrak{p}_k^{(N)})_{k \in \mathbb{N}_0}$ for the distribution of $L^{(\tau_0)}$. Let $\mathcal{T}$ be a random variable with standard exponential distribution (playing the role of the first splitting time of $\mathscr{Y}$), put

$$\mathcal{T}_0 := \min(\mathcal{T}, \tau_0)$$

and given $\mathcal{T}_0$ let M_0 be a Poisson random variable with parameter $\mathcal{T}_0(1 - N^{-\delta})$.

Decomposing $\mathscr{Y}^{(\tau_0)}$ at the time of the first branching of $\mathscr{Y}$ (provided this time is smaller than τ_0) we obtain that $L^{(\tau_0)}$ solves the stochastic fixed point equation

$$L^{(\tau_0)} \stackrel{d}{=} M_0 + \min(L_1^{(\tau_0 - \mathcal{T}_0)}, L_2^{(\tau_0 - \mathcal{T}_0)}), \tag{59}$$

where, given $\mathcal{T}_0$, the random variables M_0, $L_1^{(\tau_0 - \mathcal{T}_0)}$ and $L_2^{(\tau_0 - \mathcal{T}_0)}$ are independent.

We are now going to establish a stochastic fixed point equation whose solution is asymptotically equivalent to that of (59). To this purpose let $M^{(N)} + 1$ be a geometrically distributed random variable with parameter $\frac{1 - N^{-\delta}}{2 - N^{-\delta}} = \frac{\rho}{1+\rho}$. Let $L^{(N)}$ be the solution of the stochastic fixed point equation

$$L^{(N)} \stackrel{d}{=} M^{(N)} + \min(L_1^{(N)}, L_2^{(N)})$$

where $L_1^{(N)}$ and $L_2^{(N)}$ are independent copies of $L^{(N)}$ that are also independent of $M^{(N)}$. Let $\pi^{(N)}$ be the distribution of $L^{(N)}$. The fact that the distribution weights of the random

variables $L^{(\tau_0)}$ and $L^{(\tau_0-\mathcal{T}_0)}$ are asymptotically equivalent as $N \to \infty$ strongly suggests that for all $k \in \mathbb{N}_0$

$$\mathfrak{p}_k^{(N)} \sim \pi_k^{(N)} \quad \text{as } N \to \infty. \tag{60}$$

By applying step 4 of the proof of [GSW23] Proposition 11.1 to the random variable $L^{(N)}$ (instead of the random variable L figuring there) and using (60) we obtain the asymptotic validity of the recursion

$$\mathfrak{p}_0 = N^{-\delta}, \qquad (1 - N^{-\delta})(\mathfrak{p}_k - \mathfrak{p}_{k-1}) = \mathfrak{p}_k \left(1 - \mathfrak{p}_k - 2 \sum_{k'=0}^{k-1} \mathfrak{p}_{k'} \right).$$

Together with (58) this gives the assertion (52) of the Claim.

5. Proofs

5.1. **Proof of Theorem 3.2.** Denote by $\mathcal{Y}$ the process whose jump rates are obtained through multiplying (17), (18) by $\frac{1}{m} = N^\beta$. Then Theorem 3.2 is an immediate corollary of the following result, which gives a refined statement on occupation times of $\mathcal{Y}$, depending on various starting conditions.

Theorem 5.1.   *a) For $(\beta, \delta) \in \mathcal{E}$,*

$$\mathbb{E}_{N^\alpha} \left[\int_0^{T_0} \mathbf{1}_{\{1 \leq \mathcal{Y}_t \leq N^\alpha\}} \, \mathrm{d}t \right] \sim 2N^{\alpha - \beta}, \qquad \alpha < \beta + \delta$$

$$\mathbb{E}_{N^\alpha} \left[\int_0^{T_0} \mathbf{1}_{\{1 \leq \mathcal{Y}_t \leq N^\alpha\}} \, \mathrm{d}t \right] \approx e^{2N^{\alpha - \beta - \delta}}, \qquad \beta + \delta < \alpha < 1 - \delta$$

$$\mathbb{E}_{KN^{1-\delta}} \left[\int_0^{T_0} \mathbf{1}_{\{1 \leq \mathcal{Y}_t \leq N^\alpha\}} \, \mathrm{d}t \right] \approx \exp\left(N^{1-\beta-2\delta} K \left(\frac{2}{\rho} - K \right) \right), \qquad K \leq \frac{1}{\rho}$$

$$\mathbb{E}_{\mathcal{Y}_0} [T_0] \approx \exp\left(N^{1-\beta-2\delta} \frac{1}{\rho^2} \right), \qquad \mathcal{Y}_0 \geq \frac{1}{\rho} N^{1-\delta}.$$

b) For $(\beta, \delta) \in \mathcal{P}$,

$$\mathbb{E}_{N^\alpha} [T_0] \sim 2 \left((1+\beta)/2 - \alpha \right) \log N N^{\alpha-\beta}, \qquad \alpha < \frac{1+\beta}{2} \tag{61}$$

$$\mathbb{E}_{N^\alpha} [T_0] \sim N^{(1-\beta)/2} \frac{\pi^{3/2}}{2}, \qquad \frac{1+\beta}{2} \leq \alpha. \tag{62}$$

This we prove now. We denote the time-discrete embedded process corresponding to $\mathcal{Y}$ by $\widetilde{\mathcal{Y}}$, and write $\widetilde{G}(m, n)$ for the expected number of visits at n of $\widetilde{\mathcal{Y}}$ when starting in m. Denote by jp_j and jq_j the upward resp. the downward jump rate of $\mathcal{Y}$ from j. Let us start with an anlysis of $\widetilde{G}(n, n)$. By standard arguments we have

$$\widetilde{G}(n, n) = \frac{1}{\phi(n)}, \tag{63}$$

where $\phi(n)$ is the escape probability of $\widetilde{\mathcal{Y}}$ from the state n, i.e.

$$\phi(n) = \frac{q_n}{q_n + p_n}(1 - h^{(n)}(n-1)), \tag{64}$$

where $h^{(n)} : \{0, 1, \ldots, n\} \to [0, 1]$ is $\widetilde{\mathcal{Y}}$-harmonic on $\{1, \ldots, n-1\}$ and satisfies the boundary conditions $h^{(n)}(0) = 0$, $h^{(n)}(n) = 1$. Hence (cf. e.g. [GSW23])

$$h^{(n)}(\ell) = \frac{\sum_{j=0}^{\ell-1} d(j)}{\sum_{k=0}^{n-1} d(k)}, \quad \ell = 0, \ldots, n, \tag{65}$$

where $d(k)$ is the oddsratio product

$$d(k) := \prod_{j=1}^{k} \frac{q_j}{p_j}.$$

From (65) we obtain

$$1 - h^{(n)}(n-1) = \frac{d(n-1)}{\sum_{k=0}^{n-1} d(k)}. \tag{66}$$

Writing $G(n, n)$ for the expected time spent by $\mathcal{Y}$ in n when starting in n we observe the relation

$$G(n, n) = \frac{1}{\phi(n)} \cdot \frac{1}{n} \cdot \frac{1}{q_n + p_n}.$$

Combining this with (63), (64) and (66) we arrive at

$$G(n, n) = \frac{1}{nq_n} \cdot \sum_{k=0}^{n-1} \frac{d(k)}{d(n-1)}.$$

Let us now turn to $G(m, n)$ with $m \neq n$. For $n \leq m$ we have

$$G(m, n) = G(n, n),$$

such that

$$G(m, n) = G(n, n) = \frac{1}{q_n \cdot n} \sum_{k=0}^{n-1} \frac{d(k)}{d(n-1)}$$

$$= \frac{1}{q_n \cdot n} \sum_{l=0}^{n-1} \prod_{k=l+1}^{n-1} \frac{p_k}{q_k}.$$

while for $m > n$ we have

$$G(m, n) = G(n, n) \cdot \frac{\sum_{l=0}^{m-1} d(l)}{\sum_{l=0}^{n} d(l)}. \tag{67}$$

Writing out (67) gives

$$G(m, n) = \frac{1}{q_n \cdot n} \sum_{k=0}^{n-1} \frac{d(k)}{d(n-1)} \cdot \frac{\sum_{l=0}^{m-1} d(l)}{\sum_{l=0}^{n-1} d(l)}$$

$$= \frac{1}{q_n \cdot n} \sum_{l=0}^{m-1} \prod_{k=l+1}^{n-1} \frac{p_k}{q_k}.$$

In total this gives

$$G(m,n) = \frac{1}{q_n \cdot n} \sum_{l=0}^{m-1 \wedge n-1} \prod_{k=l+1}^{n-1} \frac{p_k}{q_k}. \tag{68}$$

We now have the following expressions for the jump rates from the state n:

$$p_n = \left(\frac{1}{2m} + \frac{1}{\rho}\right)\left(1 - \frac{n}{N}\right)$$

from n to $n+1$ and

$$q_n = \frac{1}{2m}\left(1 - \frac{n}{N}\right) + 1.$$

from n to $n-1$. We thus have:

$$
\begin{aligned}
\frac{q_n}{p_n} &= \left(1 + \frac{2m}{\rho}\right)^{-1}\left(1 + \frac{2m}{1 - n/N}\right) \\
&= \left(\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^k\right)\left(1 + \rho\frac{2m}{\rho} + 2m\sum_{k=1}^{\infty}\left(\frac{n}{N}\right)^k\right) \\
&= \left(\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^k + \rho\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^{k+1}\right) \\
&\quad + 2m\left(\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^k\right)\left(\sum_{k=1}^{\infty}\left(\frac{n}{N}\right)^k\right) \\
&= \left(1 - \sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^{k+1} + \rho\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^{k+1}\right) \\
&\quad + 2m\left(\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^k\right)\left(\sum_{k=1}^{\infty}\left(\frac{n}{N}\right)^k\right) \\
&= \left(1 - (1-\rho)\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^{k+1}\right) + \rho\left(\sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^{k+1}\right)\left(\sum_{k=1}^{\infty}\left(\frac{n}{N}\right)^k\right).
\end{aligned}
$$

Hence if we introduce

$$\mathfrak{S} := \sum_{k=0}^{\infty}(-1)^k\left(\frac{2m}{\rho}\right)^{k+1} \sim 2m,$$

then

$$\frac{q_n}{p_n} = 1 - (1 - \rho)\mathfrak{S} + \rho\mathfrak{S}\sum_{k=1}^{\infty}\left(\frac{n}{N}\right)^k.$$

A second order Taylor expansion thus gives:

$$
\begin{aligned}
\frac{1}{\mathfrak{S}}\log\frac{q_n}{p_n} &= (1-\rho) + \rho\sum_{k=1}^{3}\left(\frac{n}{N}\right)^k - \frac{(1-\rho)^2}{2}\mathfrak{S} - \frac{\rho^2}{2}\mathfrak{S}\left(\sum_{k=1}^{2}\left(\frac{n}{N}\right)^k\right)^2 \\
&\quad + O\left((1-\rho)^3\mathfrak{S}^2 + \mathfrak{S}^2\left(\frac{n}{N}\right)^3 + \left(\frac{n}{N}\right)^4\right) \\
&= (1-\rho) + \rho\sum_{k=1}^{3}\left(\frac{n}{N}\right)^k - \frac{(1-\rho)^2}{2}\mathfrak{S} - \frac{\rho^2}{2}\mathfrak{S}\left(\left(\frac{n}{N}\right)^2 + 2\left(\frac{n}{N}\right)^3\right) \\
&\quad + O\left((1-\rho)^3\mathfrak{S}^2 + \mathfrak{S}^2\left(\frac{n}{N}\right)^3 + \left(\frac{n}{N}\right)^4\right)
\end{aligned}
$$

In particular, if n is of order at most $N^{1-\delta/2}$,

$$
\begin{aligned}
\log\frac{q_n}{p_n} &= -(1-\rho)\mathfrak{S} + \rho\mathfrak{S}\sum_{k=1}^{3}\left(\frac{n}{N}\right)^k - \frac{(1-\rho)^2}{2}\mathfrak{S}^2 - \frac{\rho^2}{2}\mathfrak{S}^2\left(\sum_{k=1}^{3}\left(\frac{n}{N}\right)^k\right)^2 \\
&\quad + O\left(N^{-3\beta-3\delta/2} + N^{-\beta-2\delta}\right) \\
&= -(1-\rho)\mathfrak{S} + \rho\mathfrak{S}\sum_{k=1}^{2}\left(\frac{n}{N}\right)^k - \frac{(1-\rho)^2}{2}\mathfrak{S}^2 - \frac{\rho^2}{2}\mathfrak{S}^2\left(\frac{n}{N}\right)^2 + O\left(N^{-\beta-3\delta/2}\right) \\
&= -(1-\rho)\mathfrak{S}\left(1 - \frac{(1-\rho)}{2}\mathfrak{S}\right) + \rho\mathfrak{S}\frac{n}{N} + \rho\mathfrak{S}\left(1 - \frac{\rho}{2}\mathfrak{S}\right)\left(\frac{n}{N}\right)^2 + O\left(N^{-\beta-3\delta/2}\right).
\end{aligned}
$$

This implies for $K < \infty$ and $n \leq KN^{1-\delta/2}$

$$
\begin{aligned}
\log d(n) &= -(1-\rho)\mathfrak{S}\left(1 - \frac{(1-\rho)}{2}\mathfrak{S}\right)n + \rho\mathfrak{S}\frac{n^2}{2N} + \rho\mathfrak{S}\left(1 - \frac{\rho}{2}\mathfrak{S}\right)\frac{n^3}{3N^2} + o(1) \\
&= -(1-\rho)\mathfrak{S}n + \mathfrak{S}\frac{n^2}{2N} + \mathfrak{S}\left(1 - \frac{\mathfrak{S}}{2}\right)\frac{n^3}{3N^2} + o(1).
\end{aligned}
$$

We will first prove point b).

5.1.1. *The polynomial regime.* Here we have $1 - \beta - 2\delta < 0$. Let us first take $\alpha \leq 1 - \delta$. Then for any $n \leq N^\alpha$, $d(k) \sim 1$ and

$$G(n,n) \sim q_n^{-1} \sim 2N^{-\beta}.$$

Hence

$$\sum_{n=1}^{KN^\alpha} G(n,n) \sim 2KN^{\alpha-\beta}.$$

Now let us take $1 - \delta < \alpha \le 1 - \delta/2$:

$$\sum_{n=KN^{1-\delta}}^{N^\alpha} G(n,n) \sim 2N^{\alpha-\beta} \int_{KN^{1-\delta-\alpha}}^1 du \int_0^1 e^{(1-\rho)\mathfrak{S}N^\alpha u(1-v) - \mathfrak{S}N^{2\alpha}u^2 \frac{1-v^2}{2N} - \left(\mathfrak{S} - \frac{\mathfrak{S}^2}{2}\right)N^{3\alpha}u^3 \frac{1-v^3}{3N^2}} dv.$$

For $1 - \delta < \alpha \le 1 - \delta/2$,

$$\frac{(1-\rho)\mathfrak{S}N^\alpha u(1-v)}{\mathfrak{S}N^{2\alpha}u^2 \frac{(1-v^2)}{2N}} \le \frac{2N^{1-\delta-\alpha}}{u}$$

goes to 0 when N goes to infinity and $u \ne 0$. Likewise,

$$\frac{\mathfrak{S}\left(1 - \frac{\mathfrak{S}}{2}\right)N^{3\alpha}u^3 \frac{(1-v^3)}{3N^2}}{\mathfrak{S}N^{2\alpha}u^2 \frac{(1-v^2)}{2N}} \le 2N^{\alpha-1} \le 2N^{-\delta/2}$$

goes to 0. We deduce that

$$
\begin{aligned}
\sum_{n=KN^{1-\delta}}^{N^\alpha} G(n,n) &\sim 2N^{\alpha-\beta} \int_{KN^{1-\delta-\alpha}}^1 du \int_0^1 e^{-\mathfrak{S}N^{2\alpha}u^2 \frac{(1-v^2)}{2N}} dv \\
&= 2N^{\alpha-\beta} \int_0^1 dv \int_{KN^{1-\delta-\alpha}}^1 e^{-\mathfrak{S}N^{2\alpha}u^2 \frac{(1-v^2)}{2N}} du \\
&= 2\frac{N^{\alpha-\beta}}{\sqrt{\mathfrak{S}N^{2\alpha-1}\frac{1}{2}}} \int_0^1 \frac{dv}{\sqrt{1-v^2}} \int_{K\sqrt{\mathfrak{S}N^{1-2\delta}\frac{(1-v^2)}{2}}}^{\sqrt{\mathfrak{S}N^{2\alpha-1}\frac{(1-v^2)}{2}}} e^{-u^2} du \\
&\sim 2N^{(1-\beta)/2} \int_0^1 \frac{dv}{\sqrt{1-v^2}} \int_{K\sqrt{\mathfrak{S}N^{1-2\delta}\frac{(1-v^2)}{2}}}^{\sqrt{\mathfrak{S}N^{2\alpha-1}\frac{(1-v^2)}{2}}} e^{-u^2} du.
\end{aligned}
$$

We have

$$\mathfrak{S}N^{1-2\delta}\frac{(1-v^2)}{2} \to 0, \qquad N \to \infty.$$

If $1 - \delta < \alpha < (1+\beta)/2$,

$$\sqrt{\mathfrak{S}N^{2\alpha-1}\frac{(1-v^2)}{2}} \to 0, \qquad N \to \infty$$

and

$$\sum_{n=KN^{1-\delta}}^{N^\alpha} G(n,n) \sim 2N^{\alpha-\beta},$$

whereas if $(1 + \beta)/2 < \alpha < 1 - \delta/2$,

$$\sqrt{\mathfrak{S}N^{2\alpha-1}\frac{(1-v^2)}{2}} \to \infty, \qquad N \to \infty$$

and

$$\sum_{n=KN^{1-\delta}}^{N^\alpha} G(n,n) \quad \sim \quad 2N^{(1-\beta)/2}\int_0^1 \frac{dv}{\sqrt{1-v^2}}\int_0^\infty e^{-u^2}\,du$$

$$= \quad 2N^{(1-\beta)/2}\frac{\pi^{3/2}}{4} = 2N^{(1-\beta)/2}\left(\frac{\pi}{2}\right)^{3/2}. \tag{69}$$

Let $n_0 \le n$. Then from (68),

$$G(n_0, n) = \frac{1}{nq_n}\sum_{l=0}^{n_0-1}\frac{d(l)}{d(n-1)}.$$

Hence for any $\varepsilon > 0$,

$$\sum_{n=n_0}^{N^{1-\varepsilon}} G(n_0, n) \le N^{-\beta}\left(\frac{1}{n_0}\sum_{l=0}^{n_0-1}\frac{d(l)}{d(n_0-1)}\right)\sum_{n=n_0}^{N^{1-\varepsilon}}\frac{d(n_0-1)}{d(n-1)}.$$

Now notice that for $n \ge N^{1-\delta/2}$ and N large enough,

$$\frac{q_n}{p_n} \ge 1 + N^{-\beta-\delta/2}.$$

In particular, for $\varepsilon < \delta/2$,

$$\sum_{n=N^{1-\delta/2}}^{N^{1-\varepsilon}}\frac{d(N^{1-\delta/2}-1)}{d(n-1)} \le \sum_{n=0}^{\infty}\left(1+N^{-\beta-\delta/2}\right)^{-n} \sim N^{\beta+\delta/2}.$$

Moreover we get from the previous computations,

$$\frac{1}{N^{1-\delta/2}}\sum_{l=0}^{N^{1-\delta/2}-1}\frac{d(l)}{d(N^{1-\delta/2}-1)}$$

$$\sim \quad \int_0^1 e^{(1-\rho)\mathfrak{S}N^{1-\delta/2}(1-v)-\mathfrak{S}N^{2-\delta}\frac{(1-v^2)}{2N}-\mathfrak{S}\left(1-\frac{\mathfrak{S}}{2}\right)N^{3-3\delta/2}\frac{(1-v^3)}{3N^2}}\,dv$$

$$\le \quad \int_0^1 e^{-\mathfrak{S}N^{2-\delta}\frac{(1-v^2)}{4N}}\,dv$$

$$\le \quad \int_0^1 e^{-\mathfrak{S}N^{1-\delta}\frac{(1-v)}{4}}\,dv$$

$$= \quad \frac{2}{\sqrt{\mathfrak{S}N^{1-\delta}}}\int_0^{\frac{\sqrt{\mathfrak{S}N^{1-\delta}}}{2}} e^{-v}\,dv \sim N^{(\beta+\delta-1)/2}.$$

Hence

$$\sum_{n=N^{1-\delta/2}}^{N^{1-\varepsilon}} G(N^{1-\delta/2}, n) \leq N^{-\beta} \left(\sqrt{2} N^{(\beta+\delta-1)/2} \right) N^{\beta+\delta/2} \sim \sqrt{2} N^{\delta-(1-\beta)/2}.$$

But from the condition

$$1 - \beta - \delta > 0$$

we deduce that $\delta < 1 - \beta$ and thus for large N,

$$\sum_{n=N^{1-\delta/2}}^{N^{1-\varepsilon}} G(N^{1-\delta/2}, n) = o\left(N^{(1-\beta)/2} \right).$$

Adding (69) yields for any small $\varepsilon > 0$,

$$\sum_{n=N^{\frac{1-\beta}{2}+\varepsilon}}^{N^{1-\varepsilon}} G(N^{(1-\beta)/2+\varepsilon}, n) \leq \sum_{n=N^{\frac{1-\beta}{2}+\varepsilon}}^{N^{1-\delta/2}} G(n, n) + \sum_{n=N^{1-\delta/2}}^{N^{1-\varepsilon}} G(N^{1-\delta/2}, n) = o\left(N^{\frac{1-\beta}{2}} \right).$$

This ends the proof of (62).

Let us come back to the case $\alpha < (1 + \beta)/2$. First notice that for $\varepsilon > 0$ such that $\alpha + \varepsilon < (1 + \beta)/2$,

$$\sum_{n=N^{\alpha}}^{N^{(1+\beta)/2-\varepsilon}} G(N^{\alpha}, n) \sim \sum_{n=N^{\alpha}}^{N^{(1+\beta)/2-\varepsilon}} \frac{N^{\alpha}}{nq_n} \sim 2\left((1+\beta)/2 - \varepsilon - \alpha \right) \log N N^{\alpha-\beta}.$$

Indeed we proved that $d(n) \sim 1$ for any $n \leq N^{(1+\beta)/2-\varepsilon}$. Now notice that the condition $1 - \beta - 2\delta < 0$ implies that $1 - \delta < (1+\beta)/2$. Moreover, from the definition of q_n and p_n we see that $q_n/p_n \geq 1$ as soon as $n \geq N^{1-\delta}$. Hence if ε is small enough, $1 - \delta < (1+\beta)/2 - \varepsilon$ and

$$\sum_{N^{(1+\beta)/2-\varepsilon}}^{N^{(1+\beta)/2+\varepsilon}} G(N^{\alpha}, n) = \sum_{n=N^{(1+\beta)/2-\varepsilon}}^{N^{(1+\beta)/2+\varepsilon}} \frac{1}{nq_n} \sum_{l=0}^{N^{\alpha}-1} \frac{d(l)}{d(n-1)}$$

$$\leq \frac{1}{d(N^{(1+\beta)/2-\varepsilon}-1)} \sum_{n=N^{(1+\beta)/2-\varepsilon}}^{N^{(1+\beta)/2+\varepsilon}} \frac{1}{nq_n} \sum_{l=0}^{N^{\alpha}-1} d(l)$$

$$\sim 2\varepsilon \log N N^{\alpha-\beta}.$$

Finally,

$$\sum_{N^{(1+\beta)/2+\varepsilon}}^{N} G(N^{\alpha}, n) = \sum_{N^{(1-\beta)/2+\varepsilon}}^{N} \frac{1}{nq_n} \sum_{k=0}^{N^{\alpha}} \frac{d(k)}{d(n-1)}$$

$$\leq \frac{2N^{\alpha-\beta}}{N^{(1+\beta)/2+\varepsilon}} \sum_{N^{(1+\beta)/2+\varepsilon}}^{N} \frac{d(N^\alpha-1)}{d(n-1)}$$

$$\sim \frac{2N^{\alpha-\beta}}{N^{(1+\beta)/2+\varepsilon}} \sum_{N^{(1+\beta)/2+\varepsilon}}^{N} \frac{1}{d(n-1)}.$$

But from the Taylor expansion we see that

$$\sum_{N^{(1+\beta)/2+\varepsilon}}^{N} \frac{1}{d(n-1)} \leq \int_{N^{(1+\beta)/2+\varepsilon}}^{N} e^{-\mathfrak{S}x^2\frac{1}{4N}} dx = o\left(1\right).$$

This concludes the proof of (61).

Let us now focus on point a):

5.1.2. *The exponential regime.* Here we have $1 - \beta - 2\delta > 0$. From the Taylor expansion we deduce that there are three possibilities:

- If $\alpha < 1 - \delta$,
$$\log d(N^\alpha) \sim -(1-\rho)\mathfrak{S}N^\alpha$$

- If $K \in \mathbb{R}_+$,
$$\log d(KN^{1-\delta}) = -(1-\rho)\mathfrak{S}KN^{1-\delta} + \rho\mathfrak{S}\frac{(KN^{1-\delta})^2}{2N} + o\left(N^{1-\beta-2\delta}\right)$$

- If $\alpha > 1 - \delta$,
$$\log d(N^\alpha) \sim \rho\mathfrak{S}\frac{n^2}{2N}$$

We will consider the three cases successively.

If $\alpha < \beta + \delta$,

$$\sum_{n=1}^{N^\alpha} G(n,n) \sim 2N^{\alpha-\beta} \int_0^1 du \int_0^1 e^{(1-\rho)\mathfrak{S}N^\alpha u(1-v)} dv$$

$$= \frac{2N^{\alpha-\beta}}{(1-\rho)\mathfrak{S}N^\alpha} \int_0^1 \frac{1}{1-v}\left(e^{(1-\rho)\mathfrak{S}N^\alpha(1-v)} - 1\right) dv \sim 2N^{\alpha-\beta}.$$

Recall that $\approx$ stands for logarithmical equivalences. If $\beta + \delta < \alpha < 1 - \delta$,

$$\sum_{n=1}^{N^\alpha} G(n,n) \approx \int_0^1 du \int_0^1 e^{(1-\rho)\mathfrak{S}N^\alpha u(1-v)} dv \approx e^{(1-\rho)\mathfrak{S}N^\alpha} \approx e^{2N^{\alpha-\beta-\delta}}.$$

Now, let us take $K \in \mathbb{R}_+$. Then:

$$\sum_{n=1}^{KN^{1-\delta}} G(n,n) \approx \int_1^{KN^{1-\delta}} dn \int_1^n \exp\left(\frac{x}{N^{1+\beta}}\left(x - \frac{2}{\rho}N^{1-\delta}\right) - \frac{n}{N^{1+\beta}}\left(n - \frac{2}{\rho}N^{1-\delta}\right)\right) dx$$

$$\approx \int\limits_{1}^{KN^{1-\delta}} dn \int\limits_{1}^{n} \exp\left(\frac{1}{N^{1+\beta}}(x-n)\left(x+n-\frac{2}{\rho}N^{1-\delta}\right)\right) dx$$

$$\approx \int\limits_{0}^{K} d\lambda \int\limits_{0}^{\lambda} \exp\left(N^{1-\beta-2\delta}(u-\lambda)\left(u+\lambda-\frac{2}{\rho}\right)\right) du$$

$$\approx \int\limits_{0}^{K} d\lambda \int\limits_{0}^{\lambda} \exp\left(N^{1-\beta-2\delta}u\left(u-2\lambda+\frac{2}{\rho}\right)\right) du.$$

There are then two possible cases for the last integral:

- If $\lambda \geq 2/\rho$,

$$\int_{0}^{\lambda} \exp\left(N^{1-\beta-2\delta}u\left(u-2\lambda+\frac{2}{\rho}\right)\right) du = O(1).$$

- If $\lambda \leq 2/\rho$,

$$\int_{0}^{\lambda} \exp\left(N^{1-\beta-2\delta}u\left(u-2\lambda+\frac{2}{\rho}\right)\right) du \approx \exp\left(N^{1-\beta-2\delta}\lambda\left(\frac{2}{\rho}-\lambda\right)\right).$$

We deduce that

- If $K \leq 1/\rho$

$$\sum_{n=1}^{KN^{1-\delta}} G(n,n) \approx \exp\left(N^{1-\beta-2\delta}K\left(\frac{2}{\rho}-K\right)\right)$$

- If $1/\rho \leq K$,

$$\sum_{n=1}^{KN^{1-\delta}} G(n,n) \approx \exp\left(N^{1-\beta-2\delta}\frac{1}{\rho^2}\right)$$

Notice that for $n \geq N^{1-\delta}$, $q_n/p_n \geq 1$ and recall that for $n \geq n_0$,

$$G(n_0, n) = \frac{1}{nq_n} \sum_{l=0}^{n_0-1} \prod_{k=l+1}^{n-1} \frac{p_k}{q_k}.$$

Hence for $(2/\rho)N^{1-\delta} \leq n_0 \leq n$

$$G(n_0, n) \leq G\left((2/\rho)N^{1-\delta}, n\right) \leq \frac{1}{nq_n} \sum_{l=0}^{(2/\rho)N^{1-\delta}-1} \prod_{k=l+1}^{(2/\rho)N^{1-\delta}-1} \frac{p_k}{q_k}$$

$$= \frac{(2/\rho)N^{1-\delta}q_{(2/\rho)N^{1-\delta}}}{nq_n} G((2/\rho)N^{1-\delta}, (2/\rho)N^{1-\delta})$$

$$= \frac{(2/\rho)N^{1-\delta}q_{(2/\rho)N^{1-\delta}}}{nq_n}O(1).$$

We deduce that for $(2/\rho)N^{1-\delta} \leq n_0$

$$\sum_{n=n_0}^{N} G(n_0,n) = o\left(\exp\left(N^{1-\beta-2\delta}\right)\right).$$

This ends the proof of point a).

5.2. Proof of Propositions 4.3 and 4.4.

Proof of Proposition 4.3. We abbreviate $p_k := p_k^{(N,\delta)}$, $\rho := 1 - N^{-\delta}$, and proceed by induction. For $k = 0$, (49) is true since $p_0 = 1 - \rho$ by (44). Let us now assume that for some $k \in \mathbb{N}$ the asymptotics (49) is valid for all $\ell < k$. This implies in particular the validity of the asymptotics (50) for this k. Re-arranging (44) we obtain

$$p_k = -\frac{1}{2}\left(\rho - 1 + 2\sum_{\ell<k} p_\ell\right) + \sqrt{\frac{1}{4}\left(\rho - 1 + 2\sum_{\ell<k} p_\ell\right)^2 + \rho p_{k-1}}. \tag{70}$$

Because of (50), and since $\rho - 1 = -N^{-\delta}$, we have

$$N^{\delta/2^k}\left(\rho - 1 + 2\sum_{\ell<k} p_\ell\right) \to 0 \quad \text{as } N \to \infty.$$

This together with (70) implies

$$N^{\delta/2^k}p_k \sim \sqrt{\rho N^{\delta/2^k \mathcal{D}1}p_{k-1}} \quad \text{as } N \to \infty.$$

As the r.h.s converges to 1 by induction assumption, this completes the induction step. $\square$

Proof of Proposition 4.4. First notice that for any $k \geq 1$,

$$\rho - 1 + 2\sum_{\ell<k} p_\ell \geq \rho - 1 + 2p_0 = 1 - \rho \geq 0.$$

From the inequality

$$\sqrt{a^2 + b} - a \leq \sqrt{b}$$

which holds for any $a, b \geq 0$ and from (70), we get that for any $k \geq 1$,

$$p_k \leq \sqrt{\rho p_{k-1}} \leq (\rho p_0)^{(1/2)^k} = (\rho N^{-\delta})^{(1/2)^k}. \tag{71}$$

We deduce that for any $\varepsilon > 0$ and $k \leq \log\log N/(\log 2 + \varepsilon)$,

$$\sum_{\ell=0}^{k} p_\ell \leq \sum_{\ell=0}^{k}(\rho N^{-\delta})^{(1/2)^\ell} \leq \frac{\log\log N}{\log 2 + \varepsilon}(\rho N^{-\delta})^{(1/2)^{\log\log N/(\log 2+\varepsilon)}}$$

$$= \frac{\log\log N}{\log 2 + \varepsilon}e^{-\delta(\log N)^{\frac{\varepsilon}{\log 2+\varepsilon}}} \to 0 \quad \text{as } N \to \infty.$$

We deduce that
$$\text{median}(\mathbf{p}^{(N,\delta)}) \geq \frac{\log\log N}{\log 2 + \varepsilon}$$
when N is large enough.

From (71) we get
$$\sum_{\ell=0}^{k-2} p_\ell \leq (\rho N^{-\delta})^{(1/2)^{k-2}} \sum_{\ell=0}^{k-2} (\rho N^{-\delta})^{(2^\ell)} \leq (\rho N^{-\delta})^{(1/2)^{k-2}} \sum_{\ell=0}^{\infty} (\rho N^{-\delta})^\ell = \frac{(\rho N^{-\delta})^{(1/2)^{k-2}}}{1 - \rho N^{-\delta}}.$$

Adding (70) yields
$$p_k \geq \sqrt{\rho p_{k-1}} - p_{k-1} - \frac{(\rho N^{-\delta})^{(1/2)^{k-2}}}{1 - \rho N^{-\delta}}.$$

Now let us choose a small ε and introduce K_ε such that
$$(\rho N^{-\delta})^{(1/2)^{K_\varepsilon}} \leq \varepsilon \leq (\rho N^{-\delta})^{(1/2)^{K_\varepsilon+1}}. \tag{72}$$

Then for any $k \leq K_\varepsilon$,
$$p_k \geq \left(\sqrt{\rho}(1 - 3\varepsilon/\sqrt{\rho})\right)^{\mathfrak{s}(k)} (\rho N^{-\delta})^{(1/2)^k},$$

where $\mathfrak{s}(k) = 1 + 1/2 + \ldots + 1/2^{k-1}$. Indeed this holds for $k = 0$ and if it holds for $k - 1 < K_\varepsilon$,

$$
\begin{aligned}
p_k &\geq \sqrt{\rho p_{k-1}} - p_{k-1} - \frac{(\rho N^{-\delta})^{(1/2)^{k-2}}}{1 - \rho N^{-\delta}} \\
&= \sqrt{\rho p_{k-1}} - \sqrt{\rho p_{k-1}}\frac{\sqrt{p_{k-1}}}{\sqrt{\rho}} - \frac{\sqrt{\rho p_{k-1}}}{\sqrt{\rho p_{k-1}}}\frac{(\rho N^{-\delta})^{(1/2)^{k-2}}}{1 - \rho N^{-\delta}} \\
&\geq \sqrt{\rho p_{k-1}} - \sqrt{\rho p_{k-1}}\frac{(\rho N^{-\delta})^{(1/2)^k}}{\sqrt{\rho}} - \frac{\sqrt{\rho p_{k-1}}}{\rho^{1/2 + \frac{\mathfrak{s}(k-1)}{4}}(1 - 3\frac{\varepsilon}{\sqrt{\rho}})^{\frac{\mathfrak{s}(k-1)}{2}}}\frac{(\rho N^{-\delta})^{(1/2)^{k-2}-(1/2)^k}}{1 - \rho N^{-\delta}} \\
&\geq \sqrt{\rho p_{k-1}}\left(1 - \frac{(\rho N^{-\delta})^{(1/2)^k}}{\sqrt{\rho}} - \frac{1}{\rho^{1/2 + \mathfrak{s}(k-1)/4}(1 - \frac{3\varepsilon}{\sqrt{\rho}})^{\frac{\mathfrak{s}(k-1)}{2}}}\frac{(\rho N^{-\delta})^{(1/2)^{k-1}}}{1 - \rho N^{-\delta}}\right)
\end{aligned}
$$

where we used (71) and the induction hypothesis. Adding again the induction hypothesis and (72) then leads to

$$
\begin{aligned}
p_k &\geq \sqrt{\rho p_{k-1}}\left(1 - \frac{\varepsilon}{\sqrt{\rho}} - \frac{1}{\rho^{\mathfrak{s}(k)/2}(1 - 3\varepsilon/\sqrt{\rho})^{\mathfrak{s}(k-1)/2}}\frac{\varepsilon}{1 - \rho N^{-\delta}}\right) \\
&\geq \rho^{\mathfrak{s}(k)/2}(1 - 3\varepsilon/\sqrt{\rho})^{\mathfrak{s}(k-1)/2}(\rho N^{-\delta})^{(1/2)^k}\left(1 - \frac{\varepsilon}{\sqrt{\rho}} - 2\frac{\varepsilon}{\sqrt{\rho}}\right) \\
&= \rho^{\mathfrak{s}(k)/2}(1 - 3\varepsilon/\sqrt{\rho})^{\mathfrak{s}(k)}(\rho N^{-\delta})^{(1/2)^k}.
\end{aligned}
$$

Notice that for any $k \in \mathbb{N}$,
$$\mathfrak{s}(k) \leq \mathfrak{s}(\infty) = 2$$

and thus
$$\rho^{\mathfrak{s}(k)/2}(1 - 3\varepsilon/\sqrt{\rho})^{\mathfrak{s}(k)} \geq \rho(1 - 3\varepsilon/\sqrt{\rho})^2.$$
In particular, this holds true for $k = K_\varepsilon$, and thus
$$
\begin{aligned}
p_{K_\varepsilon} &\geq \rho(1 - 3\varepsilon/\sqrt{\rho})^2(\rho N^{-\delta})^{(1/2)^{K_\varepsilon}} \\
&= \rho(1 - 3\varepsilon/\sqrt{\rho})^2\left((\rho N^{-\delta})^{(1/2)^{K_\varepsilon+1}}\right)^2 \\
&\geq \rho(1 - 3\varepsilon/\sqrt{\rho})^2\varepsilon^2.
\end{aligned}
$$

As the $(p_k)_{k \in \mathbb{N}}$ are non decreasing up to the median (recall (45)), the last inequality implies that the median is smaller than $K_\varepsilon + 4\varepsilon^{-2}$. Adding that
$$K_\varepsilon \sim \frac{\log(\log(N^\delta)/\log(2))}{\log(2)} \qquad \text{as } N \to \infty$$
ends the proof. □

REFERENCES

[AP13] J. Audiffren and E. Pardoux. Muller's ratchet clicks in finite time. *Stochastic Processes and their Applications*, 123(6):2370–2397, 2013.

[BCH18] E. Baake, F. Cordero, and S. Hummel. A probabilistic view on the deterministic mutation–selection equation: dynamics, equilibria, and ancestry via individual lines of descent. *Journal of mathematical biology*, 77:795–820, 2018.

[BFM00] T. Bäck, T. Fogel, and D. Michalewicz. Evolutionary computation 1 (Basic algorithms and operators). 2000.

[BT96] T. Blickle and L. Thiele. A comparison of selection schemes used in evolutionary algorithms. *Evolutionary Computation*, 4:361–394, 12 1996. doi:10.1162/evco.1996.4.4.361.

[CCM16] J.-R. Chazottes, P. Collet, and S. Méléard. Sharp asymptotics for the quasi-stationary distribution of birth-and-death processes. *Probability Theory and Related Fields*, 164(1-2):285–332, 2016.

[CHS22] F. Cordero, S. Hummel, and E. Schertzer. General selection models: Bernstein duality and minimal ancestral structures. *The Annals of Applied Probability*, 32(3):1499–1556, 2022.

[CS20] A. G. Casanova and C. Smadi. On λ-fleming–viot processes with general frequency-dependent selection. *Journal of Applied Probability*, 57(4):1162–1197, 2020.

[EPW09] A. Etheridge, P. Pfaffelhuber, and A. Wakolbinger. How often does the ratchet click? Facts, heuristics, asymptotics, 2009, arXiv:0709.2775.

[Fel74]   J. Felsenstein. The evolutionary advantage of recombination. *Genetics*, 78(2):737–756, 1974.

[GC00]   I. Gordo and B. Charlesworth. The degeneration of asexual haploid populations and the speed of muller's ratchet. *Genetics*, 154 3:1379–87, 2000.

[GCS18]   A. González Casanova and D. Spanò. Duality and fixation in ξ-wright–fisher processes with frequency-dependent selection. *The Annals of Applied Probability*, 2018.

[GSW23]   A. González Casanova, C. Smadi, and A. Wakolbinger. Quasi-equilibria and click times for a variant of muller's ratchet, 2023, arXiv:2211.13109.

[Hai78]   J. Haigh. The accumulation of deleterious genes in a population—muller's ratchet. *Theoretical Population Biology*, 14(2):251–267, 1978. doi:https://doi.org/10.1016/0040-5809(78)90027-8.

[JK14]   S. Jansen and N. Kurt. On the notion (s) of duality for markov processes. 2014.

[KN97]   S. M. Krone and C. Neuhauser. Ancestral processes with selection. *Theoretical population biology*, 51(3):210–237, 1997.

[Lam05]   A. Lambert. The branching process with logistic growth. *The Annals of Applied Probability*, 15(2):1506–1535, 2005. URL http://www.jstor.org/stable/30038363.

[ME13]   J. J. Metzger and S. Eule. Distribution of the fittest individuals and the rate of Muller's ratchet in a model with overlapping generations. *PLoS computational biology*, 9(11):e1003303, 2013.

[Möh99]   M. Möhle. The concept of duality and applications to Markov processes arising in neutral population genetics models. *Bernoulli*, 5(5):761 – 777, 1999.

[MPV20]   M. Mariani, E. Pardoux, and A. Velleret. Metastability between the clicks of the Muller ratchet. 2020, arXiv:2007.14715.

[Mul64]   H. Muller. The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 1(1):2–9, 1964. doi:https://doi.org/10.1016/0027-5107(64)90047-8.

[NK97]   C. Neuhauser and S. M. Krone. The genealogy of samples in models with selection. *Genetics*, 145(2):519–534, 1997.

[PBB+15]   T. Paixão, G. Badkobeh, N. Barton, D. Çörüş, D.-C. Dang, T. Friedrich, P. K. Lehre, D. Sudholt, A. M. Sutton, and B. Trubenová. Toward a unifying framework for evolutionary processes. *Journal of Theoretical Biology*, 383:28–43, 2015.

[PP13]   C. Pokalyuk and P. Pfaffelhuber. The ancestral selection graph under strong directional selection. *Theoretical population biology*, 87:25–33, 2013.

[PSW12]   P. Pfaffelhuber, P. R. Staab, and A. Wakolbinger. Muller's ratchet with compensatory mutations. *The Annals of Applied Probability*, 22(5):2108 – 2132, 2012. doi:10.1214/11-AAP836.

[SCS93]   W. Stephan, L. Chao, and J. G. Smale. The advance of muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. *Genetical research*, 61 3:225–31, 1993.

Jan Lukas Igelbrink, Institut für Mathematik, Johannes Gutenberg-Universität Mainz and Goethe-Universität Frankfurt, Germany.

Email address: jigelbri@uni-mainz.de, igelbrin@math.uni-frankfurt.de

Adrián González Casanova, Universidad Nacional Autónoma de México (UNAM), Instituto de Matemáticas, Circuito exterior, Ciudad Universitaria, 04510, México, 11; University of California, Berkeley. Department of Statistics, 367 Evans Hall, University of California, Berkeley, CA 94720-386

Email address: adriangcs@matem.unam.mx, gonzalez.casanova@berkeley.edu

Charline Smadi, Univ. Grenoble Alpes, INRAE, LESSEM, 38000 Grenoble, France and Univ. Grenoble Alpes, CNRS, Institut Fourier, 38000 Grenoble, France

Email address: charline.smadi@inrae.fr

Anton Wakolbinger, Goethe-Universität, Institut für Mathematik, 60629 Frankfurt am Main, Germany

Email address: wakolbinger@math.uni-frankfurt.de