Genetic weight of a mutant in a biparental Moran model with infinite selection

Camille Coron* Olivier Hénard†

November 3, 2026

Abstract

We study, after Coron Le Jan [6], the genetic contribution of a single strongly advantageous mutant in a biparental Moran model under infinite selection. We prove a functional limit theorem for the entire beneficial mutation spreading period: after normalization by $\sqrt{N}$, the contributions carried by the advantaged and disadvantaged subpopulations converge jointly in law to a non-degenerate random multiple of an explicit deterministic profile. Randomness is generated during the initial microscopic growth of the mutant lineage and subsequently propagated through the population along a deterministic macroscopic trajectory. At the heart of this phenomenon is an eigenmode martingale that measures the imbalance between the mean genetic weight of the founding ancestor within the two classes; its limiting value provides precisely the random amplitude of the macroscopic profile and is characterized by a smoothing equation.

1 Context

When a strongly beneficial mutation appears in a single individual, its spread through the population affects not only the selected allele itself but also the transmission of the whole genetic material carried by its original bearer, a phenomenon known as genetic hitchhiking [22]. In this work, we study this effect in a biparental Moran model with infinite viability selection.

The Moran model is a classical population genetics model in which the population size remains constant and reproduction and replacement occur one individual at a time; see [19]. In the usual monoparental version, each newborn has a single parent, so that the genetic contribution of a given ancestor (i.e. the quantity of genome of the population coming from this individual) is simply the number of its currently alive descendants. In a biparental population, the situation is fundamentally different. Each newborn has two parents, and in the model considered here, the genetic contribution from any given ancestor carried by this newborn is defined as the average of the contributions carried by those two parents. Contributions are therefore real-valued and evolve through series of averaging along the random pedigree.

We start from a single individual carrying a strongly beneficial mutation, and whose genetic contribution through time is our object of interest. Selection acts through death: individuals carrying the beneficial mutation are less likely to be replaced. This mutation is transmitted from parents to offspring following a Mendelian rule at one haploid locus (i.e. the type, advantaged or not, of a newborn is chosen uniformly at random among the types of its two parents). By contrast, the genetic contribution carried by an offspring is defined to be the average of the genetic contributions carried by its two parents. This rule may be viewed as an idealized genome-wide approximation, corresponding heuristically to a law-of-large-numbers effect across

*Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA Paris-Saclay, 91120, Palaiseau, France. Email: camille.coron@inrae.fr

†Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay, 91405, Orsay, France. Email: olivier.henard@universite-paris-saclay.fr

many independently transmitted neutral loci. Thus selection does not directly modify the transmission of genetic material at reproduction; rather, by changing which individuals survive and are replaced, it reshapes the pedigree along which this material is propagated.

Biparental genetics studies started with the articles [3] dealing with time scales to common ancestors and [7] studying the asymptotic quantity of transmitted genome in the neutral biparental Wright-Fisher model. This topic has also received some more recent interest, through the articles [2, 18, 13] studying ancestors' genetic contribution dynamics in biparental Wright-Fisher populations, the articles [15, 10] and [21, 1] studying the genome structure of sexually reproducing populations, and the articles [23, 24, 8, 20] studying the conditioning of gene genealogies by the fixed underlying pedigree. Previous work has also addressed asymptotic genetic contribution for the neutral biparental Moran pedigree [5], the expected genetic contribution of a single strongly advantaged mutant [6], and the deterministic large-population regime in which a macroscopic fraction of the initial population is advantageous [4]. Here we consider the singular regime in which the advantageous type starts from a single individual and selection is infinite. We derive a functional limit theorem describing the evolution of the founder's genetic contribution throughout the advantageous allele spreading, from the appearance of the mutant up to fixation.

A distinctive feature of this regime is that randomness survives in the large-population limit. When a positive proportion of the population is initially advantageous, the system is in a law-of-large-numbers regime and the limiting genetic profile is deterministic [4]. Starting from a single mutant produces a different mechanism. During the first generations of the allelic replacement, the small advantaged population develops randomly and generates a non-degenerate random amplitude. Once the advantaged population has reached a mesoscopic and then macroscopic size, fluctuations become negligible at the relevant scale: the random amplitude created during this initial phase is then transported through the population according to an essentially deterministic profile. The limiting process therefore separates into a random amplitude, generated near the origin of the advantageous allele spreading, and a deterministic macroscopic shape.

This separation between the random seeding phase and the subsequent bulk evolution also guides the proof of our main result, stated in Theorem 1. The early growth of the advantaged population is analysed through a coupling with a random recursive tree, which identifies the law of the random amplitude appearing in the limit as the solution to a fixed point equation. The macroscopic phase is controlled by a martingale built from the difference between the mean genetic contributions carried by advantageous and disadvantaged individuals. Finally, separate approximation arguments connect the initial regime to the bulk evolution and control the terminal window near fixation. Section 3 describes in more details the structure and main ingredients of the proof of Theorem 1.

2 Model and main result

Model We consider a discrete-time biparental Moran model with two types, advantaged and disadvantaged, and with infinite selection at death. The population size $N \geqslant 1$ is fixed, and we write

$$I_N := \{1, \dots, N\}.$$

Each individual has a type, which remains fixed throughout its lifetime. We denote by $\mathcal{Y}_0 \subset I_N$ the set of initially advantaged individuals and, unless otherwise stated, assume $\mathcal{Y}_0 = \{1\}$, corresponding to the arrival of a single advantaged mutant.

At each discrete time $n \in \mathbb{Z}_+$, two parents μ_n and π_n are sampled independently and uniformly from I_N. Independently, as long as $Y_n < N$, one individual κ_n is sampled uniformly from the set of disadvantaged individuals and is replaced by a child of μ_n and π_n. By convention, the child inherits the type of the mother μ_n: since the two parents are sampled symmetrically, this is equivalent in law to choosing uniformly which parent transmits the allele at the selected

locus. An example of simulation of this pedigree and the advantage status of individuals is given in Figure 2 for $N = 4$ and 6 time steps.

Thus the number of advantaged individuals increases precisely when μ_n is advantaged. Sampling the dying individual only among the disadvantaged class is the infinite viability selection assumption. Once the number of advantaged individuals is equal to N all individuals are equivalent and the model becomes neutral, i.e. the dying individual is chosen uniformly in the population. Asymptotic genetic contribution of ancestors in this neutral framework has been studied in [5].

More formally, let $\mathcal{Y}_n$ be the advantaged class at time n, and set

$$Y_n := |\mathcal{Y}_n|, \qquad \mathcal{D}_n := I_N \backslash \mathcal{Y}_n, \qquad D_n := N - Y_n.$$

We shall use the natural filtration of the reproduction process,

$$\mathcal{F}_n = \sigma\big(\mathcal{Y}_0, (\mu_\ell, \pi_\ell, \kappa_\ell) : 0 \leqslant \ell < n\big).$$

Thus $\mathcal{Y}_n, A_n, B_n, C_n$ are $\mathcal{F}_n$-measurable, and, conditionally on $\mathcal{F}_n$, the parents μ_n, π_n are independent uniform variables on I_N, while κ_n is uniform on $\mathcal{D}_n$ (as long as $Y_n < N$, then uniform on I_N), independently of the parents.

We encode genetic inheritance through stochastic matrices $(A_n)_{n\geqslant 0}$, with $A_n \in [0,1]^{N\times N}$, defined by

$$A_0(i,j) := \mathbf{1}_{\{i=j\}},$$

and, recursively,

$$A_{n+1}(i,j) = \begin{cases} A_n(i,j), & i \neq \kappa_n, \\ \dfrac{A_n(\mu_n,j) + A_n(\pi_n,j)}{2}, & i = \kappa_n. \end{cases} \tag{1}$$

Therefore, conditionally on $\mathcal{F}_n$, since the genealogy of a gene at a locus that is independent from the locus under selection can be seen as a random walk on the pedigree backwards in time,

$$A_n(i,j) = \mathbb{P}\left(\text{a gene sampled from individual } i \text{ at time } n \text{ descends from } j \text{ at time } 0 \,|\, \mathcal{F}_n\right).$$

In this sense, $A_n(i,j)$ is the genetic contribution of ancestor j to the genome of individual i at time n and $\sum_{i=1}^N A_n(i,j)$ is the total genetic contribution of ancestor j at time n. Figure 2 shows a simulation of the population with $N = 4$ until fixation of the advantaged type as well as the genetic contributions of the initially advantaged ancestor 1 in each individual once this fixation has occurred.

We focus on the total genetic contribution $\sum_{i=1}^N A_n(i,1)$ of the initially advantaged individual. Following [4], we decompose it according to the type of individuals carrying this contribution:

$$B_n := \sum_{i\in\mathcal{Y}_n} A_n(i,1), \qquad C_n := \sum_{i\in\mathcal{D}_n} A_n(i,1). \tag{2}$$

The variable B_n is the genetic contribution carried by advantaged individuals, while C_n is the contribution carried by disadvantaged individuals. Hence $B_n + C_n$ is the total genetic contribution of ancestor 1 at time n.

Throughout the paper, the dependence in N is usually suppressed from the notation.

Main result For $1 \leqslant k \leqslant N$, define

$$S_k := \inf\{n \geqslant 0 : Y_n = k\}.$$

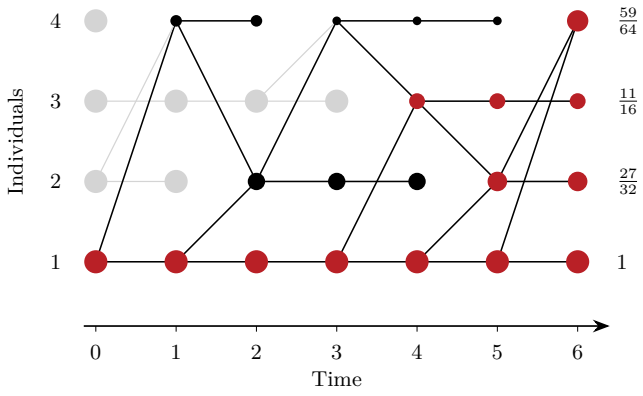


Figure 1: A small N instance of the two-type Moran model, starting from a single advantaged individual (in red). Descendants of the initial mutant are shown in black or red according to their type (disadvantaged or advantaged, respectively), which is inherited from the mother. Circle radii are proportional to the fraction of genome inherited from the initial mutant. All other individuals are shown in light grey. The numbers on the right indicate the genetic contribution of the initial mutant (ancestor 1) to each individual in the final generation, namely $A_{S_N}(1,j)$ for $1 \leqslant j \leqslant N$.

Thus $S_1 = 0$ under our standing assumption $\mathcal{Y}_0 = \{1\}$. The time S_N is the fixation time of the advantaged type. Since Y_n is nondecreasing and $\mathcal{F}_n$-measurable, S_k is an $(\mathcal{F}_n)$-stopping time. For $y \in [0,1]$, set

$$k_N(y) := 1 \vee \lfloor yN \rfloor$$

and equip $\mathbb{R}^2$ with the infinite norm $|.|$.

Theorem 1. *As $N \to \infty$,*

$$\left(\frac{1}{\sqrt{N}} \begin{pmatrix} B_{S_{k_N(y)}} \\ C_{S_{k_N(y)}} \end{pmatrix} \right)_{0 \leqslant y \leqslant 1} \Longrightarrow \left(W \begin{pmatrix} b(y) \\ c(y) \end{pmatrix} \right)_{0 \leqslant y \leqslant 1}.$$

in $D([0,1], \mathbb{R}^2)$, endowed with the Skorokhod J_1 topology, where

$$b(y) := \frac{2}{\sqrt{\pi}} \sqrt{y}\,(1+y), \qquad c(y) := \frac{2}{\sqrt{\pi}} \sqrt{y}\,(1-y), \qquad y \in [0,1].$$

The random variable W is the unique L^2 solution with mean 1 of

$$W \stackrel{d}{=} \sqrt{U}\, W^{(1)} + \frac{1}{2} \sqrt{1-U}\, W^{(2)},$$

where $U \sim \mathrm{Unif}[0,1]$, and $W^{(1)}$ and $W^{(2)}$ are independent copies of W, independent of U. Moreover, $W > 0$ almost surely.

Remark 1. In particular, at fixation,

$$\frac{B_{S_N}}{\sqrt{N}} \Longrightarrow \frac{4}{\sqrt{\pi}} W \qquad \text{as } N \to \infty.$$

The corresponding first moment also satisfies $\mathbb{E}[B_{S_N}]/\sqrt{N} \longrightarrow 4/\sqrt{\pi}$. This follows from the explicit first-moment formula of Lemma 2, applied at level $N-1$, together with the fact that the final increment from S_{N-1} to S_N is bounded. This convergence of the first moment was previously proved in [6]. Theorem 1 identifies the full limiting distribution and, more generally, the evolution of the genetic contribution throughout the fixation of the advantageous type.

Remark 2. It is useful to compare Theorem 1 with the regime of a macroscopic initial condition. Suppose that aN individuals are initially advantaged, with $a \in (0,1)$, and let B_n (resp. C_n) denote the total genetic contribution of this initial advantaged population carried by advantaged (resp. disadvantaged) individuals. In this regime the natural scaling is N, and the limit is deterministic. More precisely, the results of [4] suggest (by taking the $s \to \infty$ limit there that:

$$\sup_{y\in[a,1]} \left| \frac{1}{N} \begin{pmatrix} B_{S_{\lfloor yN \rfloor}} \\ C_{S_{\lfloor yN \rfloor}} \end{pmatrix} - \begin{pmatrix} b_a(y) \\ c_a(y) \end{pmatrix} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0,$$

where

$$b_a(y) = \sqrt{ay} + y\sqrt{a}\,(\sqrt{y} - \sqrt{a}), \qquad c_a(y) = (1-y)\sqrt{a}\,(\sqrt{y} - \sqrt{a}).$$

Formally substituting $a = 1/N$ into these profiles recovers the $\sqrt{N}$-scale and the shape $\sqrt{y}(1\pm y)$, but neither the prefactor $2/\sqrt{\pi}$ nor the random multiplier W. These are both created by the microscopic seeding phase. See also Remark 3 for a thorough comparison with [4].

3 Proof strategy

The proof is organized around a triangular drift structure for the process observed at the hitting times S_k, i.e. when the advantaged population first reaches size k. Writing

$$X_k := \overline{B}_k - \overline{C}_k, \tag{3}$$

with

$$\overline{B}_k = \frac{B_{S_k}}{k} \qquad \text{and} \qquad \overline{C}_k = \frac{C_{S_k}}{N-k},$$

an exact computation over the block $[S_k, S_{k+1}]$ gives (Proposition 3)

$$\mathbb{E}\left[\begin{pmatrix} X_{k+1} \\ \overline{C}_{k+1} \end{pmatrix} \middle| \mathcal{F}_{S_k} \right] = \begin{pmatrix} \alpha_{k,N} & 0 \\ (2N+1)^{-1} & 1 \end{pmatrix} \begin{pmatrix} X_k \\ \overline{C}_k \end{pmatrix}.$$

Thus the first coordinate evolves autonomously at the level of conditional drift, while the second is driven by the first.

After multiplication by an explicit deterministic factor $a_{k,N}$, the first coordinate becomes the martingale

$$M_k^{(N)} := a_{k,N} X_k.$$

At the same time, the second coordinate admits the exact decomposition

$$\overline{C}_{k+1} = \overline{C}_k + \frac{1}{2N+1} a_{k,N}^{-1} M_k^{(N)} + \Delta R_{k+1}, \qquad \text{where} \qquad \mathbb{E}\left[\Delta R_{k+1} \,|\, \mathcal{F}_{S_k}\right] = 0. \tag{4}$$

These two identities contain the basic mechanism of the proof: the random amplitude is carried by $M^{(N)}$, while the macroscopic profile is obtained by integrating this mode through the evolution of $\overline{C}$.

We choose a mesoscopic sequence $k_0 = k_0(N)$, such that $1 << k_0 << N$ when N goes to infinity. If paternal contributions are discarded, the maternal genealogy of the advantaged individuals is exactly a random recursive tree; see Lemma 4. The associated genetic weight, after normalization, is an L^2-bounded martingale whose limit is denoted W; see Proposition 4. The recursive decomposition of the tree gives a smoothing equation for W, whose L^2-solution with mean 1 is unique and strictly positive; see Proposition 5. The paternal contribution up to level k_0 is proven to be negligible (Lemma 6), and the disadvantaged contribution is negligible at the same scale by Lemma 2. Consequently, Proposition 6 gives

$$M_{k_0}^{(N)} \Longrightarrow W.$$

The second phase propagates this random entrance value through the macroscopic regime. The one-step quadratic-variation estimate of Proposition 7, together with the a priori bounds of Lemma 2, yield (Lemma 8).

$$\sup_{k_0 \leqslant k \leqslant (1-\varepsilon)N} \left| M_k^{(N)} - M_{k_0}^{(N)} \right| \xrightarrow[N\to\infty]{L^2} 0.$$

Obtaining this estimate is the main technical effort of the proof. It contrasts with the N-scale regime of a macroscopic initial condition studied in [4], where much rougher bounds suffice.

Proposition 8 shows in parallel that the accumulated centred remainder in the $\overline{C}$-decomposition is negligible. Hence, throughout the macroscopic phase,

$$M_k^{(N)} \approx W, \qquad X_k = a_{k,N}^{-1} M_k^{(N)} \approx \frac{2W}{\sqrt{\pi k}}.$$

Summing the drift of $\overline{C}$ in Equation (4) therefore gives, heuristically,

$$\overline{C}_k \approx \frac{W}{2N} \sum_{j \leqslant k} a_{j,N}^{-1} \approx \frac{2W}{\sqrt{\pi}} \frac{\sqrt{k}}{N}.$$

Thus, for $k \sim yN$,

$$C_k = (N-k)\overline{C}_k \approx \frac{2}{\sqrt{\pi}} W \sqrt{N}\, \sqrt{y}\, (1-y),$$

while using Equation(3) gives

$$B_k = k(\overline{C}_k + X_k) \approx \frac{2}{\sqrt{\pi}} W \sqrt{N}\, \sqrt{y}\, (1+y).$$

This is precisely the deterministic profile appearing in Theorem 1; the rigorous uniform approximation is established in Proposition 9.

It remains to connect the bulk approximation to the endpoints. Near 0, Lemma 10 shows that both contributions are negligible on the $\sqrt{N}$-scale before the mesoscopic entrance level. Near 1, the average disadvantaged contribution is no longer sufficiently controlled to continue the quadratic-variation argument. A separate terminal estimate, Lemma 11, shows instead that B_k varies little and that C_k remains small near fixation. Combining these endpoint estimates with Proposition 9 yields directly the uniform approximation

$$\sup_{0 \leqslant y \leqslant 1} \left| \frac{1}{\sqrt{N}} \begin{pmatrix} B_{k_N(y)} \\ C_{k_N(y)} \end{pmatrix} - M_{k_0}^{(N)} \begin{pmatrix} b(y) \\ c(y) \end{pmatrix} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

The theorem then follows from the convergence of the entrance value $M_{k_0}^{(N)}$ and Slutsky's lemma.

Remark 3. The macroscopic phase could also be treated as in [4], using a fluid-limit argument in the spirit of Kurtz [14]. Indeed, setting

$$\left(\frac{Y_{\lfloor Nt \rfloor}}{N}, \frac{B_{\lfloor Nt \rfloor}}{\sqrt{N}}, \frac{C_{\lfloor Nt \rfloor}}{\sqrt{N}} \right) \approx \big(y(t), b(t), c(t) \big),$$

the conditional drift identities lead to the closed system

$$\begin{cases} \dot{y} = y, \\ \dot{b} = \dfrac{1}{2}\big((1+y)b + yc\big), \\ \dot{c} = \dfrac{1-y}{2} b - \dfrac{y(3-y)}{2(1-y)} c, \end{cases}$$

A direct fluid approximation from our initial state would, however, be degenerate: a single initially advantaged individual corresponds in the limit to the initial condition $(y, b, c) = (0, 0, 0)$, whose solution remains identically zero. One could instead start the fluid approximation once the advantaged population has reached a fixed proportion $a > 0$, as in [4]. From that point on, concentration indeed holds and could be proved by quadratic-variation estimates of the same type as those used below; using y as the time variable, this system then yields the same deterministic profiles as those we obtain. This would still leave the separate problem of gluing the fluid regime to the microscopic random seeding phase near $y = 0$.

We find the martingale $M_k^{(N)}$ more convenient for our purpose, since the same quantity captures the random amplitude created in the initial regime and transports it through the macroscopic one.

4 Embedded dynamics and the fundamental martingale

All quantities depend on N, but this dependence is suppressed unless it is needed.

Generic finite constants are denoted by K, or by K_ε when they may depend on ε, and may change from line to line.

4.1 One-step and embedded dynamics

For $Y_n < N$, set

$$\overline{B}_n := \frac{B_n}{Y_n}, \qquad \overline{C}_n := \frac{C_n}{N - Y_n}, \qquad X_n := \overline{B}_n - \overline{C}_n.$$

The quantities $\overline{B}_n$ and $\overline{C}_n$ are respectively equal to the mean genetic contribution of the initial ancestor carried by advantaged (resp. disadvantaged) individuals living at time n. The quantity X_n therefore measures a form of imbalance between the proportion of genome of advantaged vs disadvantaged individuals living at time n coming from the initially advantaged individual. This quantity is initially equal to 1 and due to the even parental genome transmission we expect this difference to tend to go to 0. The following propositions give first the autonomous dynamics of the number of advantaged individuals $(Y_n)_{n\geqslant 0}$ and second that of $(B_n)_{n\geqslant 0}$, $(C_n)_{n\geqslant 0}$ and the dynamics of the expectation of their rescaled versions $(\overline{B}_n)_{n\geqslant 0}$, $(\overline{C}_n)_{n\geqslant 0}$ and of the difference $(X_n)_{n\geqslant 0}$. The two propositions are proved together for more simplicity.

Proposition 1 (Autonomous dynamics of Y)**.** *Let $w_n(i) := A_n(i, 1)$. Conditionally on $\mathcal{F}_n$, as long as $Y_n < N$,*

$$\mathbb{P}(Y_{n+1} = Y_n + 1 \mid \mathcal{F}_n) = \frac{Y_n}{N}, \qquad \mathbb{P}(Y_{n+1} = Y_n \mid \mathcal{F}_n) = 1 - \frac{Y_n}{N}.$$

Consequently, for $1 \leqslant k \leqslant N - 1$,

$$S_{k+1} - S_k - 1 \sim \mathrm{Geom}_{\mathbb{Z}_+}\left(\frac{k}{N}\right)$$

conditionally on $\mathcal{F}_{S_k}$.

Once $Y_n = N$ the model all individuals are advantaged and the model becomes the neutral version studied in [5]. Here we do not consider the dynamics of the considered processes after that time.

Proposition 2 (One-step dynamics of B and C)**.** *Let us set $Y_n = k$ and $d = N - Y_n = N - k$. On the stable event $\{Y_{n+1} = Y_n = k\}$, one has*

$$B_{n+1} = B_n, \qquad C_{n+1} = C_n + \frac{w_n(\mu_n) + w_n(\pi_n)}{2} - w_n(\kappa_n),$$

where, conditionally on $\mathcal{F}_n$ and on this event, μ_n and κ_n are uniform on the disadvantaged class, while π_n is uniform on I_N. In particular,

$$\mathbb{E}[\overline{C}_{n+1} - \overline{C}_n \mid \mathcal{F}_n,\, Y_{n+1} = Y_n = k] = \frac{k}{2Nd} X_n, \tag{5}$$

and

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n,\, Y_{n+1} = Y_n = k] = \left(1 - \frac{k}{2Nd}\right) X_n.$$

On the jump event $\{Y_{n+1} = Y_n + 1\}$, one has

$$B_{n+1} = B_n + \frac{w_n(\mu_n) + w_n(\pi_n)}{2}, \qquad C_{n+1} = C_n - w_n(\kappa_n),$$

where, conditionally on $\mathcal{F}_n$ and on this event, μ_n is uniform on the advantaged class, κ_n is uniform on the disadvantaged class, and π_n is uniform on I_N. In particular, if $k \leqslant N-2$, then

$$\mathbb{E}[\overline{C}_{n+1} - \overline{C}_n \mid \mathcal{F}_n,\, Y_{n+1} = Y_n + 1 = k+1] = 0, \tag{6}$$

and

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n,\, Y_{n+1} = Y_n + 1 = k+1] = \left(1 - \frac{d}{2N(k+1)}\right) X_n.$$

Proof. The event $Y_{n+1} = Y_n + 1$ occurs exactly when the mother μ_n is advantaged, which has conditional probability Y_n/N. Otherwise the number of advantaged individuals is unchanged. This also gives the conditional geometric law of $S_{k+1} - S_k - 1$.

The formulae for B_{n+1} and C_{n+1} follow directly from the update rule for A_n given in Equation (1) and the definitions of B_n and C_n in Equation (2).

On a stable step, $Y_{n+1} = k$ and $N - Y_{n+1} = d$, and

$$\overline{C}_{n+1} - \overline{C}_n = \frac{1}{d}\left(\frac{w_n(\mu_n) + w_n(\pi_n)}{2} - w_n(\kappa_n)\right).$$

Taking conditional expectations gives

$$\begin{aligned}\mathbb{E}[\overline{C}_{n+1} - \overline{C}_n \mid \mathcal{F}_n,\, Y_{n+1} = Y_n = k] &= \frac{1}{d}\left(\frac{1}{2}\overline{C}_n + \frac{1}{2}\frac{k\overline{B}_n + d\overline{C}_n}{N} - \overline{C}_n\right) \\ &= \frac{k}{2Nd}\left(\overline{B}_n - \overline{C}_n\right) = \frac{k}{2Nd} X_n.\end{aligned}$$

Since $B_{n+1} = B_n$ on a stable step, $\overline{B}_{n+1} = \overline{B}_n$, and therefore

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n,\, Y_{n+1} = Y_n = k] = \left(1 - \frac{k}{2Nd}\right) X_n.$$

On a jump step, assuming $k \leqslant N-2$, the disadvantaged average satisfies

$$\overline{C}_{n+1} - \overline{C}_n = \frac{\overline{C}_n - w_n(\kappa_n)}{d-1}.$$

Since κ_n is uniform on the disadvantaged class, its conditional expectation is zero. For the advantaged average,

$$\overline{B}_{n+1} = \frac{k}{k+1}\overline{B}_n + \frac{w_n(\mu_n) + w_n(\pi_n)}{2(k+1)}.$$

Thus, taking conditional expectations gives

$$\begin{aligned}\mathbb{E}[\overline{B}_{n+1} - \overline{B}_n \mid \mathcal{F}_n,\, Y_{n+1} = Y_n + 1 = k+1] &= -\frac{1}{k+1}\overline{B}_n + \frac{1}{2(k+1)}\left(\overline{B}_n + \frac{k\overline{B}_n + d\overline{C}_n}{N}\right) \\ &= -\frac{d}{2N(k+1)}\left(\overline{B}_n - \overline{C}_n\right).\end{aligned}$$

Combining this with the zero drift of $\overline{C}$ at the jump gives

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n,\, Y_{n+1} = Y_n + 1] = \left(1 - \frac{d}{2N(k+1)}\right) X_n.$$

□

Denote by

$$\alpha_{k,N} := \frac{2Nk + N + k}{(2N+1)(k+1)} = \frac{k + 1 - b_N}{k+1}, \qquad b_N := \frac{N+1}{2N+1},$$

and set

$$a_{1,N} := 1, \qquad a_{k,N} := \prod_{\ell=1}^{k-1} \alpha_{\ell,N}^{-1}, \qquad 2 \leqslant k \leqslant N. \tag{7}$$

Recall the hitting times S_k introduced in Section 2. From now on, an index n refers to calendar time, whereas an index k refers to the embedded chain observed when the advantaged population has size k. Thus

$$B_k := B_{S_k}, \qquad C_k := C_{S_k}, \qquad 1 \leqslant k \leqslant N,$$

and, for $1 \leqslant k \leqslant N-1$,

$$\overline{B}_k := \frac{B_k}{k}, \qquad \overline{C}_k := \frac{C_k}{N-k}, \qquad X_k := \overline{B}_k - \overline{C}_k.$$

Accordingly, set

$$M_k^{(N)} := a_{k,N} X_k, \qquad 1 \leqslant k \leqslant N-1. \tag{8}$$

Thus B_k and C_k always denote the values at the hitting time S_k, rather than at deterministic calendar time k.

For later use, we also introduce the block notations

$$d := N - k, \qquad q := \frac{d}{N}, \qquad \lambda := \frac{k}{2Nd}, \qquad L := S_{k+1} - S_k - 1. \tag{9}$$

Thus, L is the number of stable steps before the jump $k \to k+1$, and, conditionally on $\mathcal{F}_{S_k}$, from Proposition 1,

$$\mathbb{P}(L = \ell \mid \mathcal{F}_{S_k}) = q^\ell (1-q), \qquad \ell \geqslant 0.$$

On $\{L \geqslant j\}$, let $\overline{C}^{(j)}$ and $X^{(j)}$ denote the values of $\overline{C}$ and X after the first j stable steps of the block (i.e. at time $S_k + j$), and set

$$C^{(j)} := d\,\overline{C}^{(j)}.$$

In particular,

$$\overline{C}^{(0)} = \overline{C}_k, \qquad X^{(0)} = X_k, \qquad C^{(0)} = C_k.$$

Proposition 3 (Embedded drift structure and fundamental martingale)**.** *For every $N \geqslant 3$ and $1 \leqslant k \leqslant N-2$,*

$$\mathbb{E}\left[\begin{pmatrix} X_{k+1} \\ \overline{C}_{k+1} \end{pmatrix} \middle| \mathcal{F}_{S_k}\right] = \begin{pmatrix} \alpha_{k,N} & 0 \\ (2N+1)^{-1} & 1 \end{pmatrix} \begin{pmatrix} X_k \\ \overline{C}_k \end{pmatrix}. \tag{10}$$

Consequently, $\left(M_k^{(N)}\right)_{1 \leqslant k \leqslant N-1}$ is a martingale with respect to $(\mathcal{F}_{S_k})_{1 \leqslant k \leqslant N-1}$.

Moreover, setting

$$\Delta R_{k+1} := \overline{C}_{k+1} - \overline{C}_k - \frac{1}{2N+1} X_k,$$

one has the decomposition

$$\overline{C}_{k+1} = \overline{C}_k + \frac{1}{2N+1} X_k + \Delta R_{k+1} = \overline{C}_k + \frac{1}{2N+1} a_{k,N}^{-1} M_k^{(N)} + \Delta R_{k+1}, \tag{11}$$

where

$$\mathbb{E}\left[\Delta R_{k+1} \,|\, \mathcal{F}_{S_k}\right] = 0.$$

Proof. Fix $1 \leqslant k \leqslant N-2$. By Proposition 1, conditionally on $\mathcal{F}_{S_k}$,

$$\mathbb{P}(L = \ell \mid \mathcal{F}_{S_k}) = q^\ell (1-q), \qquad \ell \geqslant 0.$$

By Proposition 2, during each stable step,

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n,\, Y_{n+1} = Y_n = k] = (1-\lambda)\, X_n,$$

while at the final jump $k \to k+1$,

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n,\, Y_{n+1} = Y_n + 1 = k+1] = \left(1 - \frac{d}{2N(k+1)}\right) X_n.$$

Therefore

$$\begin{aligned} \mathbb{E}\left[X_{k+1} \,|\, \mathcal{F}_{S_k}\right] &= \sum_{\ell \geqslant 0} q^\ell (1-q)(1-\lambda)^\ell \left(1 - \frac{d}{2N(k+1)}\right) X_k \\ &= \alpha_{k,N} X_k. \end{aligned} \tag{12}$$

We next compute the drift of $\overline{C}$. Since the event that the next update is stable is independent of the current genetic configuration, Proposition 2 gives, by induction,

$$\mathbb{E}\left[X^{(j)} \,\middle|\, \mathcal{F}_{S_k},\, L \geqslant j\right] = (1-\lambda)^j X_k.$$

The final jump has zero conditional drift for $\overline{C}$, therefore

$$\begin{aligned} \mathbb{E}\left[\overline{C}_{k+1} - \overline{C}_k \,|\, \mathcal{F}_{S_k}\right] &= \mathbb{E}\left[\overline{C}^{(L)} - \overline{C}_k \,\middle|\, \mathcal{F}_{S_k}\right] \\ &= -\mathbb{E}\left[X^{(L)} - X_k \,\middle|\, \mathcal{F}_{S_k}\right] \\ &= -\left[\left(\sum_{j \geqslant 0} q^j (1-q)(1-\lambda)^j\right) - 1\right] X_k \\ &= \frac{q\lambda}{1 - q(1-\lambda)} X_k \\ &= \frac{1}{2N+1} X_k. \end{aligned} \tag{13}$$

Together with the preceding identity (12) for X_{k+1}, this proves (10).

Finally, since

$$a_{k+1,N} = a_{k,N} \alpha_{k,N}^{-1},$$

we obtain

$$\mathbb{E}\left[M_{k+1}^{(N)} \,\middle|\, \mathcal{F}_{S_k}\right] = a_{k+1,N} \alpha_{k,N} X_k = M_k^{(N)}$$

therefore $(M_k)_{1 \leqslant k \leqslant N-1}$ is a martingale with respect $(\mathcal{F}_{S_k})_{1 \leqslant k \leqslant N-1}$. The centering of ΔR_{k+1} follows directly from (13) and the second part of the decomposition (11) follows from the definition (8). □

The triangular form of (10) is the basic structural feature used throughout the proof: the conditional expectation of the imbalance X_k evolves autonomously after deterministic renormalization, while the dynamics of the disadvantaged average $\overline{C}_k$ is driven by this imbalance X_k.

Remark 4. The martingale $(M_k)_{1 \leqslant k \leqslant N-1}$ is defined only up to $k = N-1$, since $\overline{C}_N$ is not defined at fixation. The terminal step $N-1 \to N$ will be handled separately in Section 7.

4.2 Deterministic and first-moment estimates

We first collect the estimates on the deterministic coefficients introduced in (7).

Lemma 1 (Deterministic estimates for $a_{k,N}$)**.** *There exists a constant $K < \infty$ such that, for every $N \geqslant 2$ and every $1 \leqslant k \leqslant N$,*

$$K^{-1}\sqrt{k} \leqslant a_{k,N} \leqslant K\sqrt{k}. \tag{14}$$

Moreover, if $k_0 = k_0(N) \leqslant N$ satisfies $k_0 \to \infty$ when N goes to infinity, then

$$\sup_{k_0 \leqslant k \leqslant N} \left| \sqrt{k}\, a_{k,N}^{-1} - \frac{2}{\sqrt{\pi}} \right| \xrightarrow[N\to\infty]{} 0. \tag{15}$$

If in addition $k_0/N \to 0$ when N goes to infinity, then

$$\sqrt{N} \sup_{k_0 \leqslant k \leqslant N} \left| \frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} - \frac{2}{\sqrt{\pi}} \frac{\sqrt{k}}{N} \right| \longrightarrow 0 \qquad \text{as } N \to \infty. \tag{16}$$

Proof. From (7),

$$a_{k,N} = \Gamma(2-b_N) \frac{\Gamma(k+1)}{\Gamma(k+1-b_N)}, \qquad b_N = \frac{1}{2} + \frac{1}{4N+2}.$$

By Gautschi's inequality [12], for every $k \geqslant 1$ and $0 < b < 1$,

$$1 \leqslant \frac{\Gamma(k+1)}{k^b \Gamma(k+1-b)} \leqslant (1+k^{-1})^b \leqslant 1 + k^{-1}.$$

Consequently, since $b_N \in [1/2, 3/5]$,

$$\frac{\Gamma(k+1)}{\Gamma(k+1-b_N)} = k^{b_N}\big(1 + O(k^{-1})\big),$$

where the implicit constant is independent of N. In particular, this ratio is bounded above and below by positive constant multiples of k^{b_N}, uniformly in $N \geqslant 2$ and $k \geqslant 1$. Moreover, $\Gamma(2-b_N)$ is bounded above and away from zero, while

$$k^{b_N} = \sqrt{k} \exp\left(\frac{\log k}{4N+2}\right)$$

and, for $k \leqslant N$,

$$1 \leqslant \exp\left(\frac{\log k}{4N+2}\right) \leqslant \exp\left(\frac{\log N}{4N+2}\right).$$

The last expression is bounded uniformly in $N \geqslant 2$, and (14) follows.

Moreover,

$$\sqrt{k}\, a_{k,N}^{-1} = \frac{k^{1/2-b_N}}{\Gamma(2-b_N)}\big(1 + O(k^{-1})\big).$$

Uniformly for $k_0 \leqslant k \leqslant N$,

$$\left|\log k^{1/2-b_N}\right| = \frac{\log k}{4N+2} \leqslant \frac{\log N}{4N+2} \longrightarrow 0,$$

and the $O(k^{-1})$ term is $o(1)$ since $k_0 \to \infty$. Finally,

$$\Gamma(2-b_N) \longrightarrow \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

This proves (15).

Set

$$\varepsilon_N := \sup_{k_0 \leqslant j \leqslant N} \left| \sqrt{j}\, a_{j,N}^{-1} - \frac{2}{\sqrt{\pi}} \right|,$$

so that $\varepsilon_N \to 0$. Uniformly for $k_0 \leqslant k \leqslant N$,

$$\left| \sum_{j=k_0}^{k-1} a_{j,N}^{-1} - \frac{2}{\sqrt{\pi}} \sum_{j=k_0}^{k-1} j^{-1/2} \right| \leqslant \varepsilon_N \sum_{j=k_0}^{k-1} j^{-1/2} \leqslant 2\varepsilon_N \sqrt{N}.$$

Furthermore, by comparison with the corresponding integral,

$$\sum_{j=k_0}^{k-1} j^{-1/2} = 2\big(\sqrt{k} - \sqrt{k_0}\big) + O(k_0^{-1/2})$$

uniformly in $k \geqslant k_0$. Consequently,

$$\sqrt{N} \sup_{k_0 \leqslant k \leqslant N} \left| \frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} - \frac{2}{\sqrt{\pi}} \frac{\sqrt{k}}{N} \right| \leqslant C\varepsilon_N + C\sqrt{\frac{k_0}{N}} + \frac{C}{N} + \frac{C}{\sqrt{Nk_0}},$$

which tends to zero because $k_0/N \to 0$. This proves (16). □

Lemma 2 (First-moment calculations and bounds)**.** *There exists a constant $K < \infty$ such that, for every $1 \leqslant k \leqslant N-1$,*

$$\mathbb{E}[X_k] = a_{k,N}^{-1} \leqslant \frac{K}{\sqrt{k}}, \tag{17}$$

while

$$\mathbb{E}[\overline{C}_k] = \frac{k a_{k,N}^{-1} - 1}{N} \qquad \textit{and} \qquad \mathbb{E}[\overline{B}_k] = \frac{(N+k) a_{k,N}^{-1} - 1}{N}.$$

In particular,

$$\mathbb{E}[\overline{C}_k] \leqslant K\frac{\sqrt{k}}{N}, \qquad \mathbb{E}[\overline{B}_k + \overline{C}_k] \leqslant \frac{K}{\sqrt{k}}, \qquad \textit{and} \qquad \mathbb{E}[B_k + C_k] \leqslant K\sqrt{k}.$$

Proof. By Proposition 3,

$$\mathbb{E}[M_k^{(N)}] = \mathbb{E}[M_1^{(N)}] = 1.$$

Since $M_k^{(N)} = a_{k,N} X_k$,

$$\mathbb{E}[X_k] = a_{k,N}^{-1}.$$

The inequality given in (17) therefore follows directly from Lemma 1.

Since $\overline{C}_1 = 0$, taking expectations in (11) and summing gives

$$\mathbb{E}[\overline{C}_k] = \frac{1}{2N+1} \sum_{j=1}^{k-1} a_{j,N}^{-1}.$$

What is more by definition,

$$a_{k+1,N}^{-1} = \frac{k+1-b_N}{k+1} a_{k,N}^{-1} \qquad \text{which gives} \qquad \frac{a_{k,N}^{-1}}{2N+1} = \frac{k+1}{N} a_{k+1,N}^{-1} - \frac{k}{N} a_{k,N}^{-1}.$$

Hence

$$\mathbb{E}[\overline{C}_k] = \frac{k a_{k,N}^{-1} - 1}{N} \leqslant K\frac{\sqrt{k}}{N}.$$

Next since $\overline{B}_k = X_k + \overline{C}_k$,

$$\mathbb{E}[\overline{B}_k] = \frac{(N+k)a_{k,N}^{-1} - 1}{N}.$$

and we obtain using (17) that

$$\mathbb{E}[\overline{B}_k + \overline{C}_k] \leqslant \frac{K}{\sqrt{k}}.$$

Finally, using the non-negativity of the class averages $\overline{B}_k$ and $\overline{C}_k$,

$$\begin{aligned}\mathbb{E}[B_k + C_k] &= k\,\mathbb{E}[\overline{B}_k] + (N-k)\mathbb{E}[\overline{C}_k] \\ &\leqslant K\sqrt{k} + K\sqrt{k},\end{aligned}$$

which proves the last inequality of Lemma 2 after changing K. □

5 Initial window and random recursive tree coupling

The random amplitude is generated while the advantaged population is still small. The key observation is that, if paternal genetic input is discarded, each new advantaged individual inherits half the contribution of its advantaged mother, while that mother is uniform among the current advantaged individuals. The resulting maternal genealogy is therefore a random recursive tree. We first analyse the corresponding weighted tree, and then show that the paternal contribution is negligible on the mesoscopic scale i.e. as long as $k \ll N$.

5.1 Random recursive tree (RRT)

Let $(\mathcal{T}_k)_{k\geqslant 1}$ be the random recursive tree rooted at 1: for each $k \geqslant 2$, vertex k is attached to a parent U_k chosen uniformly in $\{1, \dots, k-1\}$, independently conditionally on the past. We denote by $\mathrm{depth}_k(v)$ the distance from v to the root in $\mathcal{T}_k$, and set

$$B_k^\star := \sum_{v\in\mathcal{T}_k} 2^{-\,\mathrm{depth}_k(v)}, \qquad \gamma_k := \prod_{\ell=1}^{k-1}\left(1 + \frac{1}{2\ell}\right), \qquad M_k^\star := \frac{B_k^\star}{\gamma_k}, \tag{18}$$

where the empty product is understood to be equal to 1, so that $\gamma_1 = 1$.

Remark 5. This martingale is a particular instance of martingales studied by Drmota and Hwang [9, Section 3.3], see also Fuchs, Hwang and Neininger [11, Equation (4) and Proposition 1] for the characterisation of their limit. We nevertheless retain our original derivation below, both to keep the paper self-contained and to establish the properties in the precise form needed here.

Proposition 4 (RRT martingale, L^2 boundedness and limit)**.** *The process $(M_k^\star)_{k\geqslant 1}$ is a non-negative martingale, and $\mathbb{E}[M_k^\star] = 1$ for every $k \geqslant 1$. What is more,*

$$\sup_{k\geqslant 1} \mathbb{E}\big[(M_k^\star)^2\big] < \infty.$$

Consequently $M_k^\star \to M_\infty^\star$ almost surely and in L^2, and $\mathbb{E}[M_\infty^\star] = 1$.

We first prove the following lemma giving the asymptotics of γ_k.

Lemma 3 (Asymptotics of γ_k)**.** *As $k \to \infty$,*

$$\frac{\gamma_k}{\sqrt{k}} \longrightarrow \frac{2}{\sqrt{\pi}}.$$

Proof. Using the Gamma function,

$$\gamma_k = \prod_{\ell=1}^{k-1} \left(1 + \frac{1}{2\ell}\right) = \frac{\Gamma(k+\frac{1}{2})}{\Gamma(\frac{3}{2})\Gamma(k)} = \frac{2}{\sqrt{\pi}} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(k)}.$$

The result follows from Stirling's formula. □

Proof of Proposition 4. Let $\mathcal{H}_k := \sigma(\mathcal{T}_k)$, and let U_{k+1} be the parent of vertex $k+1$. Then

$$B^\star_{k+1} = B^\star_k + \xi_{k+1}, \qquad \xi_{k+1} := 2^{-(\mathrm{depth}_k(U_{k+1})+1)}.$$

Conditionally on $\mathcal{H}_k$, U_{k+1} is uniform on $\mathcal{T}_k$, hence

$$\mathbb{E}[\xi_{k+1} \mid \mathcal{H}_k] = \frac{1}{k} \sum_{v \in \mathcal{T}_k} 2^{-(\mathrm{depth}_k(v)+1)} = \frac{1}{2k} B^\star_k.$$

Therefore

$$\mathbb{E}[B^\star_{k+1} \mid \mathcal{H}_k] = \left(1 + \frac{1}{2k}\right) B^\star_k,$$

and the definition of γ_k gives

$$\mathbb{E}[M^\star_{k+1} \mid \mathcal{H}_k] = M^\star_k.$$

Since $M^\star_1 = 1$, the first sentence of Proposition 4 is proved. Now let

$$R^\star_k := \sum_{v \in \mathcal{T}_k} 4^{-\mathrm{depth}_k(v)}.$$

Then

$$\mathbb{E}[\xi^2_{k+1} \mid \mathcal{H}_k] = \frac{1}{4k} R^\star_k,$$

therefore

$$\mathbb{E}\big[(B^\star_{k+1})^2 \mid \mathcal{H}_k\big] = (B^\star_k)^2 \Big(1 + \frac{1}{k}\Big) + \frac{1}{4k} R^\star_k.$$

Since

$$\gamma_{k+1} = \gamma_k \Big(1 + \frac{1}{2k}\Big),$$

we get

$$\mathbb{E}\big[(M^\star_{k+1})^2 \mid \mathcal{H}_k\big] = (M^\star_k)^2 \frac{1 + \frac{1}{k}}{(1 + \frac{1}{2k})^2} + \frac{1}{4k} \frac{R^\star_k}{\gamma_k^2 (1 + \frac{1}{2k})^2}.$$

Now

$$1 + \frac{1}{k} \leqslant \Big(1 + \frac{1}{2k}\Big)^2, \qquad \text{and} \qquad R^\star_k \leqslant B^\star_k.$$

Thus, after taking expectations,

$$\mathbb{E}\big[(M^\star_{k+1})^2\big] \leqslant \mathbb{E}\big[(M^\star_k)^2\big] + \frac{1}{4k}\, \mathbb{E}\left[\frac{B^\star_k}{\gamma_k^2}\right].$$

Since $B^\star_k = \gamma_k M^\star_k$ and $\mathbb{E}[M^\star_k] = 1$,

$$\mathbb{E}\big[(M^\star_{k+1})^2\big] \leqslant \mathbb{E}\big[(M^\star_k)^2\big] + \frac{1}{4k\gamma_k}.$$

By Lemma 3, there exists a constant $K < \infty$ such that

$$\frac{1}{\gamma_k} \leqslant \frac{K}{\sqrt{k}}, \qquad \text{for every} \qquad k \geqslant 1.$$

Consequently

$$\mathbb{E}\big[(M_{k+1}^\star)^2\big] \leqslant \mathbb{E}\big[(M_k^\star)^2\big] + \frac{K}{k^{3/2}}.$$

Iterating gives

$$\sup_{n\geqslant 1} \mathbb{E}\big[(M_n^\star)^2\big] \leqslant \mathbb{E}\big[(M_1^\star)^2\big] + K\sum_{k\geqslant 1}\frac{1}{k^{3/2}} < \infty.$$

Since $(M_k^\star)_{k\geqslant 1}$ is an L^2-bounded martingale, it converges almost surely and in L^2 to a limit $M_\infty^\star$. Finally, L^1-convergence gives

$$\mathbb{E}[M_\infty^\star] = \lim_{k\to\infty} \mathbb{E}[M_k^\star] = 1.$$

□

Proposition 5 (Smoothing equation for the RRT limit)**.** *The limiting random variable $M_\infty^\star$ is the unique L^2 solution with mean 1 of the equation in law*

$$W \overset{d}{=} \sqrt{U}\, W^{(1)} + \frac{1}{2}\sqrt{1-U}\, W^{(2)}, \tag{19}$$

where $U \sim \mathrm{Unif}[0,1]$, and $W^{(1)}, W^{(2)}$ are independent copies of W, independent of U. What is more, this solution satisfies $M_\infty^\star > 0$ almost surely.

Proof. Remove the edge between vertices 1 and 2 in the recursive tree $\mathcal{T}_n$. This gives two rooted subtrees: the component containing vertex 1, denoted by $\mathcal{T}_n^{(1)}$, and the component rooted at vertex 2, denoted by $\mathcal{T}_n^{(2)}$. Let

$$K_n := |\mathcal{T}_n^{(1)}|.$$

By the standard Pólya-urn description of the two branches born from the first split ([17]),

$$K_n \sim \mathrm{Unif}\{1,\dots,n-1\}.$$

Moreover, conditionally on $K_n = k$, the two rooted trees $\mathcal{T}_n^{(1)}$ and $\mathcal{T}_n^{(2)}$, after relabelling their vertices in increasing order, are independent recursive trees with respective sizes k and $n-k$.

The component containing vertex 1 keeps the same depths as in $\mathcal{T}_n$, whereas every vertex in the component rooted at vertex 2 has its depth increased by one. Hence

$$B_n^\star \overset{d}{=} B_{K_n}^{\star,(1)} + \frac{1}{2} B_{n-K_n}^{\star,(2)},$$

where $B^{\star,(1)}$ and $B^{\star,(2)}$ are independent processes with same law as $B^\star$. Therefore

$$M_n^\star \overset{d}{=} \frac{\gamma_{K_n}}{\gamma_n} M_{K_n}^{\star,(1)} + \frac{1}{2}\frac{\gamma_{n-K_n}}{\gamma_n} M_{n-K_n}^{\star,(2)}, \tag{20}$$

where the two RRT processes on the right-hand side are independent and independent of K_n.

We now realize the right-hand side of (20) on a common probability space as follows. Let $U \sim \mathrm{Unif}[0,1]$, independent of the two RRT processes, and set

$$\widetilde{K}_n := 1 + \lfloor (n-1)U \rfloor.$$

For every n, $\widetilde{K}_n$ has the same law as K_n, so that the right-hand side of (20) is unchanged in distribution when K_n is replaced by $\widetilde{K}_n$. Moreover,

$$\frac{\widetilde{K}_n}{n} \longrightarrow U \qquad \text{almost surely.}$$

Since $U \in (0,1)$ almost surely, both $\widetilde{K}_n$ and $n-\widetilde{K}_n$ tend to infinity almost surely. Proposition 4, applied to the two independent RRT processes, gives

$$M_m^{\star,(1)} \longrightarrow M_\infty^{*,1}, \qquad M_m^{\star,(2)} \longrightarrow M_\infty^{*,2} \qquad \text{almost surely as } m \to \infty,$$

where $M_\infty^{*,1}$ and $M_\infty^{*,2}$ are independent copies of M_∞^*, independent of U. Hence, along the random indices $\widetilde{K}_n$ and $n-\widetilde{K}_n$, and using Lemma 3,

$$\left(\frac{\gamma_{\widetilde{K}_n}}{\gamma_n}, \frac{\gamma_{n-\widetilde{K}_n}}{\gamma_n}, M_{\widetilde{K}_n}^{\star,(1)}, M_{n-\widetilde{K}_n}^{\star,(2)}\right) \longrightarrow \left(\sqrt{U}, \sqrt{1-U}, M_\infty^{*,1}, M_\infty^{*,2}\right)$$

almost surely. Therefore the right-hand side of (20), realized with $\widetilde{K}_n$ in place of K_n, converges almost surely to

$$\sqrt{U}\, M_\infty^{*,1} + \frac{1}{2}\sqrt{1-U}\, M_\infty^{*,2}.$$

Since this realization has the same distribution as the right-hand side of (20), while $M_n^\star \Rightarrow M_\infty^*$ by Proposition 4, we obtain that M_∞^* is solution to (19).

We now prove uniqueness using a contraction argument. Let Φ be the map on the set $\mathcal{P}_2(\mathbb{R})$ of probability measures on $\mathbb{R}$ with finite second moment defined by

$$\Phi(\mu) := \mathcal{L}\left(\sqrt{U}\, X_1 + \frac{1}{2}\sqrt{1-U}\, X_2\right),$$

where X_1, X_2 are independent with law μ, independently of U. If $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$ both have mean 1, take any coupling (X,Y) of μ, ν, and two independent copies $(X_1,Y_1), (X_2,Y_2)$, independent of U. Since $\mathbb{E}[X-Y]=0$,

$$\begin{aligned}\mathbb{E}\left[\left(\sqrt{U}(X_1-Y_1) + \frac{1}{2}\sqrt{1-U}(X_2-Y_2)\right)^2\right] \\ = \left(\mathbb{E}[U] + \frac{1}{4}\mathbb{E}[1-U]\right)\mathbb{E}[(X-Y)^2] = \frac{5}{8}\mathbb{E}[(X-Y)^2].\end{aligned}$$

Taking the infimum over all couplings gives

$$W_2^2(\Phi(\mu), \Phi(\nu)) \leqslant \frac{5}{8} W_2^2(\mu,\nu)$$

where W_2 denotes the L^2-Wasserstein distance on $\mathcal{P}_2(\mathbb{R})$. Thus Φ has at most one fixed point in $\mathcal{P}_2(\mathbb{R})$ with mean 1. Since Proposition 4 gives that $M_\infty^* \in L^2$ and $\mathbb{E}[M_\infty^*]=1$, the first sentence of Proposition 5 is proved.

Finally, $M_\infty^* \geqslant 0$ almost surely. Setting $p := \mathbb{P}(M_\infty^* = 0)$, the almost sure strict positivity of both coefficients $\sqrt{U}$ and $\sqrt{1-U}$ in (19) gives

$$p = p^2.$$

Since M_∞^* is not identically zero (since $\mathbb{E}[M_\infty^*]=1$), $p=0$, and hence $M_\infty^* > 0$ almost surely. □

Remark 6. Taking second moments in (19) together with the condition that $\mathbb{E}[W]=1$ gives

$$\mathbb{E}[W^2] = \left(\mathbb{E}[U] + \frac{1}{4}\mathbb{E}[1-U]\right)\mathbb{E}[W^2] + \mathbb{E}\left[\sqrt{U(1-U)}\right].$$

Since

$$\mathbb{E}[U] = \frac{1}{2}, \qquad \mathbb{E}\left[\sqrt{U(1-U)}\right] = \frac{\pi}{8},$$

we obtain

$$\mathbb{E}[W^2] = \frac{\pi}{3}, \qquad \mathrm{Var}(W) = \frac{\pi}{3} - 1.$$

More generally, by homogeneity, for every $m \geqslant 0$, the unique L^2 solution of (19) with mean m is mW. Equation (19) belongs to the class of smoothing transforms, or random weighted mean equations, studied by Liu [16]. General results from that work yield further information on the law of W, which we do not pursue here.

5.2 Paternal negligibility

We now compare the initial advantaged genealogy with the RRT. For $1 \leqslant k \leqslant N-1$, set

$$T_k := S_{k+1} - 1.$$

Since B is constant during the stable part of the block $[S_k, S_{k+1})$, its jump is

$$B_{k+1} - B_k = \frac{1}{2}\left(A_{T_k}(\mu_{T_k}, 1) + A_{T_k}(\pi_{T_k}, 1)\right),$$

where μ_{T_k} is the advantaged mother and π_{T_k} is the uniformly chosen father.

Define the maternal-only weights $A^{(\mathrm{m})}$ on the same pedigree by

$$A_0^{(\mathrm{m})}(i,1) := \mathbf{1}_{\{i=1\}},$$

and

$$A_{n+1}^{(\mathrm{m})}(i,1) = \begin{cases} A_n^{(\mathrm{m})}(i,1), & i \neq \kappa_n, \\ \dfrac{1}{2} A_n^{(\mathrm{m})}(\mu_n, 1), & i = \kappa_n. \end{cases}$$

Let

$$B_k^{(\mathrm{m})} := \sum_{i \in \mathcal{Y}_{S_k}} A_{S_k}^{(\mathrm{m})}(i,1).$$

Thus $A^{(\mathrm{m})}$ is constructed on exactly the same population pedigree as A, but with all paternal genetic contributions discarded: at each reproduction event, the newborn retains only the half of the founder's contribution inherited from its mother. In particular, the type dynamics and the set of advantaged individuals are unchanged. Since advantaged individuals are never replaced and each new advantaged individual is born from an advantaged mother, the maternal genealogy of the advantaged class couples to a growing random recursive tree.

Lemma 4 (Maternal-only coupling)**.** *For every* $1 \leqslant k \leqslant N$*,*

$$0 \leqslant B_k^{(\mathrm{m})} \leqslant B_k.$$

Moreover,

$$\left(B_k^{(\mathrm{m})}\right)_{1 \leqslant k \leqslant N} \overset{d}{=} \left(B_k^{\star}\right)_{1 \leqslant k \leqslant N}.$$

Proof. The inequality follows by induction from

$$0 \leqslant A_n^{(\mathrm{m})}(i,1) \leqslant A_n(i,1).$$

Now label the advantaged individuals by their order of appearance and attach each new individual to its advantaged mother. Fix $1 \leqslant k \leqslant N-1$ and write $L_k := S_{k+1} - S_k - 1$. On $\{L_k = \ell\}$, conditionally on $\mathcal{F}_{S_k+\ell}$, the mother producing the jump $k \to k+1$ is uniform on the k advantaged individuals. Hence, conditionally on the genealogy up to size k, the parent of vertex $k+1$ is uniform on $\{1, \ldots, k\}$, exactly as in a random recursive tree.

Moreover, stable steps do not modify the maternal-only weights carried by advantaged individuals, while a child receives half the weight of its mother. Thus a vertex v carries weight $2^{-\operatorname{depth}(v)}$, and therefore

$$B_k^{(\mathrm{m})} = \sum_{v \in \mathcal{T}_k} 2^{-\operatorname{depth}(v)}.$$

Consequently,

$$\left(B_k^{(\mathrm{m})}\right)_{1 \leqslant k \leqslant N} \overset{d}{=} \left(B_k^{*}\right)_{1 \leqslant k \leqslant N}.$$

□

We shall use the following simple estimate on the disadvantaged mass immediately before a jump.

Lemma 5. *There exists a constant $K < \infty$, independent of N and k, such that, for all $1 \leqslant k \leqslant N-2$,*

$$\mathbb{E}[C_{T_k}] \leqslant K\sqrt{k+1}.$$

Proof. Recall that $L := S_{k+1} - S_k - 1$.

For every $\ell \geqslant 0$, on $\{L = \ell\}$ one has $T_k = S_k + \ell$. Conditionally on $\mathcal{F}_{S_k+\ell}$ and on $\{L = \ell\}$, the final mother is advantaged. From Proposition 2 (Equation (6)),

$$\mathbb{E}\left[C_{S_{k+1}} \,\middle|\, \mathcal{F}_{S_k+\ell},\, L = \ell\right] = \left(1 - \frac{1}{N-k}\right) C_{S_k+\ell}.$$

Multiplying by $\mathbf{1}_{\{L=\ell\}}$, taking expectations, and summing over $\ell \geqslant 0$, we obtain

$$\begin{aligned}\mathbb{E}[C_{S_{k+1}}] &= \left(1 - \frac{1}{N-k}\right) \sum_{\ell \geqslant 0} \mathbb{E}\left[C_{S_k+\ell} \mathbf{1}_{\{L=\ell\}}\right] \\ &= \left(1 - \frac{1}{N-k}\right) \mathbb{E}[C_{T_k}].\end{aligned}$$

Hence, since $k \leqslant N-2$,

$$\mathbb{E}[C_{T_k}] = \frac{N-k}{N-k-1} \mathbb{E}[C_{S_{k+1}}] \leqslant 2\,\mathbb{E}[C_{S_{k+1}}].$$

The conclusion follows from Lemma 2, after increasing K if necessary. □

The following lemma shows that, up to any mesoscopic level $k_0 \ll N$, paternal genetic input is asymptotically negligible. Thus the genetic contribution carried by the advantaged class is, to first order, governed by the maternal-only dynamics.

Lemma 6 (Negligibility of paternal input)**.** *Let $B_k^{(\mathrm{m})}$ denote the maternal-only contribution. There exists a constant $K < \infty$ such that, for every sequence $k_0 = k_0(N)$ satisfying $k_0 \to \infty$ and $k_0/N \to 0$,*

$$\mathbb{E}\left[B_{k_0} - B_{k_0}^{(\mathrm{m})}\right] \leqslant K \frac{k_0^{3/2}}{N}.$$

Consequently,

$$\frac{B_{k_0} - B_{k_0}^{(\mathrm{m})}}{\sqrt{k_0}} \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Proof. Set

$$E_k := B_k - B_k^{(\mathrm{m})}.$$

Stable steps do not modify the contribution carried by advantaged individuals. Hence, at the jump from k to $k+1$,

$$E_{k+1} - E_k = \frac{1}{2}\left(A_{T_k}(\mu_{T_k}, 1) - A_{T_k}^{(\mathrm{m})}(\mu_{T_k}, 1)\right) + \frac{1}{2} A_{T_k}(\pi_{T_k}, 1).$$

Conditionally on the pre-jump configuration, the mother is uniform on the k advantaged individuals, whereas the father is uniform on the whole population. Therefore

$$\mathbb{E}\left[E_{k+1} \,|\, \mathcal{F}_{S_k}\right] = \left(1 + \frac{1}{2k}\right) E_k + \frac{1}{2N} \mathbb{E}\left[B_k + C_{T_k} \,|\, \mathcal{F}_{S_k}\right].$$

Now from Lemma 2 and Lemma 5, there exists $K < \infty$ such that, uniformly in $1 \leqslant k < N-1$,

$$\mathbb{E}\left[B_k + C_{T_k}\right] \leqslant K\sqrt{k}.$$

Taking expectations, we now deduce that

$$\mathbb{E}[E_{k+1}] \leqslant \left(1 + \frac{1}{2k}\right)\mathbb{E}[E_k] + K\frac{\sqrt{k}}{N}. \tag{21}$$

Recall that $\gamma_k = \prod_{\ell=1}^{k-1}\left(1+\frac{1}{2\ell}\right)$ from (18). Dividing (21) by γ_{k+1} gives

$$\frac{\mathbb{E}[E_{k+1}]}{\gamma_{k+1}} \leqslant \frac{\mathbb{E}[E_k]}{\gamma_k} + K\frac{\sqrt{k}}{N\gamma_{k+1}}.$$

Since $E_1 = 0$, iteration yields

$$\frac{\mathbb{E}[E_{k_0}]}{\gamma_{k_0}} \leqslant \frac{K}{N}\sum_{k=1}^{k_0-1}\frac{\sqrt{k}}{\gamma_{k+1}}.$$

By Lemma 3, there exists a finite number K such that for any integer $k \geqslant 1$, $\sqrt{k}/\gamma_{k+1} \leqslant \sqrt{k+1}/\gamma_{k+1} \leqslant K$, therefore such that

$$\sum_{k=1}^{k_0-1}\frac{\sqrt{k}}{\gamma_{k+1}} \leqslant Kk_0,$$

and thus

$$\mathbb{E}[E_{k_0}] \leqslant K\frac{\gamma_{k_0}k_0}{N} \leqslant K\frac{k_0^{3/2}}{N},$$

by Lemma 3 again. This proves the announced expectation bound.

Finally, $E_{k_0} \geqslant 0$, so Markov's inequality gives, for every $\eta > 0$,

$$\mathbb{P}\left(\frac{E_{k_0}}{\sqrt{k_0}} > \eta\right) \leqslant \frac{\mathbb{E}[E_{k_0}]}{\eta\sqrt{k_0}} \leqslant \frac{K}{\eta}\frac{k_0}{N},$$

which tends to 0 as N goes to infinity, by assumption. Hence

$$\frac{B_{k_0} - B^{(\mathrm{m})}_{k_0}}{\sqrt{k_0}} \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

□

5.3 Seeding

Proposition 6 (Seeding at mesoscopic entrance)**.** *Let $k_0 = k_0(N)$ be such that $1 \ll k_0 \ll N$. Then*

$$\frac{B_{k_0}}{\sqrt{k_0}} \Longrightarrow \frac{2}{\sqrt{\pi}}\,M^\star_\infty, \qquad \text{as } N \to \infty. \tag{22}$$

Moreover,

$$M^{(N)}_{k_0} \Longrightarrow M^\star_\infty. \tag{23}$$

Proof. By Lemma 4,

$$\frac{B^{(\mathrm{m})}_{k_0}}{\sqrt{k_0}} \overset{d}{=} \frac{B^\star_{k_0}}{\sqrt{k_0}} = \frac{\gamma_{k_0}}{\sqrt{k_0}}\,M^\star_{k_0}.$$

Lemma 3 gives

$$\frac{\gamma_{k_0}}{\sqrt{k_0}} \longrightarrow \frac{2}{\sqrt{\pi}},$$

and from Proposition 4, $M_k^\star \to M_\infty^\star$ in L^2. Hence

$$\frac{B_{k_0}^{(\mathrm{m})}}{\sqrt{k_0}} \Longrightarrow \frac{2}{\sqrt{\pi}}\, M_\infty^\star.$$

Together with Lemma 6, this implies (22). It remains to identify the entrance value of the martingale. We write

$$M_{k_0}^{(N)} = a_{k_0,N}\left(\frac{B_{k_0}}{k_0} - \frac{C_{k_0}}{N-k_0}\right).$$

By Lemma 1, since $k_0 = k_0(N)$ diverges as N goes to infinity,

$$\frac{a_{k_0,N}}{\sqrt{k_0}} \longrightarrow \frac{\sqrt{\pi}}{2}.$$

Therefore,

$$a_{k_0,N}\frac{B_{k_0}}{k_0} = \frac{a_{k_0,N}}{\sqrt{k_0}}\,\frac{B_{k_0}}{\sqrt{k_0}} \Longrightarrow M_\infty^\star.$$

On the other hand, Lemma 2 gives

$$\mathbb{E}\left[a_{k_0,N}\frac{C_{k_0}}{N-k_0}\right] \leqslant K\frac{k_0}{N} \longrightarrow 0.$$

Since $a_{k_0,N}\frac{C_{k_0}}{N-k_0}$ is non-negative, it converges to 0 in probability, which gives (23). □

6 Macroscopic propagation

6.1 Quadratic variation

Proposition 7 (One-step conditional bound)**.** *There exists a constant $K < \infty$, independent of N and k, such that, for all $N \geqslant 3$ and $1 \leqslant k \leqslant N-2$, setting $d := N-k$, one has*

$$\mathbb{E}\left[\left(M_{k+1}^{(N)} - M_k^{(N)}\right)^2 \,\middle|\, \mathcal{F}_{S_k}\right] \leqslant K\left[\frac{\overline{B}_k + \overline{C}_k}{k} + \frac{\overline{B}_k + \overline{C}_k}{d} + \frac{k}{d^2}\left(\overline{B}_k + \overline{C}_k\right)\right].$$

We use the following Lemma to prove Proposition 7.

Lemma 7 (Disadvantaged average during a stable block)**.** *Fix $1 \leqslant k \leqslant N-1$, and set*

$$m_j := \mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_{S_k}, L \geqslant j\right], \qquad j \geqslant 0.$$

Then

$$m_{j+1} = (1-\lambda)m_j + \lambda\overline{B}_k, \qquad m_0 = \overline{C}_k. \tag{24}$$

Consequently,

$$m_j = \overline{B}_k + (1-\lambda)^j\left(\overline{C}_k - \overline{B}_k\right), \tag{25}$$

and in particular

$$m_j \leqslant \max\{\overline{B}_k, \overline{C}_k\} \leqslant \overline{B}_k + \overline{C}_k. \tag{26}$$

Proof of Lemma 7. On $\{L \geqslant j+1\}$, Proposition 2 (Equation (5)) gives

$$\mathbb{E}\left[\overline{C}^{(j+1)} \,\middle|\, \mathcal{F}_{S_k+j},\, L \geqslant j+1\right] = \overline{C}^{(j)} + \lambda X^{(j)} = (1-\lambda)\overline{C}^{(j)} + \lambda\overline{B}_k.$$

Moreover, on $\{L \geqslant j\}$, the conditional probability that the next update is stable is q, independently of the current genetic configuration. Hence

$$\mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_{S_k}, L \geqslant j+1\right] = \mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_{S_k}, L \geqslant j\right] = m_j.$$

Taking conditional expectations in the first display yields (24). The explicit formula (25) follows by solving this affine recurrence. The first inequality of (26) can be obtained by a disjunction case based on the sign of $\overline{C}_k - \overline{B}_k$ and the second inequality comes from the positivity of the mean contributions $\overline{C}_k$ and $\overline{B}_k$. □

Proof of Proposition 7. Throughout the proof, K denotes a finite constant whose value may change from line to line.

Since by definition (7), $a_{k,N} = a_{k+1,N}\alpha_{k,N}$ for every $1 \leqslant k \leqslant N-1$, then from the definition (8) of $M_k^{(N)}$,

$$M_{k+1}^{(N)} - M_k^{(N)} = a_{k+1,N}\big(X_{k+1} - \alpha_{k,N}X_k\big). \tag{27}$$

By Lemma 1, there exists $K \geqslant 0$ such that for any $1 \leqslant k \leqslant N-2$,

$$a_{k+1,N}^2 \leqslant Kk. \tag{28}$$

We now study the difference $X_{k+1} - \alpha_{k,N}X_k$, and to this aim analyse the block $[S_k, S_{k+1})$. Recall the block notations (9) and set $T := S_k$ for a better readability. Conditionally on $\mathcal{F}_T$,

$$\mathbb{P}(L = \ell \mid \mathcal{F}_T) = q^\ell \frac{k}{N}, \qquad \ell \geqslant 0,$$

and therefore

$$\mathbb{E}[L \mid \mathcal{F}_T] = \frac{d}{k}, \qquad \mathrm{Var}(L \mid \mathcal{F}_T) = \frac{Nd}{k^2}. \tag{29}$$

Set also

$$\lambda := \frac{k}{2Nd}, \qquad r := 1 - \lambda, \qquad s := 1 - \frac{d}{2N(k+1)}.$$

For $j \geqslant 0$, let

$$\mathcal{B}_j := \mathcal{F}_T \vee \sigma\big((\mu_{T+i}, \pi_{T+i}, \kappa_{T+i}) : 0 \leqslant i < j\big).$$

On $\{L \geqslant j\}$, denote by $X^{(j)}$ and $\overline{C}^{(j)}$ the values of X and $\overline{C}$ after the first j stable steps of the block i.e. at time $T + j$.

On $\{L \geqslant j+1\}$, Proposition 2 gives

$$X^{(j+1)} = rX^{(j)} + \xi_{j+1}, \qquad \text{where} \qquad \mathbb{E}\left[\xi_{j+1} \,|\, \mathcal{B}_j, L \geqslant j+1\right] = 0.$$

At the final jump, on $\{L = \ell\}$, Proposition 2 again gives

$$X_{k+1} = sX^{(\ell)} + \zeta, \qquad \text{where} \qquad \mathbb{E}\left[\zeta \,|\, \mathcal{B}_\ell, L = \ell\right] = 0.$$

It follows that

$$\begin{aligned} X_{k+1} - \alpha_{k,N}X_k &= (sr^L - \alpha_{k,N})X_k + s\sum_{j=1}^{L} r^{L-j}\xi_j + \zeta \\ &= P + Q + J, \end{aligned} \tag{30}$$

where

$$P := (sr^L - \alpha_{k,N})X_k, \qquad Q := s\sum_{j=1}^{L} r^{L-j}\xi_j, \qquad J := \zeta.$$

Since

$$(P + Q + J)^2 \leqslant 3P^2 + 3Q^2 + 3J^2, \tag{31}$$

the three terms can be dealt with separately.

We first control P. Note that,

$$\mathbb{E}[sr^L \mid \mathcal{F}_T] = \frac{s(k/N)}{1 - qr} = \frac{2Nk + N + k}{(2N+1)(k+1)} = \alpha_{k,N}.$$

Using (29), since $0 < s \leqslant 1$ and $\ell \mapsto r^\ell$ is λ-Lipschitz,

$$\operatorname{Var}(sr^L \mid \mathcal{F}_T) \leqslant \lambda^2 \operatorname{Var}(L \mid \mathcal{F}_T) \leqslant \frac{K}{Nd}. \tag{32}$$

Furthermore, since $0 \leqslant \overline{B}_k, \overline{C}_k \leqslant 1$,

$$X_k^2 = (\overline{B}_k - \overline{C}_k)^2 \leqslant \overline{B}_k + \overline{C}_k. \tag{33}$$

Hence, by (28), (32), and (33),

$$a_{k+1,N}^2 \mathbb{E}[P^2 \mid \mathcal{F}_T] \leqslant K \frac{k}{Nd} \big(\overline{B}_k + \overline{C}_k\big) \leqslant K \frac{\overline{B}_k + \overline{C}_k}{d} \tag{34}$$

since $k \leqslant N$. We next control the stable steps martingale term Q. During a stable step, recall from the fundamental dynamics described in Proposition 2 that

$$X^{(j+1)} - X^{(j)} = -\frac{1}{d}\left(\frac{w(\mu) + w(\pi)}{2} - w(\kappa)\right).$$

Since by definition

$$\xi_{j+1} = X^{(j+1)} - X^{(j)} + \lambda X^{(j)} \tag{35}$$

and $\mathbb{E}\left[\xi_{j+1} \,|\, \mathcal{B}_j, L \geqslant j+1\right] = 0$,

$$\begin{aligned}
\mathbb{E}\left[\xi_{j+1}^2 \,\middle|\, \mathcal{B}_j, L \geqslant j+1\right] &= \operatorname{Var}\left(\xi_{j+1} \,|\, \mathcal{B}_j, L \geqslant j+1\right) \\
&= \operatorname{Var}\left(X^{(j+1)} - X^{(j)} \,\middle|\, \mathcal{B}_j, L \geqslant j+1\right) \\
&\leqslant \frac{1}{d^2}\mathbb{E}\left[\Big]\Big(\frac{w(\mu) + w(\pi)}{2} - w(\kappa)\Big)^2 \,\middle|\, \mathcal{B}_j, L \geqslant j+1\right] \\
&\leqslant \frac{K}{d^2}\mathbb{E}\left[\Big]\frac{w(\mu) + w(\pi)}{2} - w(\kappa)\,\middle|\, \mathcal{B}_j, L \geqslant j+1\right] \qquad \text{since } w(i)^2 \leqslant w(i) \\
&\leqslant \frac{K}{d^2}\left(\overline{C}^{(j)} + \overline{B}_k\right).
\end{aligned} \tag{36}$$

Indeed, conditionally on a stable step, the mother and the killed individual are uniform on the disadvantaged class, whereas the father is uniform on I_N, and $w(i)^2 \leqslant w(i)$.

Conditionally on $L = \ell$, the increments $\xi_1, \ldots, \xi_\ell$ are martingale differences, so the cross terms vanish. Since $0 \leqslant r, s \leqslant 1$, (36) gives

$$\mathbb{E}[Q^2 \mid \mathcal{F}_T] \leqslant \frac{K}{d^2}\mathbb{E}\left[\sum_{j=0}^{L-1}\left(\overline{C}^{(j)} + \overline{B}_k\right) \,\middle|\, \mathcal{F}_T\right].$$

Let

$$m_j := \mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_T, L \geqslant j\right].$$

By Lemma 7,

$$m_j \leqslant \overline{B}_k + \overline{C}_k.$$

Since $\mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_T, L \geqslant j+1\right] = \mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_T, L \geqslant j\right]$,

$$\begin{aligned}
\mathbb{E}\left[\sum_{j=0}^{L-1}\left(\overline{C}^{(j)} + \overline{B}_k\right) \,\middle|\, \mathcal{F}_T\right] &= \sum_{j \geqslant 0} E\left[\mathbf{1}_{L \geqslant j+1}\left(\overline{C}^{(j)} + \overline{B}_k\right) \,\middle|\, \mathcal{F}_T\right] \\
&= \sum_{j \geqslant 0} q^{j+1} E\left[\left(\overline{C}^{(j)} + \overline{B}_k\right) \,\middle|\, \mathcal{F}_T, L \geqslant j+1\right] \\
&= \sum_{j \geqslant 0} q^{j+1}\left(m_j + \overline{B}_k\right) \\
&\leqslant K\frac{d}{k}\left(\overline{B}_k + \overline{C}_k\right).
\end{aligned}$$

Therefore using that $q = d/N$,

$$\mathbb{E}[Q^2 \mid \mathcal{F}_T] \leqslant \frac{K}{kd}\left(\overline{B}_k + \overline{C}_k\right)$$

and hence, by (28),

$$a_{k+1,N}^2 \mathbb{E}[Q^2 \mid \mathcal{F}_T] \leqslant \frac{K}{d}\left(\overline{B}_k + \overline{C}_k\right). \tag{37}$$

It remains to control the final jump term $J = \zeta = X_{k+1} - sX^{(\ell)}$. Let $\overline{C}_-$ denote the disadvantaged average immediately before that jump. Proposition 2 gives

$$X_{k+1} - X^{(L)} = -\frac{\overline{B}_k}{k+1} - \frac{\overline{C}_-}{d-1} + \frac{w(\mu)+w(\pi)}{2(k+1)} + \frac{w(\kappa)}{d-1}.$$

We have as previously

$$\begin{aligned}
\mathbb{E}\left[J^2 \,\middle|\, \mathcal{B}_\ell, L=\ell\right] &= Var(J \,|\, \mathcal{B}_\ell, L=\ell) \\
&= Var(X_{S_{k+1}} - X_{S_{k+1}-1} \,|\, \mathcal{B}_\ell, L=\ell) \\
&= Var\Big(\frac{w(\mu)+w(\pi)}{2(k+1)} + \frac{w(\kappa)}{d-1} \,|\, \mathcal{B}_\ell, L=\ell\Big) \\
&\leqslant K\left[\frac{\overline{B}_k + \overline{C}_-}{k^2} + \frac{\overline{C}_-}{d^2}\right] \qquad \text{using again } w(i)^2 \leqslant w(i) \text{ and } k \leqslant N-2.
\end{aligned}$$

Since $\overline{C}_- = \overline{C}^{(L)}$, conditioning on L and using Lemma 7 gives

$$\mathbb{E}[\overline{C}_- \mid \mathcal{F}_T] = \sum_{\ell \geqslant 0} q^\ell \frac{k}{N} m_\ell \leqslant \overline{B}_k + \overline{C}_k.$$

Consequently,

$$\mathbb{E}[J^2 \mid \mathcal{F}_T] \leqslant K\left[\frac{\overline{B}_k + \overline{C}_k}{k^2} + \frac{\overline{B}_k + \overline{C}_k}{d^2}\right],$$

and therefore

$$a_{k+1,N}^2 \mathbb{E}[J^2 \mid \mathcal{F}_T] \leqslant K\left[\frac{\overline{B}_k + \overline{C}_k}{k} + \frac{k}{d^2}\left(\overline{B}_k + \overline{C}_k\right)\right]. \tag{38}$$

Recalling the Equations (27), (30), (31) and using (34), (37), and (38), proves the result. □

Lemma 8 (Negligibility of the martingale increments away from the origin)**.** *Let $k_0 = k_0(N)$ be an integer sequence such that $1 \ll k_0 \ll N$. Then, for every $\varepsilon \in (0,1)$,*

$$\sup_{k_0 \leqslant k \leqslant \lfloor (1-\varepsilon)N \rfloor} \left| M_k^{(N)} - M_{k_0}^{(N)} \right| \xrightarrow[N\to\infty]{L^2} 0.$$

In particular, the convergence holds in probability.

Proof. Write

$$\Delta M_k := M_{k+1}^{(N)} - M_k^{(N)}.$$

Set

$$m_N := \lfloor (1-\varepsilon)N \rfloor.$$

For $k \leqslant m_N$, we have

$$d = N - k \geqslant N - (1-\varepsilon)N = \varepsilon N \geqslant \varepsilon k.$$

Consequently,

$$\frac{1}{d} \leqslant \frac{1}{\varepsilon k}, \qquad \frac{k}{d^2} \leqslant \frac{1}{\varepsilon^2 k}.$$

Taking expectations in Proposition 7 and then using Lemma 2, we obtain, for $k \leqslant m_N - 1$,

$$\begin{aligned}\mathbb{E}\left[(\Delta M_k)^2\right] &\leqslant K_\varepsilon\, \mathbb{E}\left[\frac{\overline{B}_k + \overline{C}_k}{k}\right] \\ &\leqslant \frac{6K_\varepsilon}{k^{3/2}},\end{aligned}$$

where $K_\varepsilon < \infty$ is independent of N and k. Therefore,

$$\begin{aligned}\sum_{k=k_0}^{m_N-1} \mathbb{E}\left[(\Delta M_k)^2\right] &\leqslant K_\varepsilon \sum_{k=k_0}^{\infty} \frac{1}{k^{3/2}} \\ &\leqslant \frac{K_\varepsilon}{\sqrt{k_0}}.\end{aligned}$$

The right-hand side tends to 0 as N goes to infinity.

Since $(M_k^{(N)})_k$ is a martingale with respect to $(\mathcal{F}_{S_k})_k$, Doob's L^2 inequality gives

$$\mathbb{E}\left[\sup_{k_0 \leqslant k \leqslant m_N} \left|M_k^{(N)} - M_{k_0}^{(N)}\right|^2\right] \leqslant 4\,\mathbb{E}\left[\left|M_{m_N}^{(N)} - M_{k_0}^{(N)}\right|^2\right] = 4 \sum_{k=k_0}^{m_N-1} \mathbb{E}\left[(\Delta M_k)^2\right],$$

which tends to 0. This proves the claim. □

6.2 B and C reconstruction

The exact decomposition (11) identifies the drift of $\overline{C}$ over the whole block $[S_k, S_{k+1}]$. We now estimate the conditional variance of its centred remainder.

Proposition 8 (Quadratic bound for the $\overline{C}$-remainder)**.** *There exists a constant $K < \infty$ such that, for every $N \geqslant 3$ and $1 \leqslant k \leqslant N-2$, setting $d := N - k$, one has*

$$\mathbb{E}\left[(\Delta R_{k+1})^2 \,\middle|\, \mathcal{F}_{S_k}\right] \leqslant K\left(\frac{1}{kd} + \frac{1}{d^2}\right)\left(\overline{B}_k + \overline{C}_k\right), \tag{39}$$

where ΔR_{k+1} is the centred increment defined in (11)*.*

Consequently, for every $\varepsilon \in (0,1)$, for $N > 2/\varepsilon$, there exists $K_\varepsilon < \infty$ such that, whenever

$$1 \leqslant k \leqslant \lfloor (1-\varepsilon)N \rfloor,$$

one has

$$\mathbb{E}\left[(\Delta R_{k+1})^2 \,\middle|\, \mathcal{F}_{S_k}\right] \leqslant \frac{K_\varepsilon}{Nk}\left(\overline{B}_k + \overline{C}_k\right). \tag{40}$$

Proof. The proof uses the same type of arguments and the same outline than that of Proposition 7. We use the block notation (9) as well as, as in the proof of Proposition 7, the notations $T := S_k$ and $\mathcal{B}_j := \mathcal{F}_T \vee \sigma\big((\mu_{T+i}, \pi_{T+i}, \kappa_{T+i}) : 0 \leqslant i < j\big)$ for all $j \geqslant 0$. Similarly, on $\{L \geqslant j\}$, $X^{(j)}$ and $\overline{C}^{(j)}$ denote the values of X and $\overline{C}$ after the first j stable steps of the block, i.e. at time $S_k + j = T + j$.

Recall that

$$\Delta R_{k+1} = \overline{C}_{k+1} - \overline{C}_k - \frac{1}{2N+1} X_k. \tag{41}$$

Conditionally on $\mathcal{F}_T$, recall from Proposition 1 that L follows a geometric law with parameter $1-q$, with values in $\mathbb{Z}_+$. In particular,

$$\mathbb{E}[L^2 \mid \mathcal{F}_T] \leqslant \frac{KNd}{k^2}. \tag{42}$$

During a stable step, write as in Equation (35)

$$\overline{C}^{(j+1)} - \overline{C}^{(j)} = -(X^{(j+1)} - X^{(j)}) = \lambda X^{(j)} - \xi_{j+1}. \tag{43}$$

Then, on $\{L \geqslant j+1\}$, $\mathbb{E}\left[\xi_{j+1} \,|\, \mathcal{B}_j, L \geqslant j+1\right] = 0$, and from Equation (36)

$$\mathbb{E}\left[\xi_{j+1}^2 \,\middle|\, \mathcal{B}_j, L \geqslant j+1\right] \leqslant \frac{K}{d^2}\left(\overline{C}^{(j)} + \overline{B}_k\right). \tag{44}$$

At the final jump, on $\{L = \ell\}$, set

$$J := \overline{C}_{k+1} - \overline{C}^{(\ell)} = \frac{\overline{C}^{(\ell)} - w(\kappa)}{d-1} \qquad \text{from Proposition 2.}$$

Set

$$P := \sum_{j=0}^{L-1} \lambda X^{(j)}, \qquad Q := \sum_{j=0}^{L-1} \xi_{j+1}.$$

Then by (43),

$$\overline{C}_{k+1} - \overline{C}_k = P - Q + J.$$

From Proposition 2 (Equations (5) and (6) respectively), the terms Q and J are conditionally centred. Hence, by Proposition 3,

$$\mathbb{E}[P \mid \mathcal{F}_T] = \frac{1}{2N+1} X_k,$$

and therefore by (41),

$$\Delta R_{k+1} = \big(P - \mathbb{E}[P \mid \mathcal{F}_T]\big) - Q + J. \tag{45}$$

We again use the fact that $(P+Q+J)^2 \leqslant 3P^2 + 3Q^2 + 3J^2$ and control each term separately. We first control Q. Let

$$m_j := \mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_T, L \geqslant j\right].$$

By Lemma 7,

$$m_j \leqslant \overline{B}_k + \overline{C}_k, \qquad j \geqslant 0. \tag{46}$$

Writing

$$Q = \sum_{j \geqslant 0} \mathbf{1}_{\{L \geqslant j+1\}} \xi_{j+1},$$

and using the martingale-difference property, (44), and (46), we obtain

$$\begin{aligned}
\mathbb{E}[Q^2 \mid \mathcal{F}_T] &\leqslant \frac{K}{d^2} \sum_{j \geqslant 0} q^{j+1} \left(m_j + \overline{B}_k\right) \\
&\leqslant \frac{K}{d^2} \left(\overline{B}_k + \overline{C}_k\right) \sum_{j \geqslant 0} q^{j+1} \\
&= \frac{K}{kd} \left(\overline{B}_k + \overline{C}_k\right).
\end{aligned} \tag{47}$$

We now focus on the final jump term J. Conditionally on $\mathcal{B}_\ell$ and $\{L = \ell\}$,

$$\mathbb{E}[J \mid \mathcal{B}_\ell, L = \ell] = 0,$$

and, since $d \geqslant 2$,

$$\mathbb{E}[J^2 \mid \mathcal{B}_\ell, L = \ell] \leqslant \frac{K}{d^2} \overline{C}^{(\ell)}. \tag{48}$$

Therefore conditioning on L and using (46) gives

$$\mathbb{E}\left[\overline{C}^{(L)} \,\middle|\, \mathcal{F}_T\right] = \sum_{\ell \geqslant 0} (1-q) q^{\ell} m_{\ell} \leqslant \overline{B}_k + \overline{C}_k. \tag{49}$$

Thus, (48) and (49) give

$$\mathbb{E}[J^2 \mid \mathcal{F}_T] \leqslant \frac{K}{d^2} \left(\overline{B}_k + \overline{C}_k\right). \tag{50}$$

It remains to control the fluctuation of P. Since

$$\mathbb{E}\left[(P - \mathbb{E}[P \mid \mathcal{F}_T])^2 \,\middle|\, \mathcal{F}_T\right] \leqslant \mathbb{E}[P^2 \mid \mathcal{F}_T] \tag{51}$$

and

$$(X^{(j)})^2 \leqslant \overline{B}_k + \overline{C}^{(j)},$$

as in Equation (33). Cauchy's inequality gives

$$P^2 \leqslant \lambda^2 L \sum_{j=0}^{L-1} \left(\overline{B}_k + \overline{C}^{(j)}\right). \tag{52}$$

Conditioning on L, using that this conditioning does not bias the first stable steps, and applying (46), (42), and (52), we get

$$\begin{aligned} \mathbb{E}[P^2 \mid \mathcal{F}_T] &\leqslant K\lambda^2 \left(\overline{B}_k + \overline{C}_k\right) \mathbb{E}[L^2 \mid \mathcal{F}_T] \\ &\leqslant \frac{K}{Nd} \left(\overline{B}_k + \overline{C}_k\right). \end{aligned} \tag{53}$$

Since $k \leqslant N$, (51) and (53) imply

$$\mathbb{E}\left[(P - \mathbb{E}[P \mid \mathcal{F}_T])^2 \,\middle|\, \mathcal{F}_T\right] \leqslant \frac{K}{kd} \left(\overline{B}_k + \overline{C}_k\right). \tag{54}$$

Combining (45), (47), (50), and (54) proves (39).
Now fix $\varepsilon \in (0,1)$ and assume $k \leqslant \lfloor (1-\varepsilon)N \rfloor$. Then

$$d = N - k \geqslant \varepsilon N.$$

Consequently,

$$\frac{1}{kd} \leqslant \frac{1}{\varepsilon N k}, \qquad \frac{1}{d^2} \leqslant \frac{1}{\varepsilon^2 N^2} \leqslant \frac{1}{\varepsilon^2 N k},$$

because $k \leqslant N$. Applying these bounds in (39) proves (40). □

Proposition 9 (Propagation through the deterministic phase)**.** *Let $k_0 = k_0(N)$ be an integer sequence such that $1 \ll k_0 \ll N$. Then, for every $\varepsilon \in (0,1)$,*

$$\max_{k_0 \leqslant k \leqslant \lfloor (1-\varepsilon)N \rfloor} \left| \frac{C_k}{\sqrt{N}} - \frac{2}{\sqrt{\pi}} \sqrt{\frac{k}{N}} \left(1 - \frac{k}{N}\right) M_{k_0}^{(N)} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0, \tag{55}$$

and

$$\max_{k_0 \leqslant k \leqslant \lfloor (1-\varepsilon)N \rfloor} \left| \frac{B_k}{\sqrt{N}} - \frac{2}{\sqrt{\pi}} \sqrt{\frac{k}{N}} \left(1 + \frac{k}{N}\right) M_{k_0}^{(N)} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0. \tag{56}$$

Proof. Fix $\varepsilon \in (0,1)$, and write

$$m_N := \lfloor (1-\varepsilon)N \rfloor, \qquad y_k := \frac{k}{N}.$$

Iterating (11), for $k \geqslant k_0$,

$$\overline{C}_k = \overline{C}_{k_0} + \frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} M_j^{(N)} + R_{k_0,k}. \qquad R_{k_0,k} := \sum_{j=k_0}^{k-1} \Delta R_{j+1}. \tag{57}$$

The process $(R_{k_0,k})_{k \geqslant k_0}$ is a martingale with respect to $(\mathcal{F}_{S_k})_{k \geqslant k_0}$, starting from 0. By Doob's L^2 inequality,

$$\begin{aligned}
\mathbb{E}\left[\sup_{k_0 \leqslant k \leqslant m_N} |R_{k_0,k}|^2 \right] &\leqslant 4 \sum_{j=k_0}^{m_N-1} \mathbb{E}[(\Delta R_{j+1})^2] \\
&\leqslant \frac{K_\varepsilon}{N} \sum_{j=k_0}^{m_N-1} \frac{\mathbb{E}[\overline{B}_j + \overline{C}_j]}{j} \qquad \text{from Proposition 8 and since } m_N = \lfloor (1-\epsilon)N \rfloor \\
&\leqslant \frac{K_\varepsilon}{N} \sum_{j=k_0}^{\infty} \frac{1}{j^{3/2}} \qquad \text{from Lemma 2} \\
&\leqslant \frac{K_\varepsilon}{N\sqrt{k_0}}.
\end{aligned}$$

Since $(N-k)/\sqrt{N} \leqslant \sqrt{N}$ and $k_0 \gg 1$, it follows that

$$\sup_{k_0 \leqslant k \leqslant m_N} \frac{N-k}{\sqrt{N}} |R_{k_0,k}| \xrightarrow[N\to\infty]{\mathbb{P}} 0. \tag{58}$$

We now estimate the adapted part of the process $(\overline{C}_k)_{k_0 \leqslant k \leqslant m_N}$. Write

$$\frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} M_j^{(N)} = M_{k_0}^{(N)} \frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} + \frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} \big(M_j^{(N)} - M_{k_0}^{(N)} \big). \tag{59}$$

By Lemma 8,

$$\sup_{k_0 \leqslant j \leqslant m_N} |M_j^{(N)} - M_{k_0}^{(N)}| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Moreover, by Lemma 1 (Equation (16)),

$$\frac{1}{2N+1} \sum_{j=k_0}^{m_N} a_{j,N}^{-1} = O(N^{-1/2}).$$

Hence,

$$\sup_{k_0 \leqslant k \leqslant m_N} \frac{1}{2N+1} \left| \sum_{j=k_0}^{k-1} a_{j,N}^{-1} \big(M_j^{(N)} - M_{k_0}^{(N)} \big) \right| = o_{\mathbb{P}}(N^{-1/2}). \tag{60}$$

Also, by Lemma 1 (Equation (16)) again,

$$\frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} = \frac{2}{\sqrt{\pi}} \frac{\sqrt{k}}{N} + o(N^{-1/2})$$

uniformly in k. Since the sequence $\big(M_{k_0}^{(N)}\big)_N$ is tight from Proposition 6, we obtain from (59) and (60) that

$$\frac{1}{2N+1} \sum_{j=k_0}^{k-1} a_{j,N}^{-1} M_j^{(N)} = \frac{2}{\sqrt{\pi}} \frac{\sqrt{k}}{N} M_{k_0}^{(N)} + o_{\mathbb{P}}(N^{-1/2}), \tag{61}$$

uniformly for $k_0 \leqslant k \leqslant m_N$.

It remains to handle the initial term $\overline{C}_{k_0}$. By Lemma 2, there exists a constant K such that

$$\mathbb{E}[\sqrt{N}\,\overline{C}_{k_0}] \leqslant K\sqrt{\frac{k_0}{N}} \xrightarrow[N\to\infty]{} 0. \tag{62}$$

Therefore from Equations (57), (62), (61) and (58) we get

$$\sup_{k_0\leqslant k\leqslant m_N} \frac{N-k}{\sqrt{N}} \left| \overline{C}_k - \frac{2}{\sqrt{\pi}} \frac{\sqrt{k}}{N} M_{k_0}^{(N)} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0. \tag{63}$$

Since $C_k = (N-k)\overline{C}_k$, this proves (55).

We now prove (56). Since

$$\overline{B}_k = \overline{C}_k + X_k, \qquad X_k = \frac{M_k^{(N)}}{a_{k,N}} \qquad \text{and} \qquad B_k = k\overline{B}_k,$$

we have

$$\frac{B_k}{\sqrt{N}} = \frac{k\overline{C}_k}{\sqrt{N}} + \frac{k}{\sqrt{N}} \frac{M_k^{(N)}}{a_{k,N}}. \tag{64}$$

For the first term, using $k/(N-k) \leqslant \varepsilon^{-1}$ as long as $k \leqslant m_N \leqslant (1-\epsilon)N$, we get

$$\sup_{k_0\leqslant k\leqslant m_N} \left| \frac{k\overline{C}_k}{\sqrt{N}} - \frac{2}{\sqrt{\pi}} y_k^{3/2} M_{k_0}^{(N)} \right| \leqslant \frac{1}{\varepsilon} \sup_{k_0\leqslant k\leqslant m_N} \frac{N-k}{\sqrt{N}} \left| \overline{C}_k - \frac{2}{\sqrt{\pi}} \frac{\sqrt{k}}{N} M_{k_0}^{(N)} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0 \tag{65}$$

from Equation (63).

For the second term, we decompose

$$\begin{aligned} &\sup_{k_0\leqslant k\leqslant m_N} \left| \frac{k}{\sqrt{N}} \frac{M_k^{(N)}}{a_{k,N}} - \frac{2}{\sqrt{\pi}} \sqrt{y_k}\, M_{k_0}^{(N)} \right| \\ &\qquad \leqslant \sup_{k_0\leqslant k\leqslant m_N} \frac{k}{\sqrt{N} a_{k,N}} \left| M_k^{(N)} - M_{k_0}^{(N)} \right| + \left| M_{k_0}^{(N)} \right| \sup_{k_0\leqslant k\leqslant m_N} \left| \frac{k}{\sqrt{N}} \frac{1}{a_{k,N}} - \frac{2}{\sqrt{\pi}} \sqrt{y_k} \right| \end{aligned}$$

The second term converges to 0 in probability as N goes to infinity from Lemma 1 (Equation (15)) and since the sequence $(M_{k_0}^{(N)})_N$ is tight from Proposition 6. The first term converges to 0 in probability as N goes to infinity since $(\frac{k}{\sqrt{N} a_{k,N}})$ is uniformly bounded when $k_0 \leqslant k \leqslant m_N$ and from Lemma 8. Thus

$$\sup_{k_0\leqslant k\leqslant m_N} \left| \frac{k}{\sqrt{N}} \frac{M_k^{(N)}}{a_{k,N}} - \frac{2}{\sqrt{\pi}} \sqrt{y_k} M_{k_0}^{(N)} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0. \tag{66}$$

Combining the decomposition (64) with the two estimates (65) and (66) gives (56). □

7 Gluing

7.1 Initial window

We first prove the following Lemma.

Lemma 9 (Positive variation during a stable block)**.** *Fix $1 \leqslant k \leqslant N-1$, and define*

$$P_k := \sum_{j=0}^{L-1} \big(C^{(j+1)} - C^{(j)}\big)^+.$$

Then

$$\mathbb{E}\left[P_k \,|\, \mathcal{F}_{S_k}\right] \leqslant \frac{N-k}{k}\,\overline{C}_k + \overline{B}_k.$$

Consequently, there exists $K < \infty$, independent of N and k, such that

$$\mathbb{E}[P_k] \leqslant \frac{K}{\sqrt{k}}.$$

Proof. Recall that $d := N-k$, $q := d/N$, and $\lambda = \frac{k}{2Nd}$. During the $(j+1)$-st stable step,

$$\big(C^{(j+1)} - C^{(j)}\big)^+ \leqslant \frac{1}{2}\big(w(\mu) + w(\pi)\big).$$

Hence, on $\{L \geqslant j+1\}$,

$$\begin{aligned}
&\mathbb{E}\left[\big(C^{(j+1)} - C^{(j)}\big)^+ \,\middle|\, \mathcal{F}_{S_k+j}, L \geqslant j+1\right] \\
&\quad\leqslant \frac{1}{2}\left(\overline{C}^{(j)} + \frac{k}{N}\overline{B}_k + \frac{d}{N}\overline{C}^{(j)}\right) \\
&\quad\leqslant \overline{C}^{(j)} + \frac{k}{2N}\overline{B}_k.
\end{aligned}$$

Let

$$m_j := \mathbb{E}\left[\overline{C}^{(j)} \,\middle|\, \mathcal{F}_{S_k},\, L \geqslant j\right].$$

By Lemma 7,

$$m_j = \overline{B}_k + (1-\lambda)^j\big(\overline{C}_k - \overline{B}_k\big).$$

Since

$$\mathbb{P}(L \geqslant j+1 \mid \mathcal{F}_{S_k}) = q^{j+1},$$

we obtain

$$\mathbb{E}\left[P_k \,|\, \mathcal{F}_{S_k}\right] \leqslant \sum_{j\geqslant 0} q^{j+1} m_j + \frac{k}{2N}\frac{q}{1-q}\overline{B}_k.$$

Moreover,

$$\sum_{j\geqslant 0} q^{j+1} m_j = \frac{q}{1-q(1-\lambda)}\overline{C}_k + \left(\frac{q}{1-q} - \frac{q}{1-q(1-\lambda)}\right)\overline{B}_k.$$

Using that

$$\frac{q}{1-q(1-\lambda)} = \frac{2N(N-k)}{k(2N+1)},$$

$$\frac{q}{1-q} - \frac{q}{1-q(1-\lambda)} = \frac{N-k}{k(2N+1)},$$

and

$$\frac{k}{2N}\frac{q}{1-q} = \frac{N-k}{2N},$$

we conclude that

$$\mathbb{E}\left[P_k \,|\, \mathcal{F}_{S_k}\right] \leqslant \frac{N-k}{k}\,\overline{C}_k + \overline{B}_k.$$

Finally, Lemma 2 gives

$$\mathbb{E}\left[\frac{N-k}{k}\overline{C}_k\right] = \frac{\mathbb{E}[C_k]}{k} \leqslant \frac{K}{\sqrt{k}},$$

and

$$\mathbb{E}[\overline{B}_k] \leqslant \mathbb{E}[\overline{B}_k + \overline{C}_k] \leqslant \frac{K}{\sqrt{k}}.$$

This proves the second assertion. □

Lemma 9 enables the control of variations of the stochastic processes B and C during the initial window. This small estimate near the origin will be used to glue the mesoscopic entrance regime with the macroscopic one.

Lemma 10 (Initial window)**.** *Let $k_0 = k_0(N)$ be an integer sequence such that*

$$\frac{k_0}{N} \longrightarrow 0.$$

Then

$$\sup_{0 \leqslant y \leqslant k_0/N} \left| \frac{1}{\sqrt{N}} \begin{pmatrix} B_{k_N(y)} \\ C_{k_N(y)} \end{pmatrix} \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Proof. Since $(B_k)_{1 \leqslant k \leqslant N}$ is nondecreasing,

$$\sup_{1 \leqslant k \leqslant k_0} B_k = B_{k_0}.$$

By Lemma 2,

$$\mathbb{E}[B_{k_0}] \leqslant K\sqrt{k_0}.$$

For C, let P_k be the positive variation accumulated during the stable part of the block $[S_k, S_{k+1})$, as in Lemma 9. Since the final jump of each block can only decrease C and $C_1 = 0$,

$$\sup_{1 \leqslant k \leqslant k_0} C_k \leqslant \sum_{k=1}^{k_0 - 1} P_k.$$

Hence, by Lemma 9,

$$\mathbb{E}\left[\sup_{1 \leqslant k \leqslant k_0} C_k \right] \leqslant K \sum_{k=1}^{k_0 - 1} \frac{1}{\sqrt{k}} \leqslant K\sqrt{k_0}.$$

Therefore

$$\frac{1}{\sqrt{N}} \mathbb{E}\left[\sup_{1 \leqslant k \leqslant k_0} (B_k + C_k) \right] \leqslant K\sqrt{\frac{k_0}{N}} \longrightarrow 0.$$

The conclusion follows from Markov's inequality. □

7.2 Terminal window and continuity at $y = 1$

We close the proof at the endpoint without using any martingale estimate. We retain the notation introduced above and set $C_N := 0$. Lemma 9 enables the control of variations of the stochastic processes B and C during the terminal window.

Lemma 11 (Small variation in the terminal window)**.** *For $\varepsilon \in (0, 1/2)$, set*

$$k_\varepsilon := \lfloor (1 - \varepsilon)N \rfloor.$$

There exists $K < \infty$, independent of N and ε, such that, for every fixed $\varepsilon \in (0, 1/2)$ and all sufficiently large N,

$$\mathbb{E}[B_N - B_{k_\varepsilon}] \leqslant K\varepsilon\sqrt{N} \tag{67}$$

and

$$\mathbb{E}\left[\sup_{k_\varepsilon \leqslant k \leqslant N} C_k \right] \leqslant K\varepsilon\sqrt{N}. \tag{68}$$

Consequently, for every $\eta > 0$, and after changing the constant K,

$$\limsup_{N\to\infty} \mathbb{P}\left(\sup_{k_\varepsilon \leqslant k \leqslant N} \frac{|B_k - B_{k_\varepsilon}| + C_k}{\sqrt{N}} > \eta \right) \leqslant \frac{K\varepsilon}{\eta}. \tag{69}$$

Proof. Fix $\varepsilon \in (0, 1/2)$. For all sufficiently large N,

$$N - k_\varepsilon \leqslant \varepsilon N + 1 \leqslant 2\varepsilon N, \qquad k_\varepsilon \geqslant \frac{N}{3}. \tag{70}$$

We first control the process $(B_k)_{k_\epsilon \leqslant k \leqslant N}$. As mentioned at the beginning of the proof of Lemma 10, the process $(B_k)_{1 \leqslant k \leqslant N}$ is nondecreasing: stable steps do not modify it, while at the jump $k \to k+1$,

$$B_{k+1} - B_k = \frac{1}{2}\big(w(\mu_{T_k}) + w(\pi_{T_k})\big),$$

where $T_k := S_{k+1} - 1$. Conditionally on the pre-jump configuration, the mother is uniform on the advantaged class, whereas the father is uniform on the whole population. Hence

$$\mathbb{E}[B_{k+1} - B_k] \leqslant \frac{1}{2}\mathbb{E}[\overline{B}_k] + \frac{1}{2N}\mathbb{E}[B_k + C_{T_k}], \qquad 1 \leqslant k \leqslant N-1,$$

For $k \leqslant N-2$, Lemma 5 gives

$$\mathbb{E}[C_{T_k}] \leqslant K\sqrt{k+1},$$

while, for $k = N-1$, the trivial bound $C_{T_k} \leqslant 1$ suffices. Together with Lemma 2 and $k \geqslant k_\varepsilon \geqslant N/3$, this yields

$$\mathbb{E}[B_{k+1} - B_k] \leqslant \frac{K}{\sqrt{N}}, \qquad k_\varepsilon \leqslant k \leqslant N-1.$$

Summing and using (70), we obtain (67).

We next control the process $(C_k)_{k_\epsilon \leqslant k \leqslant N}$. Let P_k denote the positive variation accumulated during the stable part of the block $[S_k, S_{k+1})$, as in Lemma 9. Since the final jump of each block can only decrease C, for every $m \geqslant k_\varepsilon$,

$$C_m \leqslant C_{k_\varepsilon} + \sum_{k=k_\varepsilon}^{m-1} P_k.$$

Therefore, by Lemmas 2 and 9,

$$\begin{aligned}\mathbb{E}\left[\sup_{k_\varepsilon \leqslant m \leqslant N} C_m\right] &\leqslant \mathbb{E}[C_{k_\varepsilon}] + \sum_{k=k_\varepsilon}^{N-1} \mathbb{E}[P_k] \\ &\leqslant K\varepsilon\sqrt{N} + K\sum_{k=k_\varepsilon}^{N-1} \frac{1}{\sqrt{k}}.\end{aligned}$$

Using (70),

$$\sum_{k=k_\varepsilon}^{N-1} \frac{1}{\sqrt{k}} \leqslant \frac{N - k_\varepsilon}{\sqrt{k_\varepsilon}} \leqslant K\varepsilon\sqrt{N}.$$

This proves (68).

Finally, by monotonicity of B,

$$\sup_{k_\varepsilon \leqslant k \leqslant N} |B_k - B_{k_\varepsilon}| = B_N - B_{k_\varepsilon}.$$

Markov's inequality, together with (67) and (68), gives, for all sufficiently large N,

$$\mathbb{P}\left(\sup_{k_\varepsilon \leqslant k \leqslant N} \frac{|B_k - B_{k_\varepsilon}| + C_k}{\sqrt{N}} > \eta\right) \leqslant \frac{K\varepsilon}{\eta}.$$

Taking the upper limit as $N \to \infty$ proves (69). □

7.3 Proof of the Theorem

Proof of Theorem 1. Set

$$f(y) := \begin{pmatrix} b(y) \\ c(y) \end{pmatrix}, \qquad Z_N(y) := \frac{1}{\sqrt{N}} \begin{pmatrix} B_{k_N(y)} \\ C_{k_N(y)} \end{pmatrix}, \qquad 0 \leqslant y \leqslant 1.$$

Let $k_0 = k_0(N)$ be such that

$$1 \ll k_0 \ll N,$$

and set

$$A_N := M_{k_0}^{(N)}.$$

By Proposition 6,

$$A_N \Longrightarrow W. \tag{71}$$

In particular, $(A_N)_{N \geqslant 1}$ is tight.

Fix $\varepsilon \in (0, 1/2)$. By Proposition 9,

$$\sup_{k_0 \leqslant k \leqslant \lfloor (1-\varepsilon)N \rfloor} \left| \frac{1}{\sqrt{N}} \begin{pmatrix} B_k \\ C_k \end{pmatrix} - A_N f\left(\frac{k}{N}\right) \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Since f is uniformly continuous and A_N is tight,

$$\sup_{k_0/N \leqslant y \leqslant 1-\varepsilon} |Z_N(y) - A_N f(y)| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Moreover, Lemma 10, the continuity of f, together with the fact that $f(0) = 0$, $k_0/N \to 0$, and the tightness of A_N, give

$$\sup_{0 \leqslant y \leqslant k_0/N} |Z_N(y) - A_N f(y)| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Consequently,

$$D_N(\varepsilon) := \sup_{0 \leqslant y \leqslant 1-\varepsilon} |Z_N(y) - A_N f(y)| \xrightarrow[N\to\infty]{\mathbb{P}} 0. \tag{72}$$

It remains to control the terminal interval. Set

$$k_\varepsilon := \lfloor (1-\varepsilon)N \rfloor$$

and

$$T_N(\varepsilon) := \sup_{1-\varepsilon \leqslant y \leqslant 1} |Z_N(y) - Z_N(1-\varepsilon)|.$$

Since B_k is nondecreasing and $C_k \geqslant 0$,

$$T_N(\varepsilon) \leqslant \frac{B_N - B_{k_\varepsilon}}{\sqrt{N}} + \frac{1}{\sqrt{N}} \sup_{k_\varepsilon \leqslant k \leqslant N} C_k.$$

Therefore, by Lemma 11 (Equations (67) and (68)), for all sufficiently large N,

$$\mathbb{E}[T_N(\varepsilon)] \leqslant K\varepsilon. \tag{73}$$

Define

$$\omega_f(\varepsilon) := \sup_{1-\varepsilon \leqslant y \leqslant 1} |f(y) - f(1-\varepsilon)|.$$

Since f is continuous on $[0, 1]$,

$$\omega_f(\varepsilon) \longrightarrow 0 \qquad \text{as } \varepsilon \downarrow 0.$$

For $1-\varepsilon \leqslant y \leqslant 1$,

$$|Z_N(y) - A_N f(y)| \leqslant |Z_N(y) - Z_N(1-\varepsilon)| \quad + |Z_N(1-\varepsilon) - A_N f(1-\varepsilon)| \quad + |A_N|\,|f(y) - f(1-\varepsilon)|.$$

Together with (72), this gives

$$\|Z_N - A_N f\|_\infty \leqslant T_N(\varepsilon) + D_N(\varepsilon) + |A_N|\,\omega_f(\varepsilon). \tag{74}$$

Let $\eta > 0$. By (72), (73), (74) and Markov's inequality,

$$\limsup_{N\to\infty} \mathbb{P}\left(\|Z_N - A_N f\|_\infty > 3\eta\right) \leqslant \frac{K\varepsilon}{\eta} + \limsup_{N\to\infty} \mathbb{P}\left(|A_N|\,\omega_f(\varepsilon) > \eta\right).$$

Since (A_N) is tight and $\omega_f(\varepsilon) \to 0$, the second term on the right tends to 0 as $\varepsilon \downarrow 0$. Letting $\varepsilon \downarrow 0$ therefore yields

$$\|Z_N - A_N f\|_\infty \xrightarrow[N\to\infty]{\mathbb{P}} 0. \tag{75}$$

Now the map

$$r \longmapsto rf$$

is continuous from $\mathbb{R}$ into $D([0,1], \mathbb{R}^2)$ endowed with the Skorokhod J_1 topology. Hence, by (71) and the continuous mapping theorem,

$$A_N f \Longrightarrow Wf.$$

Moreover,

$$d_{J_1}(Z_N, A_N f) \leqslant \|Z_N - A_N f\|_\infty \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Slutsky's lemma therefore gives

$$Z_N \Longrightarrow Wf$$

in $D([0,1], \mathbb{R}^2)$, endowed with the Skorokhod J_1 topology.

The characterization of W, its uniqueness among L^2 solutions with mean 1, and its strict positivity follow from Proposition 5. □

AI use declaration This work was initiated in early 2025. By the summer, we had identified a roadmap for the proof: the early phase described by a RRT, followed by a regime in which the martingales plateau, together with a strategy for gluing the two phases. Many technical details, however, remained unchecked. AI (ChatGPT 5 and its subsequent versions) was subsequently used to help us derive the numerous lemmas needed to turn this roadmap into a complete argument. In particular, it was especially helpful in establishing the estimates surrounding the quadratic variation in the second phase (Section 6). The authors carefully verified all proofs and rewrote them where necessary. They take full responsibility for the final text and all mathematical results.

References

[1] Samuel Ayomide Adeosun and Peter Pfaffelhuber. Markov processes forced on a subspace by a large drift, with applications to population genetics, 2026.

[2] Nicholas H. Barton and Alison M. Etheridge. The relation between reproductive value and genetic contribution. *Genetics*, 188(4):953–973, 2011.

[3] Joseph T. Chang. Recent common ancestors of all present-day individuals. *Advances in Applied Probability*, 31(4):1002–1026, 1999.

[4] Camille Coron and Yves Le Jan. Genetic contribution of advantaged ancestors in the biparental moran model – finite selection. *arXiv:2502.01178*, 2025.

[5] Camille Coron and Yves Le Jan. Pedigree in the biparental moran model. *Journal of Mathematical Biology*, 84:51, 2022.

[6] Camille Coron and Yves Le Jan. Genetic contribution of an advantaged mutant in the biparental moran model. *Ukrainian Mathematical Journal*, 75(11), 2024.

[7] Bernard Derrida, Susanna C. Manrubia, and Damian H. Zanette. On the genealogy of a population of biparental individuals. *Journal of Theoretical Biology*, 203(3):303 – 315, 2000.

[8] Dimitrios Diamantidis, Wai-Tong (Louis) Fan, Matthias Birkner, and John Wakeley. Bursts of coalescence within population pedigrees whenever big families occur. *Genetics*, 227(1):iyae030, 2024.

[9] Michael Drmota and Hsien-Kuei Hwang. Profiles of random trees: correlation and width of random recursive trees and binary search trees. *Advances in Applied Probability*, 37(2):321–341, 2005.

[10] F. Foutel-Rodier, N. Barton, and A. Etheridge. Introgression under linear selection on continuous genomes. *bioRxiv*, 2026.01.30.702779, 2026.

[11] Michael Fuchs, Hsien-Kuei Hwang, and Ralph Neininger. Profiles of random trees: limit theorems for random recursive trees and binary search trees. *Algorithmica*, 46(3–4):367–407, 2006.

[12] Walter Gautschi. Some elementary inequalities relating to the gamma and incomplete gamma function. *Journal of Mathematics and Physics*, 38(1):77–81, 1959.

[13] S. Gravel and M. Steel. The existence and abundance of ghost ancestors in biparental populations. *Theor Popul Biol*, 101:47–53, 2015.

[14] Thomas G Kurtz. Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of applied Probability*, 7(1):49–58, 1970.

[15] Amaury Lambert, Verónica Miró Pina, and Emmanuel Schertzer. Chromosome painting: How recombination mixes ancestral colors. *The Annals of Applied Probability*, 31(2):826–864, 2021.

[16] Quansheng Liu. Asymptotic properties and absolute continuity of laws stable by random weighted mean. *Stochastic Processes and their Applications*, 95(1):83–107, 2001.

[17] Hosam M. Mahmoud. *Pólya Urn Models*. Chapman & Hall/CRC Texts in Statistical Science. Chapman & Hall/CRC, Boca Raton, FL, 2008.

[18] Frederick A. Matsen and Steven A. Evans. To what extent does genealogical ancestry imply genetic ancestry? *Theoretical Population Biology*, 2008.

[19] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

[20] Maximillian Newman, John Wakeley, and Wai-Tong (Louis) Fan. Conditional gene genealogies given the population pedigree for a diploid moran model with selfing. *Theoretical Population Biology*, 165:29–44, 2025.

[21] P. Pfaffelhuber and A. Wakolbinger. A diploid population model for copy number variation of genetic elements. *Electron. J. Probab.*, 28:1–15, 2023.

[22] John Maynard Smith and John Haigh. The hitch-hiking effect of a favourable gene. *Genetical Research*, 23(1):23–35, 1974.

[23] John Wakeley, Léandra King, Bobbi S. Low, and Sohini Ramachandran. Gene genealogies within a fixed pedigree, and the robustness of Kingman's coalescent. *Genetics*, 190(4):1433–1445, 2012.

[24] John Wakeley, Léandra King, and Peter R. Wilton. Effects of the population pedigree on genetic signatures of historical demographic events. *Proceedings of the National Academy of Sciences*, 113(29):7994–8001, 2016.