Electronic Journal of Probability

Electron. J. Probab. **18** (2013), no. 47, 1–17.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-2422

Fluctuations of martingales and winning probabilities of game contestants*

David Aldous† Mykhaylo Shkolnikov‡

Abstract

Within a contest there is some probability $M_i(t)$ that contestant i will be the winner, given information available at time t, and $M_i(t)$ must be a martingale in t. Assume continuous paths, to capture the idea that relevant information is acquired slowly. Provided each contestant's initial winning probability is at most b, one can easily calculate, without needing further model specification, the expectations of the random variables N_b = number of contestants whose winning probability ever exceeds b, and D_{ab} = total number of downcrossings of the martingales over an interval $[a, b]$. The distributions of N_b and D_{ab} do depend on further model details, and we study how concentrated or spread out the distributions can be. The extremal models for N_b correspond to two contrasting intuitively natural methods for determining a winner: progressively shorten a list of remaining candidates, or sequentially examine candidates to be declared winner or eliminated. We give less precise bounds on the variability of D_{ab}. We formalize the setting of infinitely many contestants each with infinitesimally small chance of winning, in which the explicit results are more elegant. A canonical process in this setting is the Wright-Fisher diffusion associated with an infinite population of initially distinct alleles; we show how this process fits our setting and raise the problem of finding the distributions of N_b and D_{ab} for this process.

Keywords: entrance boundary, fluctuations, martingale; upcrossing; Wright-Fisher diffusion.
AMS MSC 2010: Primary 60G44, Secondary 91A60.
Submitted to EJP on November 7, 2012, final version accepted on April 4, 2013.

1 Introduction

Given a probability distribution $\mathbf{p} = (p_i, i \geq 1)$ consider a collection of processes $(M_i(t), 0 \leq t < \infty, i \geq 1)$ adapted to a filtration $(\mathcal{F}_t)$ and satisfying
(i) $M_i(0) = p_i, i \geq 1$;
(ii) for each $t > 0$ we have $0 \leq M_i(t) \leq 1 \ \forall i$ and $\sum_i M_i(t) = 1$;
(iii) for each $i \geq 1$, $(M_i(t), t \geq 0)$ is a continuous path martingale;
(iv) there exists a random time $T < \infty$ a.s. such that, for some random I, $M_I(T) = 1$ and $M_j(T) = 0 \ \forall j \neq I$.

*Aldous's research supported by N.S.F Grant DMS-0704159.
†U.C. Berkeley, USA. E-mail: aldous@stat.berkeley.edu
‡E-mail: mshkolni@gmail.com

Call such a collection a $\mathbf{p}$-*feasible* process, and call the $M_i(\cdot)$ its *component martingales*. To motivate this definition, consider contestants in a contest which will have one winner at some random future time. Then the probability $M_i(t)$ that contestant i will be the winner, given information known at time t, must be a martingale as t increases. In this scenario all the assumptions will hold automatically except for path-continuity, which expresses the idea that information becomes known slowly.

In view of the fact that continuous-path martingales have long been a central concept in mathematical probability, it seems curious that this particular "contest" setting has apparently not previously been studied systematically. Moreover the topic is appealing at the expository level because it can be treated at any technical level. In an accompanying non-technical article for undergraduates [1] we show data on probabilities (from the Intrade prediction market) for candidates for the 2012 Republican U.S. Presidential Nomination. The data is observed values of the variables N_b and D_{ab} below, and one can examine the question of whether there was an unusually large number of candidates that year whose fortunes rose and fell substantially. In this paper, the proof in section 4 of distributional bounds on N_b is mostly accessible to a student taking a first course in continuous-time martingales, and subsequent sections slowly become more technically sophisticated.

The starting point for this paper is the observation that there are certain random variables associated with a $\mathbf{p}$-feasible process whose *expectations* do not depend on the actual joint distribution of the component martingales, and indeed depend very little on $\mathbf{p}$. For $0 < a < b < 1$ consider

$$N_b := \text{ number of } i \text{ such that } \sup_t M_i(t) \geq b$$

$$D_{a,b} := \text{ sum over } i \text{ of the number of downcrossings of } M_i(\cdot) \text{ over } [a,b].$$

Straightforward uses of the optional sampling theorem (described verbally in [1] as gambling strategies) establish

Lemma 1.1. *If $\max_i p_i \leq b$ then for any $\mathbf{p}$-feasible process,*

$$\mathbb{E}[N_b] = 1/b, \quad \mathbb{E}[D_{a,b}] = (1-b)/(b-a).$$

In contrast, the *distributions* of N_b and $D_{a,b}$ will depend on the joint distributions of the component martingales, and one goal of this paper is to study the extremal possibilities. Here is our result for N_b.

Proposition 1.2. *(a) If $\max_i p_i \leq b$ then there exists a $\mathbf{p}$-feasible process for which the distribution of $N_b^{\mathbf{p}}$ is supported on the integers $\lfloor 1/b \rfloor$ and $\lceil 1/b \rceil$ bracketing its mean $1/b$.*
(b) There exists a family, that is a $\mathbf{p}$-feasible process for each $\mathbf{p}$, such that the distributions of $N_b^{\mathbf{p}}$ satisfy

$$\text{dist}(N_b^{\mathbf{p}}) \to \text{Geometric}(b) \text{ as } \max_i p_i \to 0. \tag{1.1}$$

(c) Any possible limit distribution for $N_b^{\mathbf{p}}$ as $\max_i p_i \to 0$ has variance at most $(1-b)/b^2$, the variance of $\text{Geometric}(b)$.

Clearly the distribution in (a) is the "most concentrated" possible, and part (c) gives a sense in which the $\text{Geometric}(b)$ distribution is the "most spread out" distribution possible. The proof will be given in section 4. The construction for (a) formalizes the idea that we maintain a list of candidates still under consideration, and at each stage choose one candidate to be eliminated. The construction for (b) formalizes the idea that we examine candidates sequentially, deciding to declare the current candidate to be the winner or to be eliminated. Returning briefly to the theme that this topic is amenable to popular exposition, with some imagination one can relate these two alternate ideas to those used in season-long television shows. Shows like *Survivor* overtly follow the idea for (a), whereas the idea for (b) would correspond to a variant of ...... *Millionaire* in which contestants were required to try for the million dollar prize and where the season ends when the prize is won.

We give an analysis of downcrossings D_{ab} in section 5, though with less precise results. The construction that gave the Geometric limit distribution for N_b in (1.1) also gives a Geometric limit distribution for D_{ab} (Proposition 5.1). We conjecture this is the maximum-variance possible limit, but can give only a weaker bound in Proposition 5.3. As for minimum-variance constructions, Proposition 5.4 shows one can construct feasible processes for which, in the limit as $b \to 0$ with a/b bounded away from 1, the variance of D_{ab} is bounded by a constant depending only on a/b. The case $a/b \approx 1$ remains mysterious, but prompts novel open problems about negative correlations for Brownian local times – see section 7.

1.1 $\mathbf{0}$-feasible processes

As a second goal of this paper, it seems intuitively clear that the concept of $\mathbf{p}$-feasible process can be taken to the limit as $\max_i p_i \to 0$, to represent the idea of starting with an infinite number of contestants each with only infinitesimal chance of winning. Informally, we define a $\mathbf{0}$-feasible process as a process with the properties:

(i) for each $t_0 > 0$, conditional on $M_i(t_0) = p_i, i \geq 1$, the process $(M_i(t_0 + t), 0 \leq t < \infty, i \geq 1)$ is a $\mathbf{p}$-feasible process;

(ii) $\sup_i M_i(t) \to 0$ a.s. as $t \downarrow 0$.

There is some subtlety in devising a precise definition, which we will give in section 3. The theory developed there will allow us to deduce results for general $\mathbf{0}$-feasible processes as limits of results for $\mathbf{p}$-feasible processes under the regime $\max_i p_i \to 0$, and also to construct specific $\mathbf{0}$-feasible processes by splicing together specific $\mathbf{p}$-feasible processes under the same regime (Proposition 3.6).

By eliminating any dependence on $\mathbf{p}$, results often become cleaner for $\mathbf{0}$-feasible processes. For instance Proposition 1.2 becomes

Corollary 1.3. *(a) There exists a $\mathbf{0}$-feasible process such that, for each $0 < b < 1$, the distribution N_b is supported on the integers $\lfloor 1/b \rfloor$ and $\lceil 1/b \rceil$ bracketing its mean $1/b$.*
(b) Given $0 < b_0 < 1$, there exists a $\mathbf{0}$-feasible process such that, for each $b_0 \leq b < 1$, N_b has $\mathrm{Geometric}(b)$ distribution.
(c) Moreover for any $\mathbf{0}$-feasible process and any $0 < b < 1$ the variance of N_b is at most $(1-b)/b^2$, the variance of $\mathrm{Geometric}(b)$.

Setting aside the “extremal" questions we have discussed so far, another motivation for considering the class of $\mathbf{0}$-feasible processes is that there is one particular such process which we regard intuitively as the “canonical" choice, and this is the $\mathbf{0}$-*Wright-Fisher process* discussed in section 6. This connection between (a corner of) the large literature on processes inspired by population genetics and our game contest setting seems not to have been developed before. In particular, questions about the fluctuation behavior of the $\mathbf{0}$-Wright-Fisher process – the distributions of N_b and D_{ab} – arise more naturally in the contest setting, though it seems hard to get quantitative estimates of these distributions.

2 Preliminary observations

2.1 The downcrossing formula

In our setting of a continuous-path martingale $M(\cdot)$ ultimately stopped at 0 or 1, recall the "fair game formula"

$$\mathbb{P}(M(t) \text{ hits } b \text{ before } a \,|M(0)=x) = \tfrac{x-a}{b-a},\ 0 \le a \le x \le b \le 1 \tag{2.1}$$

from which one can readily derive the well known formula for the expectation of the number D of downcrossings of $M(\cdot)$ over $[a,b]$: for $0 \le a \le b \le 1$,

$$\mathbb{E}[D|M(0)=x] = \tfrac{x(1-b)}{b-a} \quad \text{if } 0 \le x \le b \tag{2.2}$$

$$= \tfrac{b(1-x)}{b-a} \quad \text{if } b \le x \le 1. \tag{2.3}$$

Moreover, starting from b there is a modified Geometric distribution for D:

$$\mathbb{P}(D=0|M(0)=b) = \tfrac{b-a}{1-a}$$

$$\mathbb{P}(D=d|M(0)=b) = \tfrac{1-b}{1-a} \left(\tfrac{a(1-b)}{b(1-a)}\right)^{d-1} \left(1-\tfrac{a(1-b)}{b(1-a)}\right), \quad d \ge 1. \tag{2.4}$$

2.2 The multivariate Wright-Fisher diffusion

Textbooks introducing discrete-time martingales often use as an example (e.g. [7] Example 10.2.6) the discrete-time Wright-Fisher model for genetic drift of a single allele. Note that throughout what follows, we consider only the case of no mutation and no selection. It is classical that the infinite population limit of the k-allele model is the multivariate Wright-Fisher diffusion on the $k-1$-dimensional simplex, that is with generator

$$\tfrac{1}{2} \sum_{i,j=1}^{k} x_i(\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \partial x_j}. \tag{2.5}$$

Each component is a martingale, the one-dimensional diffusion on $[0,1]$ with drift rate zero and variance rate $x(1-x)$. There has been extensive work since the 1970s on the infinitely-many-alleles case, but this has focussed on the case of positive mutation rates to novel alleles, in which case the martingale property no longer holds. In our setting (no mutation and no selection) it is straightforward to show directly (see section 6) that for any $\mathbf{p} = (p_i, i \ge 1)$ with countable support there exists what we will call the $\mathbf{p}$-Wright-Fisher process, the infinite-dimensional diffusion with generator analogous to (2.5) starting from state $\mathbf{p}$, and that this is a $\mathbf{p}$-feasible process. So we know that $\mathbf{p}$-feasible processes do actually exist, and these $\mathbf{p}$-Wright-Fisher processes will be useful ingredients in later constructions. (When $\mathbf{p}$ has finite support we could use instead Brownian motion on the finite-dimensional simplex, whose components are killed at 0 and 1, but this does not extend so readily to the infinite-dimensional setting).

It is convenient to adopt from genetics the phrase *fixation time* for the time T at which the winner is determined.

2.3 Constructions using Wright-Fisher

In a Wright-Fisher diffusion we have $\sum_i M_i(t) \equiv 1$, but trivially we can consider a rescaled Wright-Fisher diffusion for which $\sum_i M_i(t)$ is a prescribed constant.

Our constructions of feasible processes typically proceed in stages. Within a stage we may declare that some component martingales are "frozen" (held constant) and the others evolve as a rescaled Wright-Fisher process. In particular if only two component martingales are unfrozen, say at the start S of the stage we have $M_i(S) = x_i$ and

$M_j(S) = x_j$, then during the stage we have a "reflection coupling" with $M_i(t) + M_j(t) = x_i + x_j$, and we can choose to continue the stage until the processes reach $x_i + x_j$ and 0, or we can choose to stop earlier.

An alternative construction method is to select one component martingale $M_i(S)$ at the start of the stage, let $(M_i(\cdot), 1-M_i(\cdot))$ evolve as the two-allele Wright-Fisher process during the stage, and set $M_j(\cdot) = \frac{M_j(S)}{1-M_i(S)} \times (1 - M_i(\cdot))$. We describe this construction by saying that the processes $(M_j(\cdot), j \neq i)$ are *tied*.

Both constructions clearly give continuous-path martingale components.

3 0-feasible processes

In section 3.2 we will give one formalization of the notion of a $\mathbf{0}$-feasible process introduced informally in section 1.1, and in sections 3.4 and 3.5 we give results allowing one to relate constructions and properties of $\mathbf{0}$-feasible processes to those of $\mathbf{p}$-feasible processes. A reader who is content with the informal description may skip this section, because our main results in sections 4 and 5 are based on concrete calculations and constructions with $\mathbf{p}$-feasible processes.

3.1 What is the issue?

The following observation shows that the most naive formalization does not work.

Lemma 3.1. *Let I be countable, There does not exist any process $(M_i(t), 0 \le t < \infty, i \in I)$ adapted to a filtration $(\mathcal{F}_t)$ and satisfying*
(i) for each $t > 0$ we have $0 \le M_i(t) \le 1\ \forall i$ and $\sum_i M_i(t) = 1$;
(ii) for each i, $(M_i(t), t \ge 0)$ is a martingale;
(iii) $\sup_i M_i(t) \to 0$ a.s. as $t \downarrow 0$.

Proof. The martingale property implies $\mathbb{E}[M_i(t)]$ is constant in t. But by (i) and (iii) we have

$$\lim_{t \downarrow 0} \mathbb{E}[M_i(t)] = 0.$$

So $\mathbb{E}[M_i(t)] = 0$ for all i and t, contradicting (i). □

Another way to see the difficulty is to consider how to formalize the "Survivor" model mentioned in the Introduction: maintain a list of contestants still under consideration, and at each stage choose (uniformly at random) one contestant to be eliminated. This process clearly makes sense (in discete time) for any finite initial number k of contestants. We would like to argue "these processes are Kolmogorov-consistent as k increases, and therefore there exists a single process which contains these processes simultaneously for all initial $2 \le k < \infty$". But Lemma 3.1 implies there is no way to define such a process using a countable set I of names of contestants, because then the conditional probabilities $M_i(t)$ of contestant i being the winner would (after an easy embedding into continuous time) satisfy the conditions of the lemma.

Our formalization issue is very similar to those arising in the theory of stochastic fragmentation and coagulation processes [3], where a state of the process is (conceptually) an unordered collection of "clusters" whose masses x_α sum to 1. Of the three formalizations described in [4], two could be used in our setting:
(i) ranking (decreasing-ordering) the cluster sizes (x_α);
(ii) using an interval-partition (a partition of the unit interval into sub-intervals of lengths x_α).
A third device, using the induced partition on sampled "atoms" and exploiting exchangeability, seems not applicable in our setting.

The formalization we choose below combines ranking and a point process representation. We defer further discussion until section 3.3.

3.2 A formalization

A probability distribution $\mathbf{p}$ with $p_1 \geq p_2 \geq p_3 \geq \ldots$ is called *ranked*; write ∇ for the space of ranked probability distributions. For a general discrete distribution $\mathbf{q} = (q_j, j \in J)$ write $\mathrm{rank}(\mathbf{q})$ for its decreasing ordering, where zero entries are omitted. More generally, for a collection $(A_j, j \in J)$ of objects with the same index set as $(q_j, j \in J)$, write $\mathrm{rank}(A_j, j \in J||\mathbf{q})$ for the collection re-ordered so that $\mathbf{q}$ is ranked (this is not completely specified if the values q_j are not distinct, but the arbitrariness does not matter for our purposes).

Write C_0 for the space of continuous functions $f : [0, \infty) \to [0, 1]$ with $f(0) = 0$. Consider a random point process on C_0. That is, a realization of the process is (informally) an unordered countable set $\{f_\alpha(\cdot)\}$ of functions or (formally) the counting measure associated with that set. We will use the former notation, which is more intuitive. We define a $\mathbf{0}$-feasible process to be a random point process $\{M_\alpha(\cdot)\}$ on C_0 such that

$$0 \leq M_\alpha(t) \leq 1; \quad \sum_\alpha M_\alpha(t) = 1, \quad 0 < t < \infty$$

$$\max_\alpha M_\alpha(t) \to 0 \text{ a.s. as } t \downarrow 0$$

and with the following property. For each $t_0 > 0$ and each ranked $\mathbf{p}$,

Conditional on $\mathrm{rank}(M_\alpha(t_0)) = \mathbf{p}$ and on $\mathcal{F}(t_0)$, the ranked process

$$\mathrm{rank}(M_\alpha(t_0 + \cdot))||\{M_\alpha(t_0)\}) \text{ is } \mathbf{p}\text{-feasible.} \tag{3.1}$$

In words, given t_0 we simply label component martingales as $1, 2, 3, \ldots$ in decreasing order of their values at t_0, and we can use this labeling over $t_0 \leq t < \infty$ to define a process $(M_i(t_0 + t), t \geq 0, i \geq 1)$ which we require to be a $\mathbf{p}$-feasible process, where $\mathbf{p}$ is the ranked ordering of $(M_\alpha(t_0))$. For $\mathcal{F}_t$ we take the natural filtration, generated by the restriction of the point process to $(0, t]$.

By standard arguments, property (3.1) extends to any stopping time S with $0 < S < \infty$:

Conditional on $\mathrm{rank}(M_\alpha(S)) = \mathbf{p}$ and on $\mathcal{F}(S)$, the ranked process

$$\mathrm{rank}(M_\alpha(S + \cdot))||\{M_\alpha(S)\}) \text{ is } \mathbf{p}\text{-feasible.} \tag{3.2}$$

In our initial definition of a $\mathbf{p}$-feasible process we assumed the initial configuration $\mathbf{p}$ was deterministic. Now define a $\oplus$-feasible process to be a mixture over $\mathbf{p}$ of $\mathbf{p}$-feasible processes; in other words, a process $(M_i(t), i \geq 1, t \geq 0)$ which, conditional on $(M_i(0), i \geq 1) = (p_i, i \geq 1)$, is a $\mathbf{p}$-feasible process. So the ranked process $\mathrm{rank}(M_\alpha(S + \cdot))||\{M_\alpha(S)\})$ in (3.2), considered unconditionally, is a $\oplus$-feasible process, and we describe the relationship (3.2) by saying this $\oplus$-feasible process is *embedded* into the $\mathbf{0}$-feasible process via the stopping time S. Similarly, any stopping time within a $\oplus$-feasible process specifies an embedded $\oplus$-feasible process.

3.3 Remarks of the formalization

Remark 3.2. *Our formalization above is admittedly somewhat* ad hoc. *An essentially equivalent formalization would be to assign random $U[0, 1]$ labels U_α to component martingales, so the state of the process at t is described via the pairs $(U_\alpha, M_\alpha(t))$ for which $M_\alpha(t) > 0$, and this can in turn be described via the probability measure $\sum_\alpha M_\alpha(t)\delta_{U_\alpha}$ or its distribution function. We will use this "random labels" idea in an argument below.*

Remark 3.3. *There are several possible choices for the level of generality we might adopt. The "canonical" examples of the* **0***-Wright-Fisher process (section 6), and the "*Survivor*" process featuring in the proof (section 4) of Corollary 1.3(a), have the property that at times $t > 0$ the process has only finitely many non-zero components, so we could make this a requirement. But our formalization allows a countable number of non-zero components. In the other direction, consider the construction of reflecting Brownian motion $R(t)$ from standard Brownian motion $W(t)$ as*

$$R(t) := W(t) - \min_{s \le t} W(s)$$

and run the process until $R(\cdot)$ hits 1. Within our setting, interpret this as saying that at time t there is one contestant with chance $R(t)$ of winning, the remaining chance $1 - R(t)$ being split amongst an infinite number of unidentified contestants each with only infinitesimal chance of winning. Informally this is a **0***-feasible process such that*

$$N_b \text{ has Geometric}(b) \text{ distribution for every } 0 < b < 1, \tag{3.3}$$

strengthening the assertion of Corollary 1.3(b), but it does not fit our set-up which requires the unit mass to be split as a random discrete distribution at times $t > 0$. In fact Corollary 4.2 implies that, within our formalization, no **0***-feasible process can have property (3.3). One could perhaps choose a more general formalization, in the spirit of the interval-partitions used in stochastic coalescence [4], which does allow such "dust", but we have not pursued that idea.*

Remark 3.4. *The existing classes of processes in the literature with somewhat similar qualitative behavior – in the theory of stochastic fragmentation and coagulation processes, or in population genetics inspired processes associated with Kingman's coalescent, are (to our knowledge) explicitly Markovian, in which context the question becomes determining the* entrance boundary *of a specific Markov process [2, 5]. Our setting differs in that we wish to continue making only the "martingale" assumptions (ii,iii,iv) at the start of the Introduction, and we are seeking to define a class of processes.*

Remark 3.5. *Within a* **p***-feasible process, we refer to a particular component process $(M_i(t), 0 \le t < \infty)$ as a* martingale component*. For a* **0***-feasible process this terminology is somewhat misleading, in that the technical issue emphasized in section 3.1 is that we cannot label the components to make each component process be a martingale over the entire interval $0 < t < \infty$. However, from the definition we can do so over any interval $\varepsilon \le t < \infty$.*

3.4 A general construction of 0-feasible processes

Given a **0**-feasible process and stopping times $S_k \downarrow 0$ a.s., the associated embedded $\oplus$-feasible processes are embedded within each other, and their initial values $(M_i^{(k)}, i \ge 1)$ satisfy $\max_i M_i^{(k)} \to 0$ a.s.. The following result formalizes the converse idea: one can construct a **0**-feasible process from a sequence of **p**-feasible or more generally $\oplus$-feasible processes embedded into each other, via Kolmogorov consistency.

Proposition 3.6. *Suppose that $(\mu_k, k \ge 1)$ are probability measures on ∇ and that for each k there are families $(M_i^k(t), i \ge 1, 0 \le t < \infty)$ such that*
(i) $(M_i^k(0), i \ge 1)$ has distribution μ_k.
(ii) Conditional on $(M_i^k(0), i \ge 1) = \mathbf{p}$, the process $\mathbf{M}^k = (M_i^k(t), 0 \le t < \infty, i \ge 1)$ is **p***-feasible.*
(iii) For $k \ge 2$ there is a stopping time T_k for $\mathbf{M}^k$ such that $t_k := \mathbb{E}[T_k] < \infty$ and

$\text{rank}(M_i^k(T_k), i \geq 1)$ *has distribution* μ_{k-1}.
(iv) $\sum_k t_k < \infty$.
(v) $M_1^k(0) \to_p 0$ *as* $k \to \infty$.
Then there exists a $\mathbf{0}$*-feasible process* $\{M_\alpha(\cdot)\}$ *which is consistent with the families above, in the following sense. There exist stopping times* S_k *such that for each* $k \geq 1$

$$\mathbb{E}[S_k] = \sum_{j>k} t_j, \quad S_k - S_{k+1} \stackrel{d}{=} T_{k+1}$$

and the embedded process $\text{rank}(M_\alpha(S_k + \cdot))||\{M_\alpha(S_k)\})$ *is distributed as* $M^k(\cdot)$.

Proof. By conditions (i)-(iii), for each $k \geq 2$ we can represent the process $\mathbf{M}^{k-1}$ as the process $\mathbf{M}^k(T_k + \cdot)$; more precisely, we can couple the two processes such that

$$M_i^{k-1}(t) = \text{rank}(M_i^k(T_k + t)||(M_i^k(T_k),\ i \geq 1)). \tag{3.4}$$

Then by the Kolmogorov consistency theorem we can assume this representation holds simultaneously for all k. We now attach labels α to the component martingales by the following inductive scheme. For $k = 1$, to each of the indices i designating a component martingale $M_i^1(\cdot)$ we associate an independent $\text{Uniform}(0, 1)$ label. For $k = 2$, a component martingale $M_i^2(\cdot)$ might be zero or non-zero at T_2. If non-zero then we copy the label already associated within $M^1(\cdot)$ via the coupling (3.4). If zero the we create a new independent $\text{Uniform}(0, 1)$ label.

Continue for each k this scheme of copying or creating labels. For each label α, the sample path of that martingale component in the process M^{k+1} is obtained from the sample path in M^k by inserting an extra initial segment. By (iv) the path converges as $k \to \infty$ to a function $M_\alpha(t), 0 \leq t < \infty$, and by (v) we must have $M_\alpha(0) = 0$. The remaining properties are straightforward. □

3.5 All $\mathbf{p}$-feasible processes embed

Proposition 3.6 enables construction of specific $\mathbf{0}$-feasible processes. The following result implies that any $\mathbf{p}$-feasible process can be embedded into some $\mathbf{0}$-feasible process – simply splice the $\mathbf{0}$-feasible process in the proposition to the given $\mathbf{p}$-feasible process at time S. We will use this fact in the proofs of Corollary 1.3(b) and Proposition 5.4.

Proposition 3.7. *Given any ranked* $\mathbf{p}$*, there exists a* $\mathbf{0}$*-feasible process* $\{M_\alpha(\cdot)\}$ *and a stopping time* S *such that* $\text{rank}(\ \{M_\alpha(S)\}\ ) = \mathbf{p}$.

For the proof it is convenient to use Brownian-type process instead of Wright-Fisher. Write

$$Q(t) := \sum_i (M_i(t))^2.$$

We will use constructions with the property

$$\text{At each time } 0 \leq t \leq S, \text{ at least one component martingale } M_i(t) \text{ is evolving as Brownian motion} \tag{3.5}$$

for a specified stopping time S. That is, our constructions can be written as

$$dM_i(t) = \sigma_i(t) dW_i(t)$$

for (dependent) standard Brownian motions $W_i(t)$, and we require that some $\sigma_i(t)$ equals 1. In general $Q(t) - \int_0^t \sum_i \sigma_i^2(s)\ ds$ is a martingale, so the advantage of property (3.5) is that $Q(t) - t$ is a submartingale, implying

Lemma 3.8. *Let $(M_i(t))$ be a $\mathbf{p}$-feasible process satisfying (3.5) for a stopping time S. Then $\mathbb{E}[S] \le \mathbb{E}[Q(S)] - Q(0)$.*

A simple construction satisfying (3.5) is the *Brownian reflection coupling* of two component martingales. That is, on $0 \le t \le S$ we freeze components other than i, j, and set

$$M_i(t) - M_i(0) = W_i(t), \quad M_j(t) - M_j(0) = -W_i(t).$$

Lemma 3.9. *Let I_0 be countable, and I_1 and I_2 be finite, index sets. Let $(p_i, i \in I_0 \cup I_1)$ and $(q_i, i \in I_0 \cup I_2)$ be probability distributions which coincide on I_0 and satisfy $\max_{i \in I_1} p_i \le \min_{i \in I_2} q_i$. Then there exists a $\mathbf{p}$-feasible process $\{M_\alpha(\cdot)\}$ satisfying (3.5) such that for some stopping time S we have $\mathrm{rank}(\ \{M_\alpha(S)\}\ ) = \mathrm{rank}(\mathbf{q})$.*

Proof. Freeze permanently the component martingales with $i \in I_0$. Pick two arbitrary indices i', i^* in I_1 and run the Brownian reflection coupling on these two components $M_{i'}(t), M_{i^*}(t)$ until one component hits zero or $\min_{i \in I_2} q_i$. In the latter case, freeze that component permanently and delete its index from I_1 and delete $\arg\min_{i \in I_2} q_i$ from I_2. In the former case, only delete the index from I_1. The total number (originally $|I_1| + |I_2|$) of unfrozen components is now decreased by at least 1. Continue inductively, picking two components from I_1 at each stage. Eventually all components are frozen and the ranked state is $\mathrm{rank}(\mathbf{q})$. □

Proof of Proposition 3.7. Define $\mathbf{p}^0 = \mathbf{p}$ and for $k \ge 1$ construct $\mathbf{p}^k$ from $\mathbf{p}$ by
(i) retaining entries p_i with $p_i \le 2^{-k}$;
(ii) replacing other p_i by $2^{j(i)}$ copies of $2^{-j(i)} p_i$, where $j(i) \ge 1$ is the smallest integer such that $2^{-j(i)} p_i \le 2^{-k}$.
Each pair $(\mathbf{p}^k, \mathbf{p}^{k-1})$ satisfies the hypothesis of Lemma 3.9. So for each k, writting $\mu_k = \delta_{\mathbf{p}^k}$ and writing $\mathbf{M}^k$ and T_k for the $\mathbf{p}^k$-feasible process and the stopping time given by Lemma 3.9, we see that hypotheses (i)-(iii) of Proposition 3.6 are satisfied. Moreover by Lemma 3.8 we have $\mathbb{E}[T_k] \le q_{k-1} - q_k$ for $q_k := \sum_i (p_i^k)^2$, implying that hypotheses (i)-(iii) are also satisfied. The conclusion of Proposition 3.6 now establishes Proposition 3.7. □

4 Proofs of distributional bounds on N_b

In this section we will prove Proposition 1.2 and Corollary 1.3, stated in the Introduction.

Proof of Proposition 1.2(a). Fix b. Run a Wright-Fisher process started at $\mathbf{p}$ until some $M_i(\cdot)$ reaches b. Freeze that i and run the remaining processes as rescaled Wright-Fisher until some other $M_j(\cdot)$ reaches b. Freeze that j and continue. After a finite number of such stages we must reach a state where all component martingales except one are frozen at b or at 0, and the remaining one is in $[0, b]$. Because $\sum_i M_i(t) \equiv 1$ the number frozen at b must be $\lfloor 1/b \rfloor$ and the remaining martingale must be at $1 - b\lfloor 1/b \rfloor$. Finally, unfreeze and run from this configuration to fixation as Wright-Fisher. Clearly N_b takes only the values $\lfloor 1/b \rfloor$ and $\lceil 1/b \rceil$. □

Proof of Corollary 1.3(a). This construction is similar to that above, but is closer to our earlier informal description "maintain a list of candidates still under consideration, and at each stage choose one candidate to be eliminated".

For each integer $m \ge 2$, we will define a stage which starts with m component martingales at $1/m$, and ends with $m-1$ of these martingales at $1/(m-1)$ and the other frozen at 0. To construct this stage, run as Wright-Fisher until some $M_i(\cdot)$ reaches $1/(m-1)$. Freeze that i and run the remaining martingales as rescaled Wright-Fisher until some other $M_j(\cdot)$ reaches $1/(m-1)$. Freeze that j and continue. Eventually we must reach a state where $m-1$ martingales are frozen at $1/(m-1)$ and the remaining process is 0. This stage takes some random time τ_m with finite expectation; without needing to calculate it, we can simply rescale time so that $\mathbb{E}[\tau_m] = 2^{-m}$.

Intuitively, we simply put these stages together, to obtain a $\mathbf{0}$-feasible process in which, for each $M \geq 1$, at time $\sum_{m>M} \tau_m$ there are exactly M martingales at $1/M$. Proposition 3.6 formalizes this construction, as it was designed to do. The resulting $\mathbf{0}$-feasible process satisfies the assertion of the Corollary for each $b = 1/M$, and then for general b because N_b is monotone in b. □

Proof of Proposition 1.2(b). Fix b. Write $\mathbf{p}$ in ranked order $p_1 \geq p_2 \geq \ldots$, and write J for the first term (if any) such that $p_{J+1}/(1-p_1-\ldots-p_J) > b$.

We use the "tied" construction from subsection 2.3. Run $(M_1(\cdot), 1-M_1(\cdot))$ as Wright-Fisher started from $(p_1, 1-p_1)$ and stopped at $S_1 := \min\{t : M_1(t) = 0 \text{ or } 1\}$, and for $i \geq 2$ set

$$M_i(t) = \tfrac{p_i}{1-p_1}(1 - M_1(t)),\ 0 \leq t \leq S_1.$$

So $M_i(\cdot)$ is a martingale on this time interval. Note that if $J \neq 1$ then no $M_i(\cdot)$ can reach b before time S_1, for $i \geq 2$.

If $M_1(S_1) = 1$ the process stops. If $M_1(S_1) = 0$ then for $i \geq 2$ we have $M_i(S_1) = p_i/(1-p_1)$. For $t \geq S_1$ run $(M_2(\cdot), 1-M_2(\cdot))$ as Wright-Fisher started from $(\frac{p_2}{1-p_1}, \frac{1-p_1-p_2}{1-p_1})$ and stopped at $S_2 := \min\{t : M_2(t) = 0 \text{ or } 1\}$, and for $i \geq 3$ set

$$M_i(t) = \tfrac{p_i}{1-p_1-p_2}(1 - M_2(t)),\ S_1 \leq t \leq S_2.$$

If $J \neq 2$ then no $M_i(\cdot)$ can reach b before time S_2, for $i \geq 3$.

Continue in this way to define processes $(M_j(t),\ S_{j-1} \leq t \leq S_j)$ for $1 \leq j \leq J$, or until some $M_j(\cdot)$ reaches 1 and the whole process stops. If the process has not stopped by time S_J, continue in an arbitrary manner, which makes the resulting process $\mathbf{p}$-feasible. Note that, if $M_j(\cdot)$ reaches b, then with probability exactly b it will reach 1, and that with probability $1 - \sum_{j \leq J} p_j$ the process has not stopped by time S_J.

Write $N_b^{(J)} =$ number of martingales $j \leq J$ that reach b. We can now apply Lemma 4.1 below to $Z = N_b^{(J)}$ with $q_i = \frac{1}{b} \cdot \frac{p_i}{1-p_1-\ldots-p_{i-1}}$, $1 \leq i \leq J$ and deduce that $N_b^{(J)} \leq Z' \stackrel{d}{=} \mathrm{Geometric}(b)$ with Z' constructed in Lemma 4.1. Then

$$\begin{aligned}
\mathbb{P}(N_b \neq Z') &\leq \mathbb{E}[|N_b - Z'|] \\
&\leq \mathbb{E}[Z' - N_b^{(J)}] + \mathbb{E}[N_b - N_b^{(J)}] \\
&= b^{-1}(1 - \sum_{j \leq J} p_j) + b^{-1}(1 - \sum_{j \leq J} p_j) \\
&= 2b^{-1} \sum_{j>J} p_j.
\end{aligned}$$

Hereby, the expectation $\mathbb{E}[Z' - N_b^{(J)}]$ can be computed to $b^{-1}(1 - \sum_{j \leq J} p_j)$ by noting that the random variable $Z' - N_b^{(J)}$ is only non-zero if the algorithm of Lemma 4.1 below has not terminated by the J-th step and that given the latter event of probability $(1 - \sum_{j \leq J} p_j)$ the quantity $Z' - N_b^{(J)}$ has the Geometric(b) distribution. Finally, as $\mathbf{p}$ varies we have

$$\text{if } \max_i p_i \to 0 \text{ then } \sum_{j > J(\mathbf{p})} p_j \to 0$$

establishing the limit result (1.1). □

Lemma 4.1. *Given $0 < b < 1$ and probabilities $q_i, 1 \le i \le J$ define a counting process by: for each i, given not yet terminated,*

with probability $q_i b$, increment count by 1 and terminate;
with probability $q_i(1-b)$, increment count by 1 and continue;
with probability $1-q_i$, continue.

Let Z be the value of the counting process after step J or at time T (the termination time, if any), whichever occurs first. Then there exists $Z' \stackrel{d}{=} \mathrm{Geometric}(b)$ such that $Z \le Z'$.

Proof. Augment the process by setting $q_i = 1,\ i > J$ and follow the algorithm for all $i \ge 1$. The process must now terminate at some a.s. finite time T', at which time the value Z' of the counting process has exactly Geometric(b) distribution. Indeed, it suffices to consider only those values of i for which the count is incremented. For every such step, the conditional probability of incrementing the count and terminating is b, whereas the conditional probability of incrementing the count and continuing is $(1-b)$ and such events are conditionally independent for different steps. Since Z' counts the number of such steps until the process terminates, it has the Geometric(b) distribution. □

Proof of Proposition 1.2(c). Fix b and for $k \ge 1$ let $S_k \le \infty$ be the first time at which k distinct component martingales have reached b. If $N_b \ge k$, then at time S_k one martingale takes value b, the other $k-1$ that previously reached b take some values $Z_1, \ldots, Z_{k-1}$, and the remaining martingales take some values $M_j(S_k) < b$. The chance that such a remaining martingale subsequently reaches b equals $M_j(S_k)/b$, and so, on $\{N_b \ge k\}$,

$$\mathbb{E}[N_b - k|\mathcal{F}_{S_k}] = b^{-1}\sum_j M_j(S_k) = b^{-1}\left(1 - b - \sum_{j=1}^{k-1} Z_j\right) \le \frac{1-b}{b}. \tag{4.1}$$

So

$$\mathbb{E}[(N_b-k)_+] \le \tfrac{1-b}{b}\ \mathbb{P}(N_b \ge k)$$

and summing over $k \ge 1$ gives

$$\mathbb{E}\left[\tfrac{N_b(N_b-1)}{2}\right] \le \tfrac{1-b}{b}\mathbb{E}[N_b] = \tfrac{1-b}{b^2}.$$

Finally,

$$\mathrm{var}(N_b) = 2\mathbb{E}[\tfrac{N_b(N_b-1)}{2}] + \mathbb{E}[N_b] - (\mathbb{E}N_b)^2 \le \tfrac{1-b}{b^2}.$$

□

For later use (section 5) note that to have equality in the final display above we need equality in (4.1), implying each $Z_j = 0$, that is the martingale components that previously reached b have all reached zero. We deduce

Corollary 4.2. *If, for a $\mathbf{p}$-feasible process, N_b has Geometric(b) distribution, then there is no time at which more than one component martingale is in $[b, 1]$.*

Proof of Corollary 1.3(c). This follows from Proposition 1.2(c) and the definition (3.1) of $\mathbf{0}$-feasible process via embedded $\mathbf{p}$-feasible processes. □

Proof of Corollary 1.3(b). Given b_0, consider the vector $\mathbf{p}$ of Geometric probabilities with

$$p_i = b_0(1-b_0)^{i-1}, i \ge 1. \tag{4.2}$$

Recall the construction in the proof of Proposition 1.2(b) and the definition of the quantity J at the beginning of that proof. For the initial configuration of (4.2) we find $J = \infty$. Therefore, it must hold $N_b = Z'$ with Z' being defined as in that same proof. Consequently, N_b has the $\mathrm{Geometric}(b)$ distribution. So it is enough to show that there exists a $\mathbf{0}$-feasible process and a stopping time at which the values of the component martingales are $\mathbf{p}$. But Proposition 3.7 shows this is true for every $\mathbf{p}$. □

5 Distributional bounds on downcrossings

In this section we state and prove results about the variability of D_{ab}, the total number of downcrossings over $[a, b]$.

5.1 The large spread setting

Proposition 5.1. *Given $b_0 > 0$, there exists a $\mathbf{0}$-feasible process such that $D_{ab} + 1$ has $\mathrm{Geometric}(\frac{b-a}{1-a})$ distribution, for each $b_0 \leq b < 1$ and $0 < a < b$.*

The corresponding result (cf. Proposition 1.2(b)) holds for $\mathbf{p}$-feasible processes in the limit as $\max_i p_i \to 0$.

Proof. As in the proof of Corollary 1.3(b), we may start with the Geometric(b_0) distribution $\mathbf{p}$ at (4.2) and use the construction in the proof of Proposition 1.2(b). Every time a martingale component reaches b, the other components must be at positions

$$(1-b)\ b_0(1-b_0)^{i-1}, i \geq 1.$$

Similarly, each time the component completes a downcrossing of $[a, b]$ the other components must be at positions

$$(1-a)\ b_0(1-b_0)^{i-1}, i \geq 1.$$

The event that there are no further downcrossings is the event that, after the next time some component reaches b, it then reaches 1 before a, and this has probability $(b-a)/(1-a)$ by (2.1). So

$$\mathbb{P}(D_{ab} = i | D_{ab} \geq i) = (b-a)/(1-a),\ i \geq 1.$$

By the same argument $\mathbb{P}(D_{ab} = 0) = (b-a)/(1-a)$. □

The variance of the $\mathrm{Geometric}(\frac{b-a}{1-a})$ distribution can be written as

$$\left(\tfrac{1-b}{b-a}\right)^2 + \tfrac{1-b}{b-a}. \tag{5.1}$$

It is natural to guess, analogous to Corollary 1.3(c), that this is an upper bound on the variance of D_{ab} in any $\mathbf{0}$-feasible process.

Conjecture 5.2. *For any $\mathbf{0}$-feasible process,*

$$\mathrm{var}(D_{ab}) \leq \left(\tfrac{1-b}{b-a}\right)^2 + \tfrac{1-b}{b-a}.$$

The following result establishes a weaker bound. One can check that in the $a \uparrow b$ limit this bound is first order asymptotic to $(\frac{1-b}{b-a})^2$, which coincides with the first order asymptotics in (5.1).

Proposition 5.3. *For any* $\mathbf{0}$*-feasible process and any* $0 < a < b < 1$*,*

$$\mathrm{var}(D_{ab}) \le \left(\left(\frac{1-b}{b-a} + 2\,\frac{(1-b)^2}{(b-a)^2} + \mu \right)^{1/2} + \mu^{1/2} \right)^2 - \frac{(1-b)^2}{(b-a)^2}$$

where $\mu := \min((2-b)/b^2, 1/a^2)$.

Proof. Fix $0 < a < b < 1$ and consider an arbitrary $\mathbf{0}$-feasible process. Call a particular component martingale at a particular time *active* if it is potentially part of a downcrossing of $[a, b]$. That is, the martingale is initially inactive; it becomes active if and when it first reaches b; it becomes inactive if and when it next reaches a; and so on. So a martingale at x is always active if $x > b$, is always inactive if $x < a$, but may be active or inactive if $a < x < b$.

Given that a particular martingale is currently at x, the mean number of future downcrossing completions equals, by (2.2, 2.3)

$$\tfrac{x(1-b)}{b-a} \text{ if inactive}; \qquad \tfrac{(1-x)b}{b-a} \text{ if active.}$$

Analogously to the proof of Proposition 1.2(c), consider the time S_k at which the k'th downcrossing has been completed. On $\{S_k < \infty\}$,

$$(b-a)\mathbb{E}[D_{ab} - k|\mathcal{F}_{S_k}] = (1-b) \sum_{i\ inactive} M_i(S_k) + b \sum_{i\ active} (1 - M_i(S_k))$$

and because $\sum_i M_i(\cdot) = 1$ this becomes

$$b + (b-a)\mathbb{E}[D_{ab} - k|\mathcal{F}_{S_k}] = \sum_{i\ inactive} M_i(S_k) + \sum_{i\ active} b.$$

The number of active martingales at time S_k is at most $N_b^{(k)} :=$ number of martingales that reached b before time S_k. So the right side cannot be larger than the value taken when $\min(N_b^{(k)}, 1/a)$ active martingales take values just above a and the remaining value of $1 - a\min(N_b^{(k)}, 1/a)$ is distributed among the inactive martingales. This gives the upper bound

$$b + (b-a)\mathbb{E}[D_{ab} - k|\mathcal{F}_{S_k}] \le 1 - a\min(N_b^{(k)}, 1/a) + b\min(N_b^{(k)}, 1/a) \text{ on } \{S_k < \infty\}.$$

The event $\{S_k < \infty\}$ is the event $\{D_{ab} \ge k\}$, so taking expectations and rearranging gives

$$\mathbb{E}[(D_{ab} - k)_+] \le \tfrac{1-b}{b-a}\,\mathbb{P}(D_{ab} \ge k) + \mathbb{E}[\min(N_b^{(k)}, 1/a)\,\mathbf{1}_{\{D_{ab} \ge k\}}].$$

Because $N_b^{(k)} \le N_b$, summing over all $k \ge 1$ gives

$$\tfrac{1}{2}\mathbb{E}[D_{ab}(D_{ab} - 1)] \le \tfrac{1-b}{b-a}\,\mathbb{E}[D_{ab}] + \mathbb{E}[\min(N_b, 1/a)\,D_{ab}]. \tag{5.2}$$

Apply the Cauchy-Schwarz inequality to the second summand on the right side and use $\mathbb{E}[D_{ab}] = \frac{1-b}{b-a}$ to conclude

$$\mathbb{E}[D_{ab}^2] \le \frac{1-b}{b-a} + 2\,\frac{(1-b)^2}{(b-a)^2} + 2\,\mathbb{E}\big[\min(N_b, 1/a)^2\big]^{1/2}\,\mathbb{E}\big[D_{ab}^2\big]^{1/2}. \tag{5.3}$$

Next, for positive constants C_1, C_2 we have the elementary implication

$$\text{if } 0 \le a \le C_1 + 2\,C_2\sqrt{a} \text{ then } \sqrt{a} \le \sqrt{C_1 + C_2^2} + C_2.$$

In our situation, this gives

$$\sqrt{\mathbb{E}[D_{ab}^2]} \le \left(\frac{1-b}{b-a} + 2\,\frac{(1-b)^2}{(b-a)^2} + \mathbb{E}\big[\min(N_b, 1/a)^2\big] \right)^{1/2} + \mathbb{E}\big[\min(N_b, 1/a)^2\big]^{1/2}.$$

Using first Jensen's inequality and then the result (Corollary 1.3(c)) that $\mathrm{var}(N_b) \le (1-b)/b^2$, we see

$$\mathbb{E}[\min(N_b, 1/a)^2] \le \min(\mathbb{E}[N_b^2], 1/a^2) \le \min((2-b)/b^2, 1/a^2)$$

from which the inequality in the proposition readily follows. □

5.2 The small spread setting

Proposition 1.2(a) showed that the spread of N_b could be very small. To see that the case of D_{ab} must be somewhat different, recall that for a martingale component which reaches b, its number of downcrossings has the modified Geometric distribution (2.4) with mean $b(1-b)/(b-a)$. So if we fix b and consider limits in distribution as $a \uparrow b$, we must obtain a limit of the form

$$\frac{b-a}{b(1-b)} D_{ab} \ \to_d \ \sum_{i=1}^{N_b} \xi_i$$

where each ξ_i has Exponential(1) distribution. And although there will be some complicated dependence between $(N_b, \xi_1, \xi_2, \ldots)$ it is clear that the limit cannot be a constant, and therefore in any $\mathbf{p}$-feasible process the variance of D_{ab} as $a \uparrow b$ must grow at least as order $(b-a)^{-2}$. We will not consider that case further here (but see an open problem in section 7), instead turning to the case where a/b is bounded away from 1. Here, in a $\mathbf{0}$-feasible process, $\mathbb{E}[D_{ab}]$ grows as order $1/b$ as $b \downarrow 0$. The next result shows there exist $\mathbf{0}$-feasible processes for which the variance of D_{ab} remains $O(1)$.

The idea behind the construction is to exploit reflection coupling. For instance, starting with $2m$ components at b, a reflection coupling moves the process to a configuration with m components at a and m at $2b-a$ while adding m downcrossings; one can extend this kind of construction to make the process pass through a deterministic sequence of configurations while adding a deterministic number of downcrossings.

Proposition 5.4. *For each $0 \le \alpha < 1$ there exists a constant $C(\alpha) < \infty$ such that: given $0 < a_k < b_k \to 0$ with $a_k/b_k \to \alpha$, there exist $\mathbf{0}$-feasible processes such that*

$$\limsup_k \ \mathrm{var}(D_{a_k, b_k}) \le C(\alpha). \tag{5.4}$$

Proof. Fix k, set $(a, b) = (a_k, b_k)$ and with an abuse of notation write $\alpha = a_k/b_k$. By Proposition 3.7 we may assume we have a $\mathbf{p}_0$-feasible process, where $\mathbf{p}_0$ has finite support and its components are in $(0, b)$.

The proof makes repeated use of the following kind of construction. Specify an interval $[a_0, b_0]$, freeze martingale components initially outside that interval, run the other components as a rescaled Wright-Fisher process and freeze them upon reaching a_0 or b_0 (typically there will be one component ending within (a_0, b_0)). Note this construction has a particular “deterministic" property, that in the final random configuration $(M_i(t), i \ge 1)$ the ranked (decreasing ordered) values $\mathrm{rank}(M_i(t), i \ge 1)$ are non-random, determined by the (ranked) initial values. This holds because $\sum_i M_i(t) = 1$.

The central idea of the proof is the following lemma.

Lemma 5.5. *Write $K = K(\alpha) = 6\lfloor \frac{1}{1-\alpha} \rfloor - 1$. There exists a $\mathbf{p}_0$-feasible process which reaches a configuration $\mathbf{p}_1$ with at most one martingale with value in $(b, 1]$ and at most K martingales taking values in $(0, b]$, having accomplished a deterministic number of downcrossings before that time.*

Proof. We construct the process in stages. At the start of each stage, we consider the first case in the list below which holds, and do the construction specified below for that case. If no case holds then stop; note the property "at most K martingales taking values in $(0, b]$" will then be satisfied.

Case 1. There are at least $1 + \lfloor \frac{1}{1-\alpha} \rfloor$ martingales at b;

Case 2. There are at least $2\lfloor \frac{1}{1-\alpha} \rfloor + 1$ active martingales in (a, b);

Case 3. There are at least $2\lfloor \frac{1}{1-\alpha} \rfloor + 1$ inactive martingales in (a, b);

Case 4. There are at least $\lfloor \frac{1}{1-\alpha} \rfloor$ martingales in $(0, a]$.

Construction in case 1. We let the martingales at b evolve according to the appropriately rescaled Wright-Fisher diffusion, while freezing all other martingales, and then freeze the evolving martingales that reach level a. At least $\lfloor \frac{1}{1-\alpha} \rfloor$ martingales will reach level a, and exactly one will be above b. Once all martingales are frozen, we let those at a evolve as the rescaled Wright-Fisher diffusion until they reach 0 or b. Finally, if initially there were martingales above b, then we let all the martingales above b evolve as the appropriate Wright-Fisher diffusion and freeze those that reach b. This procedure adds a deterministic number of downcrossings (all in the first step), and leaves exactly one martingale above b.

Construction in cases 2 and 3. In case 2 we let the active martingales in (a, b) evolve until they either reach a or b and freeze them at that time. All except one of these martingales reach a or b, so either at least $\lfloor \frac{1}{1-\alpha} \rfloor + 1$ martingales end at b, or at least $\lfloor \frac{1}{1-\alpha} \rfloor$ martingales end at a, adding a deterministic number of downcrossings. So the ending configuration will fit case 1 or case 4. In case 3 we do the same but with the inactive martingales instead; no additional downcrossings are added.

Construction in case 4. We let the martingales in $(0, a]$ evolve until they reach 0 or b and freeze them at that time. At least one of them must reach 0, and no additional downcrossings are added.

The sequence of stages must end, because: in each case 4 stage at least one martingale is stopped at 0, and each case 1 stage creates at least one downcrossing, so there can be only a finite number of such stages; and each case 2 or 3 stage is followed by such a stage.

Moreover each stage is "deterministic", in the previous sense that the ranked configuration at the end of the stage is determined by the ranked configuration at the start, and therefore the ranked configuration $\mathbf{p}_1$ at the termination of the entire construction is non-random, determined by the initial configuration $\mathbf{p}_0$. This implies the total number of downcrossings is deterministic, because the number within each stage is determined by that stage's starting configuration. As already mentioned, $\mathbf{p}_1$ has the property "at most K martingales taking values in $(0, b]$" by the termination condition. The number of martingale components taking values in $(b, 1]$ is at most 1, because each case 1 stage ends that way and the other cases do not allow components to exceed level b. □

In view of Lemma 5.5, to complete the proof of the proposition it suffices to show (5.4) for some $\mathbf{p}_1$-feasible process with $\mathbf{p}_1$ as in Lemma 5.5. In fact we can take an arbitrary such process. The point is that (as noted earlier) the number of downcrossings

D_{ab} has a representation of the form

$$D_{ab} = \sum_{i=1}^{N^*} G_i$$

where N^* is the number of martingale components that hit b, and each G_i has the modified Geometric distribution (2.4). Without any knowledge of the dependence between $(N^*, G_1, G_2, \ldots)$, the fact $N^* \leq K + 1$ implies

$$\mathrm{var}(D_{ab}) \leq \mathbb{E}[D_{ab}^2] \leq (K+1)^2 \mathbb{E}[G_1^2].$$

It is easy to check that $\mathbb{E}[G_1^2]$ is bounded in the limit as $b \to 0$ with $a/b \to \alpha < 1$, and (5.4) follows. □

6 The 0-Wright-Fisher process

Write Δ for the (unranked) infinite simplex $\{(p_i, 1 \leq i < \infty) : \ p_i \geq 0, \sum_i p_i = 1\}$. As mentioned in section 2.2, for each $\mathbf{p} \in \Delta$ there exists the $\mathbf{p}$-Wright-Fisher process, a process with sample paths in $C([0,\infty), \Delta)$ and initial state $\mathbf{p}$, which is the infinite-dimensional diffusion with generator analogous to (2.5) starting from state $\mathbf{p}$, and that this is a $\mathbf{p}$-feasible process. This has a straightforward construction: given $\mathbf{p} \in \Delta$, set $\mathbf{p}^n = (p_1, \ldots, p_{n-1}, \sum_{m \geq n} p_m)$, so the $\mathbf{p}^n$-process exists as a finite-dimensional diffusion. But there is a natural coupling between the $\mathbf{p}^{n-1}$- and the $\mathbf{p}^n$-processes in which the first $n-2$ coordinate processes coincide, and appealing to Kolmogorov consistency for the infinite sequences of processes we immediately obtain the $\mathbf{p}$-process.

Intuitively, we want to think of the $\mathbf{0}$-Wright-Fisher process as a suitable limit of the $(1/n, 1/n, \ldots, 1/n)$-Wright-Fisher processes as $n \to \infty$. But in fact the limit in distribution, in the compactified space $\overline{\Delta} = \{(p_i, 1 \leq i < \infty) : \ p_i \geq 0, \sum_i p_i \leq 1\}$, is the process which is identically $(0,0,0,\ldots)$. The foundational 1981 paper of Ethier and Kurtz [6] shows that a non-trivial limit $\mathbf{X}(t) = (X_i(t), i \geq 1)$ starting from $(0,0,0,\ldots)$ does exist if we work in the ranked infinite simplex ∇; more precisely the limit process has sample paths in $C([0,\infty), \overline{\nabla})$ for the compactifed ranked simplex $\overline{\nabla}$, but for $t > 0$ takes values in ∇.

That process is in some senses the process we want, but that formalization does not suffice for our purposes because it does not keep track of the identities of components as t varies. That is, we want the $\mathbf{0}$-feasible process $\{M_\alpha(t)\}$ for which

$$\mathbf{X}(t) = \mathrm{rank}(\{M_\alpha(t)\}) \text{ with a separate ranking for each } t. \tag{6.1}$$

The component processes $X_i(\cdot)$ are not martingales and we cannot define quantities like N_b and D_{ab} in terms of $\mathbf{X}$. Note that by Lemma 3.1 we cannot represent $\mathbf{X}(t)$ as $\mathrm{rank}(\mathbf{M}(t))$ for any process in $C([0,\infty), \overline{\Delta})$ with martingale components.

Fortunately we can fit the $\mathbf{0}$-Wright-Fisher process into our abstract set-up by combining the existence of the process $\mathbf{X}(t)$ with our Proposition 3.6. Take times $s_k \downarrow 0$ and let μ_k be the distribution of $\mathbf{X}(s_k)$. Then there is a $\oplus$-feasible Wright-Fisher process $\mathbf{M}^k$ with initial distribution μ_k, and existence of the ranked Wright-Fisher process $\mathbf{X}$ implies that consistency condition (iii) of Proposition 3.6 holds with $T_k = s_{k-1} - s_k$, and the conclusion of that proposition is that a $\mathbf{0}$-feasible process satisfying (6.1) exists.

6.1 Distributions associated with the 0-Wright-Fisher process

Problem 6.1. *What are the distributions of N_b and D_{ab} for the $\mathbf{0}$-Wright-Fisher process?*

We remark that, if one only wanted to compute $\mathrm{var}(N_b)$, it would be sufficient to determine the limiting behavior of the quantity

$$\mathbb{P}(\sup_t M_1(t) \geq b,\, \sup_t M_2(t) \geq b |\, M_1(0) = x,\, M_2(0) = y) \tag{6.2}$$

in the limit $x, y \downarrow 0$, where M_1, M_2 are the first two components of a 3-allele Wright-Fisher diffusion. We also note that the quantity (6.2) coincides with the classical solution of the PDE $\frac{1}{2}x(1-x)f_{xx} + \frac{1}{2}y(1-y)f_{yy} - xy\, f_{xy} = 0$ on $[0,b] \times [0,b]$ with the boundary conditions $f(x,0) = f(0,y) = 0$, $f(x,b) = x/b$, $f(b,y) = y/b$, provided that such a solution exists. We were not able to solve the PDE explicitly, so that even the question of finding $\mathrm{var}(N_b)$ is an open problem.

7 Final remarks and open problems

We have already stated open problem 6.1 and Conjecture 5.2. The discussion at the start of section 5.2 concerning constructions where D_{ab} has small spread suggests the following closely analogous question concerning Brownian motions.

Problem 7.1. *For each* $1 \leq i \leq k$ *let* $(B_i(t), 0 \leq t)$ *be standard Brownian motion w.r.t. the same filtration, killed upon first hitting* -1*, and let* L_i *be the total local time of* $B_i(\cdot)$ *at* 0*. How small can the ratio* $\mathrm{var}[\sum_{i=1}^k L_i] / \mathrm{var}[L_1]$ *be?*

We do not know any relevant work, though Jim Pitman (personal communication) observes that for $k = 2$ one can indeed have negative correlation between L_1 and L_2.

Acknowledgments. We thank an anonymous referee for thoughtful comments.

References

[1] David Aldous, *Using prediction market data to illustrate undergraduate probability*, Available at http://www.stat.berkeley.edu/~aldous/Papers/monthly.pdf. To appear in *Amer. Math. Monthly.*, 2012.

[2] David Aldous and Jim Pitman, *Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent*, Probab. Theory Related Fields **118** (2000), no. 4, 455–482. MR-1808372

[3] Jean Bertoin, *Random fragmentation and coagulation processes*, Cambridge Studies in Advanced Mathematics, vol. 102, Cambridge University Press, Cambridge, 2006. MR-2253162

[4] Jean Bertoin, *Exchangeable coalescents*, 2010, Nachdiplom Lectures, ETH Zurich. Available at http://www.fim.math.ethz.ch/lectures/Lectures_Bertoin.pdf.

[5] Peter Donnelly, *Weak convergence to a Markov chain with an entrance boundary: ancestral processes in population genetics*, Ann. Probab. **19** (1991), no. 3, 1102–1117. MR-1112408

[6] S. N. Ethier and Thomas G. Kurtz, *The infinitely-many-neutral-alleles diffusion model*, Adv. in Appl. Probab. **13** (1981), no. 3, 429–452. MR-615945

[7] Kenneth Lange, *Applied probability*, second ed., Springer Texts in Statistics, Springer, New York, 2010. MR-2680838