Metastability between the clicks of the Müller ratchet

Mauro Mariani, Etienne Pardoux, Aurélien Velleret

January 12, 2022

Abstract

We prove the existence and uniqueness of a quasi-stationary distribution for three stochastic processes derived from the model of the Müller ratchet. This model has been originally introduced to quantify the limitations of a purely asexual mode of reproduction in preventing, only through natural selection, the fixation and accumulation of deleterious mutations. The main considered model is clearly non-classical, as it is a stochastic diffusion evolving on an irregular set of infinite dimension with hard killing on an hyperplane. We are nonetheless able to prove exponential convergence in total variation to the quasi-stationary distribution even in this case. The parameters in this last result of convergence are directly related to the core parameters of the Müller ratchet effect. The speed of convergence to the quasi-stationary distribution deduced from the infinite dimensional model extends to the approximations with a large yet finite number of potential mutations. Likewise, we have uniform moment estimates of the empirical distribution of mutations in the population under quasi-stationarity.

1 Introduction

1.1 General presentation

Since deleterious mutations occur much more frequently than beneficial ones, it is crucial to understand how the fixation of these deleterious mutations is regulated. Notably, it is very exceptional that a subsequent mutation reverts a deleterious one, so that only natural selection can maintain some purity in the population. In this respect, there is a major distinction to be made between sexual and asexual reproduction. In a purely asexually reproducing population, a deleterious mutation can only be purged when the lineages carrying it go extinct. In a sexually reproducing population, such a deleterious mutation can be avoided through recombination, without getting rid of the whole set of other mutations carried by the lineages. There is actually no strong evidence that deleterious mutations are specifically targeted during this process of recombination. It appears sufficient that at random some lineages do not carry the mutation any longer and that natural selection comes into play. This aspect of purging deleterious mutations is often cited to explain the success of sexual reproduction (see [16] for more details). Such an advantage for sexual reproduction is to be confronted with the cost (in terms of reproduction efficacy) of requiring two parents. The above scheme for purging deleterious mutations in asexual populations is the main object of study of the current paper.

We plan to justify the existence and uniqueness of a metastable state in which selective effects are able to maintain the population from having an additional deleterious mutation fixed. We call a click such an event of fixation. It has been shown in [1] that clicks happen in finite time a.s. even for the infinite dimensional diffusion model (with infinitely many types of individuals). Rigorous definitions of such metastable state (characterized by the absence of click) can be obtained in a broad generality by a conditioning of stochastic processes. We refer to Subsection 2.1 for the definition of several crucial characteristic of metastability, especially the notion of quasi-stationary distribution (or QSD).

We treat in this paper three models of the Müller ratchet: the first one is discrete both in time and space, the second is a finite dimensional diffusion and the third one is an infinite dimensional diffusion, see Section 1.2 below. We prove the existence and uniqueness of a QSD for those three stochastic representations see resp. Theorems 2.1, 2.2 and 2.3. To our knowledge, the existence and uniqueness of a QSD has not been rigorously proved until now except in the case of a finite state space. This result was nonetheless implicitly exploited for the approximations provided in [17].

We shall see that these QSD are concentrated on distributions with light tails, meaning that the proportion of the population carrying a large number of mutations remains negligible under the QSD. This claim is supported by our Proposition 2.3.2.

We also address the classical issue of specifying the conditions under which metastability is the "most common" observable. A generally accepted answer is to compare the so-called relaxation time t_R, which quantifies the rate at which the dependence in the past conditions vanishes, and the average clicking time t_C of the system. Metastability between clicks would be the most common observation provided $t_R \ll t_C$, so that a sequence of i.i.d. exponential law provides an accurate description of the sequence of intervals between clicks. This is where the comparison with the clicks of a ratchet comes from. If t_R is of the same order as t_C or larger, we a priori can not exclude that trains of short interdependent intervals could alter this observed distribution of interval length. But already if t_R is of the same order as t_C, there shall still be long realizations of inter-click intervals after which we can say that the dependence in the past is forgotten. This discussion is pursued in more details in Subsection 3.

The above mentioned theorems provide a proper definition of these two main quantities. The clicking time is given by the extinction rate of the QSD. On the other hand, the QSD is approached at an exponential rate, from which we derive t_R, by the marginal law of the process conditioned upon the fact that the click has not occurred. We establish these results for the three versions of the model.

As compared to the other models that we have treated by similar techniques as in the current paper, the proof of Theorem 2.3 is particularly difficult. It specifically exploits the effect of selection to obtain practical bounds on the maximal number of accumulated mutations. The argument is technical because at any time an infinitesimal proportion of heavily counter-selected mutants cannot be completely neglect.

A simplified version of such bounds is already needed for the proof of Theorem 2.1. This concerns the process defined in Section 1.2.1 and treated with Theorem 2.1 in Section 2.2. The fact that the process describes a discrete population greatly simplifies the argument. We then extend the justification of the relaxation time and the clicking rate for large population limiting models. Note that the results of [17] or [10] already largely exploit the fact that the population is large. In the diffusive limit defined in Section 1.2.2, an additional difficulty arises in that the diffusion is degenerate on a non-smooth boundary that is partly absorbing and partly repulsive. In order to present a simplified analysis, we introduce a limitation in the number of carried mutations for the statement of Theorem 2.2 given in Section 2.3. In the last step given in Section 2.4 with Theorem 2.3, we establish the existence and uniqueness of a QSD for the more natural infinite dimensional model.

The paper is organized as follows. In the next Section 1.2, we specify the stochastic processes under consideration, first the individual-based model in Section 1.2.1 and then its diffusive limits in Section 1.2.2. Our results of quasi-stationarity are presented in Section 2. Starting in Section 2.1 with the general notion of exponential quasi-stationarity that we aim to establish, we treat resp.

in Sections 2.2, 2.3 and 2.4 each of the three stochastic processes mentioned above. The generic assumptions and theorems on which these proofs rely are stated in Subsection 2.5. Next, we discuss more precisely the interpretation of these results in Section 3. We first justify in Subsection 3.1 to 3.3 under which conditions quasi-stationarity can be observed, then in Subsection 3.4 why we expect it to be frequent in ecology. Finally, we motivate our choice for not introducing a bound on the number of mutations in Section 3.5. The rest of the paper is dedicated to the proofs. Sections 4, 5 and 7 are devoted to the proofs of quasi-stationarity for each of the three processes, while Section 6 is devoted to uniform moment estimates of the QSDs. Such controls of the moments are exploited in Section 7, which makes this ordering natural.

1.2 The mathematical model of the Müller ratchet

1.2.1 The individual-based model as a guideline

For the origin of the models which we study, we refer to the simplified mathematical model which has been proposed by Haigh in [11] to quantify the regulation of deleterious mutations in an asexual population. The interest for this type of simplified models stems from general considerations on the evolutive advantage of recombination, as notably advanced by Müller in 1964 [19]. Since in any finite population, the ultimate fixation of deleterious mutations cannot be avoided (unless by the extinction of the population), the "mechanism" has been called the Müller ratchet.

Assuming a constant deleterious effect of mutations, at each time that the fittest individuals disappear, the ratchet clicks in the sense that the new fittest individuals carry an additional deleterious effect. If the mutation rate is slow enough for these fittest individuals to maintain the stability of the system for a while, we shall rapidly get back to the dynamics before the click, translating the empirical distribution of the number of carried mutations by one. This fixed deleterious mutation is indeed shared by the whole population (present and future). Since the population size is fixed, it does not contribute to natural selection any longer.

This first model with discrete generations and fixed population size N evolves as follows. Mutations that occur are only deleterious and they occur at rate $\lambda > 0$. The cost in fitness of each mutation is quantified by $\alpha \in (0, 1)$. Assume that the current population is distributed with N_i individuals carrying i mutations and consider an individual from the next generation. Each one chooses its parent independently from the others, where the probability that he chooses a specific parent carrying i mutations is:

$$\frac{(1-\alpha)^i}{\sum_{k\geq 0} N_k (1-\alpha)^k}.$$

In addition to the mutations of its parent, each newborn gains ξ deleterious mutations, where ξ is a Poisson random variable with mean λ, specific to the newborn. ξ is drawn independently for each newborn and of the choice of the parents.

Remark 1.2.1. *Of course, the situation is more intricate in reality. Mutations certainly do not have constant effect, and combination effects are frequent (i.e. epistasis). In many asexual populations, there is evidence of the role of horizontal gene transfers, for instance with plasmids ([13], [18], [20], [24]), which can be seen as a weak form of recombination. Moreover, the fact that mutations are deleterious is due to a change in the physiology that may be compensated by other means. It might even happen that after subsequent mutations, the carriers of an initially deleterious mutation become more adapted than the wild types [23]. Neglecting these effects enables however to gain insight on the main regulatory factor.*

1.2.2 The stochastic diffusion under consideration

In the following, we also consider a description of the model that corresponds to a limit of large population size, accelerated time-scale (for which time is continuous), thus also small selective effects and small mutation rate. In the following statements, $d \in \mathbb{N} \cup \{\infty\}$ defines an upper-bound on the number of deleterious mutations that can be carried by an individual. If $d := \infty$ in the following expression, $i \in [\![0, d]\!]$ has to be understood as $i \in \mathbb{Z}_+$.

We are interested in the following Fleming-Viot system of SDEs for the $X_i(t)$'s, $i \in [\![0, d]\!]$, where $X_i(t)$ denotes the proportion of individuals in the population who carry exactly i deleterious mutations at time t (with $X_{-1} \equiv 0$): $\forall\, i \leq d$,

$$\begin{aligned} dX_i(t) &= \alpha(M_1(t) - i)\, X_i(t)\, dt + \lambda(X_{i-1}(t) - \mathbf{1}_{\{i<d\}} X_i(t))\, dt \\ &\quad + \sqrt{X_i(t)}\, dW_t^i - X_i(t)\, dW_t \qquad\qquad (1.1){:}\,(S^{(d)}) \\ &\text{where } W_t := \textstyle\sum_{j=0}^d \int_0^t \sqrt{X_j(s)} dW_s^j, \quad M_1(t) := \sum_{i=0}^d i\, X_i(t), \end{aligned}$$

with $(W^i)_{i\geq 0}$ a family of mutually independent Brownian Motions.

This process has been introduced in [10] and it has been shown in [1] that clicks occur a.s. in finite time. In [22], a closely related process with compensatory mutations is considered. We refer to this article for a detailed presentation of the connection to related individual-based models and only sketch next the interpretation of the parameters.

The selective effect of the deleterious mutations is the term proportional to α in the drift term. Since the population size is fixed, the growth rate of the individuals is shifted to be 0 on average over the population. As we assume that all deleterious mutations carry the same burden, the growth rate of individuals carrying i mutations is proportional to the difference between i and the average number of mutations, i.e. $M_1(t)$. The appearance of new mutations is modeled by the term proportional to λ in the drift term. λ corresponds to the rate at which individuals carrying i mutations give birth to individuals carrying $i + 1$ mutations. Finally, the neutral choice of the individuals replaced at each birth events give rise to the martingale term. Our time-scale corresponds to the rescaling of time $t \mapsto t'/N_e$, where N_e is the "effective population size".

Remark 1.2.2. *This notion of "effective population size" has been largely considered to extend the properties of unstructured homogeneous individual-based models to individual-based models with a population structure that differentiates the individuals. So it is meant to be applied to real populations under ecological study. Notably, it provides the scaling of the genealogies that makes it approximate the standard Kingman's coalescent [14]. Thus, it gives for any sample in the population an estimate on the time at which their most recent common ancestor lived. It is of course natural that this quantity plays a role in such modeling of heritable factors.*

Remark 1.2.3. *For practical reasons, the current formulation of the martingales is different from the one in [1] in the aftermath of [10]. One can easily check however that the brackets of these martingale parts coincide, so that the models are actually the same.*

2 Exponential quasi-stationarity results

2.1 Exponential quasi-stationarity

The conclusions of the following theorems are expressed in terms of the notion of exponential quasi-stationarity that we borrow from [26]. For $\mathcal{X}$ a generic (Polish) space, hereafter $B(\mathcal{X})$ is the space of bounded measurable functions on $\mathcal{X}$ and $\mathcal{M}_1(\mathcal{X})$ the space of Borel probability measures.

Definition 1. *For any linear and bounded semi-group $(P_t)_{t\geq 0}$ acting on $\mathcal{M}_1(\mathcal{X})$, we say that P displays a uniform exponential quasi-stationary convergence (abbreviated as QSC) with characteristics $(\nu, h, \rho_0) \in \mathcal{M}_1(\mathcal{X}) \times B(\mathcal{X}) \times \mathbb{R}$ if $\langle \nu \,|\, h \rangle = 1$ and there exists $C, \gamma > 0$ such that for any $t > 0$ and for any measure $\mu \in \mathcal{M}(\mathcal{X})$ with $\|\mu\|_{TV} \leq 1$:*

$$\left\| e^{\rho_0 t} \mu P_t - \langle \mu \,|\, h \rangle\, \nu \right\|_{TV} \leq C e^{-\gamma t}. \tag{2.1}$$

As stated in Corollary 2.2.4 [26], this implies the following convergence result to ν.

Corollary 2.1.1. *Assume* (2.1)*. Then for any* $t \geq 0$ *and* $\mu \in \mathcal{M}_1(\mathcal{X})$ *such that* $\langle \mu \,|\, h \rangle > 0$:

$$\| \, \mathbb{P}_\mu \left(X_t \in dx \mid t < \tau_\partial \right) - \nu(dx) \, \|_{TV} \leq C \frac{\|\mu - \nu\|_{TV}}{\langle \mu \,|\, h \rangle} \, e^{-\gamma \, t}. \tag{2.2}$$

Remark 2.1.2. *Choosing* $\mu = \nu$ *in* (2.1)*, it is not hard to deduce the following relation:*

$$\forall \, t \geq 0, \quad \nu P_t = e^{-\rho_0 \, t} \, \nu, \qquad \text{and in particular } \mathbb{P}_\nu(t < \tau_\partial) = e^{-\rho_0 \, t}, \tag{2.3}$$

cf Fact 2.2.2. in [26]. This relation is what characterizes ν *as a QSD since it implies that for any* $t \geq 0$, $\mathbb{P}_\nu \left(X_t \in dx \mid t < \tau_\partial \right) = \nu(dx)$*. By restricting the convergence stated in* (2.1) *on the evaluation of the measure on* $\mathcal{X}$*, we obtain a similar characterization of* h*. This latter convergence is what makes us call* h *the survival capacity.*

There is an additional related notion that will be useful to describe the behavior of the process with the requirement of a long inter-click interval. This process is generically defined through the survival capacity h, on the state space: $\mathcal{H} := \{x \in \mathcal{X} \; ; \; h(x) > 0\}$.

Definition 2. *We say that the Q-process exists if there exists a family* $(\mathbb{Q}_x)_{x \in \mathcal{H}}$ *of probability measures on* Ω *defined by:*

$$\lim_{t \to \infty} \mathbb{P}_x(\Lambda_s \,|\, t < \tau_\partial) = \mathbb{Q}_x(\Lambda_s) \tag{2.4}$$

for any $\mathcal{F}_s$*-measurable set* Λ_s*. We also implicitly assume that the process* $(\Omega; (\mathcal{F}_t)_{t \geq 0}; (X_t)_{t \geq 0}; (\mathbb{Q}_x)_{x \in \mathcal{H}})$ *is a homogeneous strong Markov process.*

Remark 2.1.3. *The transition kernel of the Q-process is given by:*

$$q(x; t; dy) = e^{\rho_0 \, t} \, \frac{h(y)}{h(x)} \, p(x; t; dy), \tag{2.5}$$

where $p(x; t; dy)$ *is the transition kernel of the Markov process* (X) *under* $(\mathbb{P}_x)$*. Note that* $\mathcal{X} \setminus \mathcal{H}$ *is generally avoided by the process* X *under* $\mathbb{Q}_x$*. In the examples of the current article,* h *is actually positive while* ν *is unique as a QSD. No distinction has then to be made between* $\mathcal{X}$ *and* $\mathcal{H}$ *regarding the Q-process.*

Thanks to Corollary 2.2.8 of [26], our justification for the proof of QSC actually implies related results of convergence for the Q-process. Notably $\beta(dx) := h(x) \, \nu(dx)$ *is the unique invariant probability measure of this process.*

2.2 The discrete population case

Let $N \geq 1$ be the number of individuals in the population, and $D_n(t)$ for $n \leq N$ and $t \in \mathbb{Z}_+$ be the number of mutations carried by the n-th individual. We consider the empirical measure at time $t > 0$ defined as follow:

$$\mathcal{Z}_t^N := (1/N) \textstyle\sum_{n \leq N} \delta_{D_n(t)}, \tag{2.6}$$

so that $\mathcal{Z}_t^N(i) \in N^{-1} \times [\![0, N]\!]$ specifies the proportion of individuals with exactly i mutations (since everything is discrete, we identify $\mathcal{Z}_t^N$ as a function from $\mathbb{Z}_+$ to $\mathbb{R}_+$). From the rules describing the next generation from the previous one, see Subsection 1.2.1, $\mathcal{Z}^N$ is clearly a Markov process evolving on $\mathcal{M}_1^N(\mathbb{Z}_+)$, where:

$$\begin{aligned} \mathcal{M}_1^N(\mathbb{Z}_+) &:= \{(1/N) \textstyle\sum_{i \leq N} \delta_{d_i} \; ; \; d_i \in \mathbb{Z}_+ \, , \, \textstyle\sum_{i \in \mathbb{Z}_+} d_i = N\} \\ &\equiv \{z : \mathbb{Z}_+ \mapsto (1/N) \times [\![0, N]\!] \; ; \; \sum_{i \in \mathbb{Z}_+} z(i) = 1\}. \end{aligned} \tag{2.7}$$

The clicking time under consideration comes from the extinction of the fittest individuals, i.e.:

$$\tau_\partial := \inf\{t \geq 0 \ ; \ \mathcal{Z}_t^N(0) = 0\} = \inf\{t \geq 0 \ ; \ \mathcal{Z}_t^N \notin \mathcal{M}_1^{(0),N}(\mathbb{Z}_+)\} \tag{2.8}$$

$$\text{where } \mathcal{M}_1^{(0),N}(\mathbb{Z}_+) = \{z \in \mathcal{M}_1^N(\mathbb{Z}_+), z(0) \geq 1/N\}. \tag{2.9}$$

Remark 2.2.1. *Classical theory on quasi-stationarity can be exploited by interpreting the clicking time τ_∂ as an extinction time. We implicitly rely on the process $\bar{\mathcal{Z}}_t^N := \mathcal{Z}_t^N \mathbf{1}_{\{t<\tau_\partial\}} + \partial \mathbf{1}_{\{\tau_\partial \leq t\}}$ which is clearly Markov and lives on $\mathcal{M}_1^{(0),N}(\mathbb{Z}_+) \cup \{\partial\}$. For the process $\bar{\mathcal{Z}}^N$, τ_∂ is the hitting time of the absorbing state ∂ (the cemetery).*

Our main conclusion for this process is the following theorem:

Theorem 2.1. *Consider for any N the Markov process Z^N whose transitions are prescribed as in Section 1.2.1, with clicking time τ_∂. Then, its semigroup P displays QSC with characteristics $(\nu, h, \lambda) \in \mathcal{M}_1(\mathcal{M}_1^{(0),N}(\mathbb{Z}_+)) \times B(\mathcal{M}_1^{(0),N}(\mathbb{Z}_+)) \times \mathbb{R}_+^*$.*

Moreover, h is bounded, admits a positive lower-bound, while the Q-process exists on $\mathcal{H} = \mathcal{M}_1^{(0),N}(\mathbb{Z}_+)$. The convergence to ν given in (2.2) is uniform. In particular, ν is the unique QSD.

Remark 2.2.2. *The proof of this theorem strongly relies on the criteria given in the proofs in [5] and generalized in Section 2.1 of [25], notably in order to exploit the property that lineages carrying many mutations tend to rapidly go extinct. It provides an elementary understanding of how the criteria of persistence (A3) can be deduced (cf Subsection 2.5.1). To handle both of these aspects, the discreteness of the process is however strongly involved in the proof, which makes the estimation poorly quantitative for large N.*

2.3 The finite dimensional case

In this section, we denote by τ_∂ the clicking time of the process $X^{(d)}$ solution of the system (1.1) : $(S^{(d)})$, that is:

$$\tau_\partial^d := \inf\{t \geq 0 \ ; \ X_0^{(d)}(t) = 0\}.$$

The system of SDE then evolves for finite d on:

$$\mathcal{X}_d := \{(x_k)_{k\in[\![0,d]\!]} \in [0,1]^{d+1} \ ; \ \textstyle\sum_{k=0}^d x_k = 1\}.$$

Our main conclusion for this process is the following theorem:

Theorem 2.2. *Consider the system of SDEs (1.1) : $(S^{(d)})$ for any $d \in \mathbb{N}$, with clicking time τ_∂. Then, its semigroup P displays a QSC with characteristics $(\nu_d, h_d, \lambda_d) \in \mathcal{M}_1(\mathcal{X}_d) \times B(\mathcal{X}_d) \times \mathbb{R}_+^*$.*

Moreover, h_d is bounded and the associated Q-process exists on $\mathcal{H} = \mathcal{X}_d$. In addition, for any $y_0 \in (0,1)$, h_d is lower-bounded by a positive constant on $\{x \in \mathcal{X}_d \ ; \ x_0 \geq y_0\}$. In particular, ν_d is the unique QSD.

Remark 2.3.1. *The proof of this theorem applies quite directly the ideas that we have previously exploited in [25]. As in [6], we rely mainly on the Harnack inequality. Nonetheless, we have here to be cautious in the way we handle jointly the absorbing and repulsive boundary conditions.*

Moreover, we prove the following controls on the moments of the QSDs ν_d, for $d \in \mathbb{N}$:

Proposition 2.3.2. *For any $k \geq 1$, we have uniform tightness in d over the moments of order k of the unique QSDs ν_d associated to the solution of (1.1) : $(S^{(d)})$, i.e.:*

$$\sup_{d\in\mathbb{N}} \int_{\mathcal{X}_d} \nu_d(dx) \mathbf{1}_{\{M_k(x) \geq m_k\}} \to 0 \text{ as } m_k \to \infty \quad \text{where } M_k(x) := \textstyle\sum_{i\in[\![0,d]\!]} i^k \, x_i.$$

In particular, the sequence $\hat{\nu}_d$, where the values for the coordinates larger than $k+1$ are put to 0, is tight in $\mathcal{M}_1(\mathbb{R}_+^{\mathbb{Z}})$.

Remark 2.3.3. *Thanks to the above theorem, we expect the sequence* (ν_d) *to converge as* $d \to \infty$ *to the unique QSD* ν_∞ *for the infinite system (for which the control extends). This control on the moments is actually crucial for the proof of uniqueness when* $d = \infty$.

2.4 The infinite dimensional case

We consider now the infinite dimensional case, for which we require the existence of moments. Let

$$\mathcal{X}^\eta := \{(x_k)_{k\in\mathbb{Z}_+} \in [0,1]^\infty \; ; \; \sum_{k=0}^{\infty} x_k = 1 \, , \, \sum_{k=0}^{\infty} k^\eta x_k < \infty\}$$

Thanks to Theorem 3 in [1], we know that for any initial condition x that belongs to $\mathcal{X}^\eta$, for some $\eta > 2$, $(S^{(\infty)})$ has a unique weak solution which is a. s. continuous with values in $\mathcal{X}^\eta$. Our main conclusion for this process is the following theorem:

Theorem 2.3. *Consider the system of SDEs* $(S^{(\infty)})$*, i.e. with* $d = \infty$*, defined on* $\mathcal{X}^6$ *with clicking time* τ_∂*. Then, its semigroup* P *displays a QSC with characteristics* $(\nu_\infty, h_\infty, \lambda_\infty) \in \mathcal{M}_1(\mathcal{X}^6) \times B(\mathcal{X}^6) \times \mathbb{R}_+^*$.

Moreover, h_∞ *is bounded and the Q-process exists on* $\mathcal{H} = \mathcal{X}^6$*. In addition, for any* $\epsilon \in (0,1)$*,* h_∞ *is lower-bounded by a positive constant on* $\left\{x \in \mathcal{X}^6 \; ; \; x_0 \geq \epsilon\right\}$*. In particular,* ν_∞ *is the unique QSD. Besides, there exist* $C, \gamma, d_\wedge > 0$ *such that , the convergences stated in* (2.1) *and* (2.2) *hold true with these constants for the processes given both by* (1.1)$:(S^{(d)})$ *on* $\mathcal{X}^6$ *and on* $\mathcal{X}_d$ *for any* $d \geq d_\wedge$.

Remark 2.4.1. *In practice, we shall need to control moments of order* η' *strictly larger than 2 while exploiting the finiteness of moments of order* $2\eta'$*. For simplicity, we thus restrict ourselves to* $\mathcal{X}^6$ *although our proof could certainly be adapted provided* $\eta > 4$*. The core of our proof is based on the intuition that the slower the decay of the tail the more rapidly it gets erased and renewed. So we do not expect large tails to play a significant role.*

2.5 Some crucial sets of conditions ensuring exponential quasi-stationarity

The proof of QSC are exploiting the criteria given both in Section 2.1 of [25] and in Subsection 2.2.1 of [26]. In particular, in [26] the methods and statements of [25] have actually been adjusted with the current paper in mind. Assumptions (**A**) and (**AF**) are exploited for our general Theorem 2.4 as stated in Subsection 2.5.2.

2.5.1 Basic assumptions

$(A0_S)$: **Specification on the state space** There exists a sequence $(\mathcal{D}_\ell)_{\ell\geq 1}$ of closed subsets of $\mathcal{X}$ such that (with $int(\mathcal{D})$ the interior of $\mathcal{D}$):

$$\forall \, \ell \geq 1, \; \mathcal{D}_\ell \subset int(\mathcal{D}_{\ell+1}). \tag{A0}$$

Remark 2.5.1. *Originally in [25], it is also assumed that* $\bigcup_{\ell\geq 1}\mathcal{D}_\ell = \mathcal{X}$*. Due to the fact that we shall deal here with reflecting boundaries, it will however be more convenient for our proofs that none among the sets* $(\mathcal{D}_\ell)$ *includes them.*

The sequence $\mathcal{D}_\ell$ will serve as a reference in the following, and we also denote:

$$\mathbf{D} := \{\mathcal{D} \; ; \; \mathcal{D} \text{ is closed and } \exists \, \ell \geq 1, \; \mathcal{D} \subset \mathcal{D}_\ell\} \, . \tag{2.10}$$

For the next statements, we shall exploit the following notations for the exit and entry times of any set $\mathcal{D}$:

$$T_\mathcal{D} := \inf\{t \geq 0 \; ; \; X_t \notin \mathcal{D}\} \, , \qquad \tau_\mathcal{D} := \inf\{t \geq 0 \; ; \; X_t \in \mathcal{D}\} \, .$$

$(A1)$: **Mixing property** There exists a probability measure $\zeta \in \mathcal{M}_1(\mathcal{X})$ such that, for any $\ell \geq 1$, there exists $L > \ell$ and $c, t > 0$ such that:

$$\forall\, x \in \mathcal{D}_\ell, \qquad \mathbb{P}_x\left(X_t \in dx\ ;\ t < \tau_\partial \wedge T_{\mathcal{D}_L}\right) \geq c\, \zeta(dx).$$

$(A2)$: **Escape from the Transitory domain** *There exist $\rho > 0$ and $E \in \mathbf{D}$:*

$$\sup_{\{x \in \mathcal{X}\}} \mathbb{E}_x\left(\exp\left[\rho\left(\tau_\partial \wedge \tau_E\right)\right]\right) < \infty.$$

ρ in the previous exponential moment is required to be strictly larger than the following **"survival estimate"**:

$$\rho_S := \sup\left\{\gamma \geq 0 \,\middle|\, \sup_{L \geq 1} \liminf_{t > 0}\ e^{\gamma t}\, \mathbb{P}_\zeta(t < \tau_\partial \wedge T_{\mathcal{D}_L}) = 0\right\} \vee 0. \tag{2.11}$$

Remark 2.5.2. *It is proved in Theorem 2.3 of [26] that ρ_S coincide with the extinction rate ρ_0 provided the semi-group displays QSC.*

The next two assumptions are proposed as alternatives and each alternative will be exploited in the current paper. The former is the assumption first introduced in Section 2.1 of [25]. The latter provides a way to ensure the former given $(A0-2)$ as proved in [26].

$(A3)$: **"Asymptotic comparison of survival"** There exists $E \in \mathbf{D}$ and $\zeta \in \mathcal{M}_1(\mathcal{X})$:

$$\limsup_{t \to \infty} \sup_{x \in E} \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_\zeta(t < \tau_\partial)} < \infty.$$

$(A3_F)$: **"Absorption with failures"**
Given $\zeta \in \mathcal{M}_1(\mathcal{X})$, $\rho > \rho_S$ and $E \in \mathbf{D}$, for any $\epsilon \in (0,1)$, there exist $t_\vee, c > 0$ such that for any $x \in E$ there exists a stopping time U_A such that:

$$\{\tau_\partial \wedge t_\vee \leq U_A\} = \{U_A = \infty\} \qquad \text{and} \qquad \mathbb{P}_x(U_A = \infty,\, t < \tau_\partial) \leq \epsilon \exp(-\rho\, t_\vee),$$

while for a certain stopping time V_A:

$$\mathbb{P}_x\big(X(U_A) \in dx'\ ;\ U_A < \tau_\partial\big) \leq c\, \mathbb{P}_\zeta\big(X(V_A) \in dx'\ ;\ V_A < \tau_\partial\big).$$

We further require that there exists a stopping time U_A^∞ extending U_A in the following sense:

- *$U_A^\infty := U_A$ on the event $\left\{\tau_\partial \wedge U_A < \tau_E^1\right\}$, where $\tau_E^1 := \inf\{s \geq t_\vee\ ;\ X_s \in E\}$.*
- *On the event $\left\{\tau_E^1 \leq \tau_\partial \wedge U_A\right\}$ and conditionally on $\mathcal{F}_{\tau_E^1}$, the law of $U_A^\infty - \tau_E^1$ coincides with the law of $\widetilde{U}_A^\infty$ for a realization $\widetilde{X}$ of the Markov process $(X_t, t \geq 0)$ with initial condition $\widetilde{X}_0 := X(\tau_E^1)$ and conditionally independent of X, given $X(\tau_E^1)$.*

2.5.2 Theorems

Slightly adapting [25] (regarding $(A0_S)$), we say that Assumption $(\mathbf{A})$ holds, whenever:

"$(A1)$ holds for a certain $\zeta \in \mathcal{M}_1(\mathcal{X})$ and a sequence $(\mathcal{D}_\ell)$ that satisfies $(A0_S)$. Moreover, there exist $E \in \mathbf{D}$ such that (A2), holds with a certain $\rho > \rho_S$ as well as $(A3)$."

As in [26], we say that Assumption $(\mathbf{AF})$ holds, whenever:

"$(A1)$ holds for a certain $\zeta \in \mathcal{M}_1(\mathcal{X})$ and a sequence $(\mathcal{D}_\ell)$ satisfying $(A0_S)$. Moreover, there exist $\rho > \rho_S$ and $E \in \mathbf{D}$ such that assumptions $(A2)$, and $(A3_F)$ hold."

Theorem 2.2 in [26] can be restated for our purpose as:

Theorem 2.4. *Assume that either* (**A**) *or* (**AF**) *holds. Then, the semigroup P displays QSC with characteristics* $(\nu, h, \rho_0) \in \mathcal{M}_1(\mathcal{X}) \times B(\mathcal{X}) \times \mathbb{R}$ *and the Q-process exists on* $\mathcal{H}$.

Since the exploited sequence $(\mathcal{D}_\ell)_{\ell \geq 1}$ usually does not cover the whole state space, we shall exploit Proposition 2.2.5 of [26] to deduce lower-bounds of h. The next proposition recalls its statement.

Proposition 2.5.3. *Assume that* (**AF**) *or* (**A**) *holds. Then, the survival capacity h is uniformly lower-bounded on any set $H \subset \mathcal{X}$ that satisfies the following condition:*
(H_0) : *there exists* $t > 0$, $\ell \geq 1$ *such that* $\inf_{\{x \in H\}} \mathbb{P}_x(\tau_{\mathcal{D}_\ell} \leq t \wedge \tau_\partial) > 0$.
It implies the following identification :

$$\mathcal{H} := \{x \in \mathcal{X} \; ; \; h(x) > 0\} = \{x \in \mathcal{X} \; ; \; \exists \ell \geq 1, \, \mathbb{P}_x(\tau_{\mathcal{D}_\ell} < \tau_\partial) > 0\}.$$

Thanks to Proposition 2.5.3, we shall prove in our models that h is actually positive, which proves the uniqueness of the QSD thanks to Corollary 2.1.1.

3 Outlook

3.1 Interpretation of crucial parameters

For the last process, no parameter other than α and λ is introduced. We deduce from Theorem 2.3 that the QSD and the survival capacity depend only on α and λ, as well as the values $C, \gamma > 0$ in (2.1) and (2.2).

As already noted by Haigh in [11], α/λ is the average number of deleterious mutations that are established in the deterministic limit (neglecting neutral fluctuations). The deterministic distribution of mutations is a function of α/λ, and actually follows a Poisson distribution with this mean, as shown in [10]. To infer the level of fluctuations around this deterministic equilibrium, we shall look at the coefficient in front of the martingale term in a new time-scale such that the mutation rate is set at 1. This gives $1/\lambda$, which we recall to scale as $\sqrt{1/N_e}$, where N_e is the effective population size mentioned in Remark 1.2.2. This term actually quantifies the relatedness in the population of uniformly sampled individuals. A large population size thus corresponds to letting λ go to infinity, making the deviations away from the deterministic distribution more rare.

A natural scale for the time between clicks can be easily derived from this notion of QSC, with the definition $t_C := \rho_0^{-1}$. On the other hand, we can propose the following definition for the relaxation time:

$$t_R := \inf\{t_r > 0 \; ; \; \exists C > 0, \, \forall \mu, \quad \|\mu A_t - \nu\|_{TV} \leq (C/\langle \mu \,|\, h \rangle) \times e^{-t/t_r}\} \quad \leq 1/\zeta. \tag{3.12}$$

Our results justify the validity of this definition. We can deduce as in [25] that the convergence to h and β also occurs at quicker rate than $1/t_R$.

By relying on the arguments of Theorem 2.3 and Proposition 2.3.2, we expect that truncating the number of accumulated mutations is not likely to alter much this value of t_R provided the threshold is sufficiently large. Since we cannot evaluate t_R precisely and are only able to provide an upper-bound, this is still conjectural. But substantial increase of these last components are proved to be rare thanks to Proposition 2.3.2 and not so significant when we look at Section 7.7.

Remark 3.1.1. *The dependence on the initial condition in* (3.12) *is expected from the linearity of the semi-group* (P_t)*, as observed in [26]. More general dependencies could nonetheless be imagined, relying for instance on Lyapunov functions as in [7] or in [2]. We simply do not think it would change the value of* t_R *because the confinement is mainly due to extinction and immediate repulsion from the boundaries.*

3.2 Previous estimations

The study of this quasi-stationary regime arises naturally when one wishes to estimate the rate at which the ratchet clicks. To obtain quantitative estimates, several authors have justified their approach by assuming that the typical clicking time t_C is much larger than the typical relaxation time t_R of the system, usually with an empirical reference for the latter ([10], [17]).

In [17], an estimation of t_C in the context where $t_R \ll t_C$ is obtained through the characteristic equation of a certain QSD $\nu_\star$, of the form $\mathcal{L}\nu_\star = -\lambda\nu_\star$, with $\mathcal{L}$ is a certain infinitesimal generator and λ its main eigenvalue . This QSD $\nu_\star$ that they study is not the general QSD ν_∞ that we describe. The detailed description of the latter is reasonably argued to be too intricate. The former is in fact derived from a one-dimensional approximation of the process under metastability. It is argued that in the context of large populations, and given the number of fittest individuals, one can approximate the rest of the distribution as an almost deterministic profile. The dependence in this number of fittest individuals only occurs in the normalizing factor of this distribution. This latter argument of concentration could probably be made rigorous by using Large Deviation theory. Such results are beyond the scope of the present paper.

Note that the validity of this approximation relies upon the fact that $t_R \ll t_C$, where t_R is to be related to the QSD ν_∞. The relaxation rate of $\nu_\star$ is only a partial indication, although presumably carrying most of the information.

3.3 The quasi-stationary regime is generally observed for $t_R \ll t_C$

Provided $t_R \ll t_C$, we clearly expect to generally observe the quasi-stationary regime between clicks.

It is classical that with the QSD as an initial condition, the extinction time and extinction state are independent, the former being exponentially distributed, as it has been established in Theorem 2.6 of [8]. Assuming that we start the analysis at a new click after a long time-interval without click, it implies that the profile of mutations just after the click is distributed as the restriction of the QSD to the hyperplane $\{X_0 = 0\}$.

Since having large values of M_1 makes it actually harder for the process to reach the hyperplane, we expect that, under the QSD restricted to $\{X_0 = 0\}$, M_1 tends to be smaller than the prediction $1 + \lambda/\alpha$ derived from the deterministic limit (under the constraint that $x_0 = 0$). Besides, the fittest individuals are altered by first changing into the type with only one mutation. So we expect also that under the QSD restricted to $\{X_0 = 0\}$, there is an over-representation of the proportion of individuals carrying a single mutation (the new optimal trait). Thus, we expect the distribution just after the click to be less prone to a future click than would be the QSD itself. Since $t_R \ll t_C$, the quasi-stationary regime is then rapidly reached.

Let us also imagine a dramatic situation where some clicks would rapidly follow each others. Then, it would imply that these fittest classes of individuals are rapidly wiped off, while not letting much time for the others to change much. Since we have seen that we have very strong controls of moments under the QSD, cf notably Proposition 2.3.2, such succession of clicks cannot hold for long. A class that is not prone to a quick extinction should be reached quite early and generate a new quasi-stationary regime. Such dramatic situations are thus expected to be very isolated and of limited impact, while of course very rare.

Expecting an exponential law for the inter-click intervals and the independence between them should be in conclusion fairly accurate provided $t_R \ll t_C$.

As we discuss in Subsection 2.3 of [27], one can also conclude whether or not the QSD profile is likely to be observed without conditioning by comparing ν to the survival capacity h. If quasi-stationarity is stable, we do not expect that the conditioning on having a click in the far future shall substantially alter the dynamics. In most trajectories, the Q-process shall thus behave as the original process. So h should be mostly constant on the support of $\beta(dx) = h(x)\,\nu(dx)$, implying $h \approx 1$ where the density of ν is large. Yet, the QSD and the survival capacity are certainly quite difficult to specify with simulations because they live on a large dimensional space. Likewise, the convergence in total variation exploited in (3.12) is probably not very practical for numerical estimation.

3.4 Such a quasi-stationary regime is favored by natural selection

The fact that such a quasi-stationary regime can be defined does not directly imply that this state is likely to be observed: if $t_C \ll t_R$, the next click would happen too quickly after the previous one for the dependence in the transition to be lost.

In the context of a rapid succession of clicks, the population would be likely to get extinct quite early on as compared to populations able to reach a metastable regime between each click (notably by having a lower mutation rate). Provided that the process in the metastable regime ensures the maintenance of an optimal sub-population of large size, the time between clicks can get much longer than a simple scaling by the mutation rate would suggest. Indeed, the click would then be the consequence of an exceptional deviation of the process away from the metastable attractor. It thus provides opportunity for additional beneficial mutations to fix and outcompete the fixation of deleterious mutations. A larger population size is then favoring both the selection of beneficial mutations and the prevention of deleterious fixations, in a positive feedback loop.

The second scenario with metastability is thus expected to be the more likely for stable asexual populations. Although the first scenario cannot be excluded for destabilized populations or too small ones, the interest in this metastable regime is thus biologically motivated by its benefice in term of survival.

3.5 Motivation for an unbounded number of deleterious mutations

In order to prove quasi-stationarity results, the case where $d < \infty$ can be treated more easily and provides an introduction to the case $d = \infty$. Nonetheless, the arguments for having a convergence at a given rate becomes more and more artificial as d tends to infinity. The constant involved in the Harnack inequalities go toes zero as the dimension increases. By considering the case $d = \infty$, we actually handle as a whole the case where d is sufficiently large. By these means, we are able to prove that the rate of convergence can be upper-bounded by a quantity that does not depend on the specific value of d. This is to be expected since, even when a large number of deleterious mutations is permitted, we expect individuals carrying a large number of mutations to remain negligible.

Referring for instance to [10], it is not difficult to prove that in the deterministic limit of a large population, the empirical measure of the number of mutations in the population tends to a Poisson distribution. The tail of the distribution is quickly disappearing. This deterministic limit corresponds to a limiting time-change of equation (1.1):$(S^{(d)})$ of the form $t' = t/\epsilon$ with $\alpha = \alpha'/\epsilon$, $\lambda = \lambda'/\epsilon$ as ϵ tends to 0. The Poisson distribution has a mean of $\lambda'/\alpha' = \lambda/\alpha$ so that it may be possible to quantify much more precisely than we do the threshold in the number of deleterious mutations after which differentiating individuals is not so crucial. This could make it possible to obtain quantitative bounds from our arguments in the context of very large populations (in the vicinity of the deterministic limit).

4 Proof of Theorem 2.1

The proof of Theorem 2.1 relies on the criteria presented in Section 2.1 of [25]. The two following propositions provides the first two steps in this proof.

Proposition 4.0.1. *For any $N \geq 1$, $\alpha \geq 0$, $\lambda > 0$ and $z \in \mathcal{M}_1^{(0),N}(\mathbb{Z}_+)$:*

$$\inf_{z_0 \in \mathcal{M}_1^{(0),N}(\mathbb{Z}_+)} \mathbb{P}_{z_0}(Z^N(1) = z) > 0.$$

Proposition 4.0.2. *For any $N \geq 1$, $\alpha, \lambda > 0$ and $\epsilon > 0$, there exists $K \geq 1$ such that:*

$$\forall\, z \in \mathcal{M}_1^{(0),N}(\mathbb{Z}_+), \quad \mathbb{P}_z(Z^N(1) \notin E) \leq \epsilon,$$
$$\textit{where } E := \{z \in \mathcal{M}_1^{(0),N}(\mathbb{Z}_+)\ ;\ z([\![K, \infty[\![) = 0\}$$

Remark 4.0.3. *The convergence is uniform in this case. By exploiting Theorem 2.1 of [25], we thus implicitly deduce the two criteria presented in [5] for the proof of a uniform convergence.*

Step 1: proof of Proposition 4.0.1 We simply impose that all the individuals of the next generation are the offspring of an individual without any mutation, and prescribe the number of mutations that they get from the profile of z. We obtain a positive lower-bound uniform over any $z_0 \in \mathcal{M}_1^{(0),N}(\mathbb{Z}_+)$ by noticing that the probability of choosing a fittest individual as a parent is: $z_0(0)/(\sum_{i\geq 0} z_0(i) \times (1-\alpha)^i)$, which is necessarily larger than $1/N$. The number of mutations is then chosen independently of z_0, and there is indeed a positive probability that the sequence of independent Poisson distributed random variables has an empirical law distributed as z. This concludes the proof of Proposition 4.0.1. □

Step 2: proof of Proposition 4.0.2 We first prove that with a high probability, the sub-population of individuals carrying a large number of mutations leave no progeny. Let $K \geq 1$ for the threshold in the number of mutations. The probability that an individual chooses a parent with more than K mutations is upper-bounded by $N \times (1-\alpha)^K$, since $z(0) \geq 1$. For any $\epsilon > 0$, there exists indeed $K \geq 1$ such that, with a probability greater than $1 - \epsilon/2$, no individual in the next generation descends from an individual with more than K mutations. Likewise, there exists $K' \geq 1$ such that, with a probability greater than $1 - \epsilon/2$, the number of additional mutations is less than K' (for any individual, independently of the initial condition z). We deduce that:

$$\forall\, z \in \mathcal{M}_1^{(0),N}(\mathbb{Z}_+), \quad \mathbb{P}_z(Z^N(1) \notin E) \leq \epsilon,$$
$$\text{where } E := \{z \in \mathcal{M}_1^{(0),N}(\mathbb{Z}_+)\ ;\ z([\![K + K', \infty[\![) = 0\}$$

which concludes the proof of Proposition 4.0.2. □

Step 3: proof of Theorem 2.1 The choice of the sequence $\mathcal{D}_\ell$ is here degenerate, since we can simply set $\mathcal{D}_\ell$ as the whole space $\mathcal{M}_1^{(0),N}(\mathbb{Z}_+)$ for any ℓ. Note that $(A0_S)$ is satisfied even for this degenerate case. Actually, the exit time are simply infinite and the entry times in $\mathcal{D}_\ell$ always equal zero. We see that Proposition 4.0.1 clearly implies Assumption $(A1)$ of [25]. Next we prove $(A2)$, namely that for any $\rho > 0$, there exists E such that, with τ_E its first hitting time:

$$\sup_z \mathbb{E}_z(\exp[\rho(\tau_\partial \wedge \tau_E)]) < \infty.$$

This is easily deduced thanks to Proposition 4.0.2 through the Markov property and an induction over $k \geq 1$ to have a proper upper-bound on $\mathbb{P}_z(k < \tau_\partial \wedge \tau_E)$. Then, for the last criterion $(A3)$, we remark that E as defined in Proposition 4.0.2 is finite. By Proposition 4.0.1 and the Markov property, we deduce that there exists $c > 0$ such that for any $t \geq 1$:

$$\mathbb{P}_{\delta_0}(t < \tau_\partial) \geq c \sup_{z \in E} \mathbb{P}_z(t - 1 < \tau_\partial) \geq c \sup_{z \in E} \mathbb{P}_z(t < \tau_\partial).$$

This concludes $(A3)$ and that Assumption $(\mathbf{A})$ is satisfied. Theorem 2.1 is then deduced from Theorem 2.4. □

Remark 4.0.4. *If we were to replace the law of ξ by a Bernoulli distribution (mutations occurring one by one), Proposition 4.0.1 would still hold with the restriction of $z = \delta_0$, which is the only case we need. It would extend to any z provided we change the time 1 by the maximal number of mutations in z. The proof would not be much more difficult with overlapping generations, except that individuals should then be removed one by one. The proof of the equivalent of Proposition 4.0.2 would merely be slightly more difficult. The details are left to the interested reader.*

5 Proof of Theorem 2.2

5.1 Main properties leading to the proof

The proof of Theorem 2.2 relies (roughly) on the criteria stated in Susbection 2.2.1 of [26], with here a non-uniform convergence. In the following, d is fixed and we drop the notations recalling it. Let:

$$\mathcal{D}_\ell := \left\{ x = (x_i)_{0\le i\le d} \in \left(\frac{1}{2\ell d}\,,\, 1 - \frac{1}{2\ell}\right)^d \ ;\ \textstyle\sum_{0\le i\le d} x_i = 1 \right\}, \tag{5.1}$$

which have non-empty interiors. The three following propositions state that X with extinction time τ_∂ satisfies Assumption (**A**), as we show in Section 5.2. The proofs of each of them is deferred to further subsections. We recall that $T_{\mathcal{D}_\ell}$ denotes the exit time out of $\mathcal{D}_\ell$.

Proposition 5.1.1. *For any $t > 0$, there exists $\zeta \in \mathcal{M}_1(\mathcal{X}_d)$ with support in $\mathcal{D}_2$ such that for any $\ell \ge 1$, there exists $c > 0$ such that:*

$$\forall\, x \in \mathcal{D}_\ell, \quad \mathbb{E}_x\left(X_t \in dy\ ;\ t < \tau_\partial \wedge T_{\mathcal{D}_{\ell+1}}\right) \ge c\, \zeta(dy).$$

Proposition 5.1.2. *For any $\ell \ge 1$:*

$$\limsup_{t\to\infty} \sup_{x,x'\in\mathcal{D}_\ell} \frac{\mathbb{P}_x\left(t < \tau_\partial\right)}{\mathbb{P}_{x'}\left(t < \tau_\partial\right)} < \infty.$$

Proposition 5.1.3. *For any $\rho > 0$, there exists $\ell \ge 1$ such that:*

$$\sup_{x\in\mathcal{X}_d} \mathbb{E}_x \exp[\rho\,(\tau_{\mathcal{D}_\ell} \wedge \tau_\partial)] \le 16.$$

The lower-bound of the survival capacity is derived from the following lemma:

Lemma 5.1.4. *For any $y_0 > 0$, the set $H := \{x \in \mathcal{X}_d\ ;\ x_0 \ge y_0\}$ satisfies (H_0) as stated in Proposition 2.5.3.*

Since its proof is elementary, it is deferred to Subsection 5.3 just after we show that Theorem 2.2 is implied by the four above statements.

Remark 5.1.5. *The above results hold for any $d \in \mathbb{N}$ and the constant could depend dramatically on d, except for Lemma 5.1.4. The choice of 16 In Proposition 5.1.3 is arbitrary and suffices to our purpose.*

5.2 Proof of Theorem 2.2 with these propositions

For this proof, we plan to exploit Theorem 2.4 and first ensure Assumption (**A**) (cf Subsection 2.5.2). It is easily seen from their definition in (5.1) that the sets $\mathcal{D}_\ell$ satisfy $(A0_S)$.

Propositions 5.1.1, 5.1.2 and 5.1.3 ensure respectively $(A1)$, $(A3)$ and $(A2)$. Notably, for $(A3)$, since ζ has support in $\mathcal{D}_2$, for any $\ell \geq 2$, thanks to Proposition 5.1.2:

$$\limsup_{t\to\infty} \sup_{x\in\mathcal{D}_\ell} \frac{\mathbb{P}_x (t < \tau_\partial)}{\mathbb{P}_\zeta (t < \tau_\partial)} < \infty.$$

From Theorem 2.4, we thus deduce that the semi-group displays QSC. The survival capacity is actually positive, thanks to Lemma 5.1.4 combined with Proposition 2.5.3. This implies with (2.2) that ν is in fact the unique QSD because any QSD ν' must then satisfy both $\langle \nu' \,|\, h \rangle > 0$ and $\nu' A_t = \nu'$ for any t. This concludes the proof of Theorem 2.2. $\square$

5.3 Proof of Lemma 5.1.4

Let any $y_0 > 0$ and define $H := \{x \in \mathcal{X}_d \,;\, x_0 \geq y_0\}$. X_0 under $\mathbb{P}_x^{(d)}$ is lower-bounded by the solution Y (independent of d) to:

$$dY_s = -\lambda\, ds + \sqrt{Y_s\,(1 - Y_s)} dB_0(s)\,, \qquad Y(0) = y_0.$$

We consider the extinction time for Y: $\tau_\partial^Y := \inf\{s \,;\, Y_s = 0\}$ and an arbitrary time $t = 1$. We deduce the following uniform lower-bound:

$$\inf_{\{d, x\in H\}} \mathbb{P}_x^{(d)}(t < \tau_\partial) \geq c_S := \mathbb{P}_{y_0}(t < \tau_\partial^Y) > 0.$$

By the Markov inequality and Proposition 5.1.3, there exists ℓ such that for any d and $x \in \mathcal{X}_d$:

$$\mathbb{P}_x^{(d)}(t < \tau_{\mathcal{D}_\ell} \wedge \tau_\partial) \leq c_S/2.$$

This clearly implies (H_0) for H in the sense that for any $x \in H$:

$$\mathbb{P}_x^{(d)}(\tau_{\mathcal{D}_\ell} \leq t \wedge \tau_\partial) \geq \mathbb{P}_x^{(d)}(t < \tau_\partial) - \mathbb{P}_x^{(d)}(t < \tau_{\mathcal{D}_\ell} \wedge \tau_\partial) \geq c_S/2. \qquad \square$$

5.4 Harnack inequalities for Propositions 5.1.1 and 5.1.2

The proofs of Propositions 5.1.1 and 5.1.2 are actually similar to those in Subsection 4.2.2 of [25]. They exploit the Harnack inequality –the following assumption (H)– classically deduced for such an elliptic diffusion.

In the following, we say that a process (Y_t) on $\mathcal{Y}$ with generator $\mathcal{L}$ (including possibly an extinction rate ρ_c) satisfies Assumption (H) if:

For any compact sets $K, K' \subset \mathcal{Y}$ with C^2 boundaries such that $K \subset int(K')$, $0 < t_1 < t_2$ and positive C^∞ constraints: $u_{\partial K'} : (\{0\} \times K') \cup ([0, t_2] \times \partial K') \to \mathbb{R}_+$, the unique positive solution $u(t, x)$ to the Cauchy problem:

$$\begin{aligned} &\partial_t u(t, x) = \mathcal{L}u(t, x) \text{ on } [0, t_2] \times K'\,; \\ &u(t, x) = u_{\partial K'}(x) \text{ on } (\{0\} \times K') \cup ([0, t_2] \times \partial K'), \end{aligned}$$

satisfies, for a certain $C = C(t_1, t_2, K, K') > 0$ independent of $u_{\partial K'}$:

$$\inf_{x\in K} u(t_2, x) \geq C \sup_{x\in K} u(t_1, x).$$

On any $\mathcal{D}_n$, $\sigma^{(d)}$ is uniformly elliptic while $\sigma^{(d)}$ and $b^{(d)}$ are uniformly Lipschitz. Similarly as we show in Section 4.2.2 of [25], Assumption (H) holds for the generator $\mathcal{L}^{(d)}$ of any finite dimensional process $X^{(d)}$, while restricted on a certain set $\mathcal{D}_n$.

5.4.1 Proof of Proposition 5.1.1

We apply assumption (H) to $u(t,y) := \mathbb{E}_y\left(f(Y_t)\ ;\ t < \tau_\partial^{n+1}\right)$, where f is any non-negative C^∞ function with support in $\mathcal{D}_n = K$, and $\tau_\partial^{n+1} := \inf\{t \geq 0\ ;\ X_t^{(d)} \notin \mathcal{D}_{n+1}\}$. It implies that for any $y \in \mathcal{D}_n$, $y_0 \in \mathcal{D}_1$ and $0 < t_0 < t$:

$$\mathbb{E}_y\left(f(Y_t)\ ;\ t < \tau_\partial^{n+1}\right) \geq C_n\, \mathbb{E}_{y_0}\left(f(Y_{t_0})\ ;\ t_0 < \tau_\partial^2\right).$$

Since $\inf_{y_0 \in \mathcal{D}_1} \mathbb{P}_{y_0}\left(Y_t \in \mathcal{D}_1\ ;\ t < \tau_\partial^2\right) > 0$ (by choosing arbitrary y_0 and $t_0 = t/2$), we can obtain a probability measure ζ with support on $\mathcal{D}_2$, independent of n, such that (since C_n does not depend on f):

$$\forall\, y \in \mathcal{D}_n, \quad \mathbb{E}_y\left(Y_t \in dy\ ;\ t < \tau_\partial^{n+1}\right) \geq c_n\, \zeta(dy). \qquad \square$$

5.4.2 Proof of Proposition 5.1.2

The proof is similar but more technical because the reference measure is now in the upper-bound, so that we can no longer neglect trajectories exiting $\mathcal{D}_{n+1}$. We can choose $t_1 := 1$ and find two compact sets $K, K' \subset \mathcal{Y}$ with C^2 boundaries such that $\mathcal{D}_n \subset K \subset int(K') \subset int(\mathcal{D}_{n+1})$. We want to approximate the function:

$$u(t,y) := \mathbb{E}_y\left(f(Y_t)\ ;\ t < \tau_\partial\right), \quad \text{with } t \geq t_1,\ y \in K'$$

defined for some non-negative $f \in \mathbb{C}^\infty(\mathcal{Y})$. Referring to Theorem 5.1.15 in [15], we can prove that u is continuous. It is clearly non-negative. However, it is a priori not regular enough to apply the Harnack inequality directly. Thus, we approximate it on the parabolic boundary $[t_1, \infty) \times \partial K' \bigcup \{t_1\} \times K'$ by a certain family $(U_k)_{k \geq 1}$ of non-negative smooth –$\mathcal{C}_+^\infty$ w.l.o.g.– functions. We then deduce approximations of u in $[t_1, \infty) \times K'$ by (smooth) solutions of:

$$\begin{aligned} \partial_t u_k(t,y) - \mathcal{L} u_k(t,y) = 0, &\qquad t \geq t_1,\ y \in int(K') \\ u_k(t,y) = U_k(t,y), \quad t \geq t_1,\ y \in \partial K', &\qquad \text{or } t = t_1,\ y \in K'. \end{aligned}$$

By Assumption (H), the constant involved in the Harnack inequality does not depend on the values on the boundary. Thus, it applies with the same constant for the whole family of approximations u_k. With $t_2 := 2$ and $t_3 := 3$, we deduce that there exists $C_n > 0$ such that for any k and any $y, y' \in \mathcal{D}_n$:

$$u_k(t_2, y) \leq C_n u_k(t_3, y'),$$

where the constant C_n does not depend on f either. We refer to the proof in [6], Section 4, step 4, to state that such an Harnack inequality extends to the approximated function u, with the convergence of U_k on the parabolic boundary. It means that for any non-negative $f \in \mathbb{C}^\infty(\mathcal{Y})$:

$$\forall\, y, y' \in \mathcal{D}_n,\ \mathbb{E}_y\left(f(Y_{t_2})\ ;\ t_2 < \tau_\partial\right) \leq C_n\, \mathbb{E}_{y'}\left(f(Y_{t_3})\ ;\ t_3 < \tau_\partial\right)$$

It thus extends to any measurable and bounded f. We know fix $t \geq t_2$ and apply this result to the function $f(y) := \mathbb{P}_y(t - t_2 < \tau_\partial)$, together with the Markov property:

$$\forall\, y, y' \in \mathcal{D}_n,\ \forall\, t \geq t_2, \quad \mathbb{P}_y\left(t < \tau_\partial\right) \leq C_n\, \mathbb{P}_{y'}\left(t + t_3 - t_2 < \tau_\partial\right) \leq C_n\, \mathbb{P}_{y'}\left(t < \tau_\partial\right).$$

This concludes the proof of Proposition 5.1.2. $\square$

5.5 Proof of Proposition 5.1.3

Proposition 5.1.3 is proved by recursively ensuring that the k-th first coordinates have escaped from the value 0. For any $y > 0$, and $0 \leq k \leq d$, let:

$$T_y^k := \inf\{s \geq 0 \ ; \ \forall\, j \leq k, \ X_j(s) \geq y\}.$$

The proof of Proposition 5.1.3 is achieved with two following lemmas as its first steps.

Lemma 5.5.1. *For any $d \in \mathbb{N}$ and $t > 0$:*

$$\sup\left\{\mathbb{P}_x\left(t < \tau_\partial\right) \middle| x \in \mathcal{X}_d, x_0 \leq y_0\right\} \to 0 \ \textit{as} \ y_0 \to 0.$$

Lemma 5.5.2. *For any $y \in (0,1), \epsilon, t_\vee > 0$ and $1 \leq k \leq d$, there exists $y' \in (0,1)$, $t \in (0, t_\vee)$ such that:*

$$\inf\left\{\mathbb{P}_x\left(T_{y'}^k < t \wedge \tau_\partial\right) \middle| x \in \mathcal{X}_d, \forall\, j \leq k-1, \ x_j \geq y\right\} \geq 1 - \epsilon.$$

Lemma 5.5.1 is a direct consequence of the fact that X_0 is upper-bounded by the solution Y of:

$$dY_t = \alpha\, d\, Y_t dt + \sqrt{Y_t(1-Y_t)} dB_0(t) \ , \ Y_0 = y_0,$$

for which it is known that 0 is an absorbing value.

Step 1: proof of Lemma 5.5.2. Given $k, \epsilon, y, t_\vee$, choose $t \in (0, t_\vee)$ sufficiently small such that:

$$\inf\left\{\mathbb{P}_x\left(t < U_{y/2}^{k-1}\right) \middle| x \in \mathcal{X}_d, \forall\, j \leq k-1, \ x_j \geq y\right\} \geq 1 - \epsilon/2,$$
$$\text{where } U_{y/2}^{k-1} := \inf\{s \geq 0 \ ; \ \exists\, j \leq k-1, \ X_j(s) \leq y/2\}.$$

To ensure roughly the uniformity in such x, we can simply lower-bound X_j for $j \leq k-1$ by solutions to the equation:

$$dY_s^j = -(\lambda + \alpha j)\, dt + \sqrt{Y_t(1-Y_t)} dB(t) \ , \ Y_0^j = y,$$

and choose t such that Y^j stays above $y/2$ on the time-interval $[0,t]$ with probability greater than $1 - \epsilon/(2k)$.

Let $y_1 := \lambda\, y/(4\lambda + 4\alpha\, k)$. Then, for any $s \leq U_{y/2}^{k-1} \wedge T_{y_1}^k$:

$$(\alpha M_1(s) - \alpha k - \lambda)\, X_k(s) + \lambda\, X_{k-1}(s) \geq \lambda\, y/2 - (\alpha k + \lambda)\, y_1 \geq \lambda\, y/4,$$

so that X_k is lower-bounded by the solution Y_k of:

$$dY_k(s) = \lambda\, y/4\, dt + \sqrt{Y_t(1-Y_t)} dB(t) \ , \ Y_k(0) = 0.$$

Since 0 is an entrance boundary for this process, cf e.g. Subsection 3.3.3 in [12], there exists $0 < y' \leq y_1 \wedge (y/2)$ such that:

$$\mathbb{P}(\sup_{\{s \leq t\}} Y_k(s) < y') \leq \epsilon/2.$$

On the event $\left\{\sup_{s \leq t} Y_k(s) \geq y'\right\} \cap \left\{t < U_{y/2}^{k-1}\right\}$, which occurs with probability greater than $1 - \epsilon$, the condition $T_{y'}^k < t \wedge \tau_\partial$ is satisfied. This ends the proof of Lemma 5.5.2. $\square$

Step 2: concluding the proof of Proposition 5.1.3. Given $\rho > 0$, let $t_0 := \log(2)/\rho$. We can choose thanks to Lemma 5.5.1 a certain $y_0 \in (0,1)$ such that for any $x = (x_i)$ satisfying $x_0 \leq y_0$, it holds:

$$\mathbb{P}_x\,(t_0 < \tau_\partial) \leq \exp[-\rho\, t_0]/2 = 1/4.$$

By the Markov property and an induction, for any $k \geq 1$:

$$\begin{aligned} \mathbb{P}_x(k\, t_0 \leq \tau_\partial \wedge T^0_{y_0}) &\leq \mathbb{P}_x(k\, t_0 \leq \tau_\partial \;,\; (k-1)\, t_0 \leq T^0_{y_0}) \leq 1/4^k, \\ \sup_x \mathbb{E}_x \exp[\rho\,(T^0_{y_0} \wedge \tau_\partial)] &\leq \textstyle\sum_{k\geq 0} e^{\rho t_0 [k+1]}\, \mathbb{P}_x(k\, t_0 \leq \tau_\partial \wedge T^0_{y_0}) \\ &\leq \textstyle\sum_{k\geq 0} 2^{k+1}/4^k = 4 < \infty. \end{aligned}$$

By the Markov property, Lemma 5.5.2 and by induction on $0 \leq k \leq d$, there exists y_k such that, on the event $\{T^0_{y_0} \leq \tau_\partial\}$ and with $\epsilon = 1/16d$:

$$\mathbb{P}_x - \text{a.s.} \quad \mathbb{P}_x\left(T^k_{y_k} \leq (T^0_{y_0} + k\, t_0/d) \wedge \tau_\partial \,|\mathcal{F}_{T^0_{y_0}}\right) \geq 1 - k/(16d). \tag{5.2}$$

To deduce this induction, we set $t_\veebar := t_0/d$ when we apply Lemma 5.5.2 for $1 \leq k \leq d$.

Then, for some large value of $t > 0$, let $V_t := \tau_\partial \wedge T^d_{y_d} \wedge t$, and consider

$$E_t := \sup_{\{x\}} \mathbb{E}_x \exp[\rho V_t] < \infty.$$

For any x such that $x_0 \geq y_0$ (so that $T^0_{y_0} = 0$ $\mathbb{P}_x$-a.s.), we deduce from the Markov property, with $\widetilde{V}_t$ defined as V_t for the Markov process $\widetilde{X}$ starting at time 0 from X_{t_0}:

$$\begin{aligned} \mathbb{E}_x \exp[\rho V_t] &\leq e^{\rho t_0}\,(1 + \mathbb{E}_x[\mathbb{E}_{X(t_0)} \exp[\rho \widetilde{V}_t] \;;\; t_0 < V_t]) \\ &\leq 2\,(1 + E_t \times [1 - \mathbb{P}_x(T^d_{y_d} \leq t_0 \wedge \tau_\partial)]) \\ &\leq 2 + E_t/8, \end{aligned}$$

where we exploited inequality (5.2). On the other hand, for any general x:

$$\begin{aligned} \mathbb{E}_x \exp[\rho V_t] &\leq \mathbb{E}_x\left(\exp[\rho(T^0_{y_0} \wedge \tau_\partial)]; V_t \leq T^0_{y_0}\right) + \mathbb{E}_x\left(\exp[\rho T^0_{y_0}]\, \mathbb{E}_{X(T^0_{y_0})} \exp[\rho \widetilde{V}_t] \;;\; T^0_{y_0} < V_t\right) \\ &\leq (2 + E_t/8) \times \mathbb{E}_x\left(\exp[\rho(T^0_{y_0} \wedge \tau_\partial)]\right) \\ &\leq 8 + E_t/2, \end{aligned}$$

where we exploited the previous estimate with the fact that $X_0(T^0_{y_0}) \geq y_0$. Taking the supremum over x, and since $E_t < \infty$, we deduce: $E_t \leq 16$.

The limit where $t \to \infty$ ensures $\sup_x \mathbb{E}_x \exp[\rho(T^d_{y_d} \wedge \tau_\partial)] \leq 16$. This concludes the proof of Proposition 5.1.3. □

6 Proof of Proposition 2.3.2

The proof of Proposition 2.3.2 relies on the two following propositions, handled uniformly over d. The first states that descent from large values of the moment quickly occurs with probability close to one; the second states that a too large increase of the moment is unlikely to occur.

Proposition 6.0.1. *For any $k, t, \epsilon, m > 0$, there exists $m' > 0$ such that for any $d \in \mathbb{N}$ and initial condition $x \in \mathcal{X}_d$ such that $M_k(x) \leq m$:*

$$\mathbb{P}_x\left(\sup_{s\leq t} M_k(X_s) \geq m'\right) \leq \epsilon.$$

Proposition 6.0.2. *For any $t > 0$ and $k \geq 1$, with $T_k(m) := \inf\{t \geq 0\ ;\ M_k(X_t) \leq\ m\}$:*

$$\sup\left\{\mathbb{P}_x\left(t < \tau_\partial \wedge T_k(m)\right) \middle| d \in \mathbb{N}, x \in \mathcal{X}_d\right\} \underset{m\to\infty}{\longrightarrow} 0.$$

We shall first prove Proposition 2.3.2 thanks to these two propositions, then Proposition 6.0.1 and finally prove Proposition 6.0.2. This last proof relies on 3 steps of descent, the last one being iterated for each moment between 2 and k. The main result for each of these steps is given by the three following lemmas, whose proofs are deferred at the end of Subsection 6:

Lemma 6.0.3. *For any $t > 0$:*

$$\sup\left\{\mathbb{P}_x\left(t < \tau_\partial\right) \middle| d \in \mathbb{N}, x \in \mathcal{X}_d, M_1(x) \geq 1, x_0\, M_1(x) \leq \delta\right\} \to 0 \text{ as } \delta \to 0.$$

Lemma 6.0.4. *Given any $t, y_0 > 0$:*

$$\sup\left\{\mathbb{P}^{(d)}_x\left(t \leq T_1(m_1) \wedge \tau_\partial\right) \middle| d \in \mathbb{N}, x \in \mathcal{X}_d, x_0 \geq y_0\right\} \underset{m_1\to\infty}{\longrightarrow} 0.$$

Lemma 6.0.5. *Given any $k \geq 1$ and t, $m_k > 0$:*

$$\sup\left\{\mathbb{P}^{(d)}_x\left(t \leq T_{k+1}(m_{k+1}) \wedge \tau_\partial\right) \middle| d \in \mathbb{N}, x \in \mathcal{X}_d, M_k(x) \leq m_k\right\} \underset{m_{k+1}\to\infty}{\longrightarrow} 0.$$

6.1 Proof of Proposition 2.3.2

First of all, we show that we have a uniform upper-bound on the extinction rate $\rho^{(d)}_0$ associated to the system (1.1):$(S^{(d)})$: $\sup_{d\in\mathbb{N}} \rho^{(d)}_0 < \infty$. Indeed, whatever $d \in \mathbb{N}$, we can find $x \in \mathcal{X}_d$ such that $x_0 \geq 1/2$ so that X_0 under $\mathbb{P}^{(d)}_x$ is lower-bounded by the solution Y to:

$$dY_s = -\lambda\, ds + \sqrt{Y_s\ (1 - Y_s)}dB_0(s)\ , \qquad Y(0) = 1/2.$$

Note that the boundary $y = 1$ is entrance for this process so that it exits $(0, 1)$ only through 0, cf e.g. Subsection 3.3.3 in [12]. The semi-group governing Y, with extinction at τ^Y_0, corresponds exactly to the system (1.1):$(S^{(d)})$ with $d = 1$, $\alpha = 0$, $X'_0 = Y$ and $X'_1 = 1 - Y$. We know from Theorem 2.2 that the semigroup displays QSC with extinction rate $\rho_\vee$. Denoting $\mathbb{P}^Y_{1/2}$ the law of Y, we deduce from the convergences of the survival capacities:

$$\rho^{(d)}_0 = \lim_{t\to\infty} \tfrac{-1}{t}\, \log \mathbb{P}^{(d)}_{x^{(d)}}(t < \tau_\partial) \leq \lim_{t\to\infty} \tfrac{-1}{t}\, \log \mathbb{P}^Y_{1/2}(t < \tau^Y_0) := \rho_\vee.$$

Thanks to Proposition 6.0.2, we can choose m such that $T_k(m)$ satisfies:

$$\sup_{d\geq 1} \sup_{x\in\mathcal{X}_d} \mathbb{E}^{(d)}_x \exp[(\rho_\vee + 1)\, (T_k(m) \wedge \tau_\partial)] \leq C < \infty.$$

In particular, it implies that for any $t > 0$ and $d \in \mathbb{N}$:

$$\mathbb{P}^{(d)}_{\nu_d}(t < T_k(m) \wedge \tau_\partial) \leq C\, \exp[-(\rho_\vee + 1)\, t]. \tag{6.1}$$

Then, for any $\epsilon > 0$, consider $t := -\log(\epsilon/(2\, C))$. Thanks to Proposition 6.0.1, we can choose a certain $m' > 0$ such that for any initial condition x such that $M_k(x) \leq M$:

$$\mathbb{P}^{(d)}_x\left(\sup_{s\leq t} M_k(X_s) \geq m'\right) \leq \epsilon/2\, \exp[-\rho_\vee\, t]. \tag{6.2}$$

Thus, for any $d \in \mathbb{N}$: (with the QSD ν_d)

$$\begin{aligned}
\nu_d(\{M_k \geq m'\}) &= \exp[\rho_0^{(d)}\, t]\, \mathbb{P}^{(d)}_{\nu_d}(M_k(X_t) \geq m'\,;\, t \leq \tau_\partial) \\
&\leq \exp[\rho_\vee\, t] \Big(\mathbb{P}^{(d)}_{\nu_d}(T_k(m) > t\,;\, t < \tau_\partial) \\
&\qquad + \mathbb{E}^{(d)}_{\nu_d}\Big[\mathbb{P}^{(d)}_{X(T_k(m))}\Big(\sup_{s\leq t} M_k(X_s) \geq m'\Big)\,;\, T_k(m) < t \wedge \tau_\partial \Big] \Big) \\
&\leq C \times \epsilon/(2\,C) + \epsilon/2 \leq \epsilon,
\end{aligned}$$

by inequalities (6.1), (6.2) and the definition of t. This concludes the proof of Proposition 2.3.2, given Propositions 6.0.1 and 6.0.2 whose proofs follow. □

Remark 6.1.1. *We deduce thanks to Lemma 6.0.5 that we are in fact able to bound any moment with large probability under the QSD ν_d uniformly on $d \in \mathbb{N}$. These of course will extend to the limiting QSD on $\mathcal{X}^6$.*

6.2 Proof of Proposition 6.0.1

Let $k, t > 0$,$m' \geq m$, $d \in \mathbb{N}$ and $x \in \mathcal{X}_d$ such that $M_k(x) \leq m$ be fixed.

We consider the semi-martingale decomposition of M_k:

$$dM_k(t) = V_k(t)\, dt + d\mathcal{M}_k(t), \tag{6.3}$$

where $\mathcal{M}_k$ is a continuous martingale starting from 0, whose quadratic variation is

$$\langle \mathcal{M}_k \rangle_t = \int_0^t (M_{2\,k}(s) - M_k(s)^2)\, ds,$$

and V_k is a bounded variation process defined as:

$$V_k := \alpha(M_1 \times M_k - M_{k+1}) + \lambda \sum_{\ell=0}^{d-1} (\ell+1)^k X_\ell - \lambda(M_k - d^k X_d). \tag{6.4}$$

Thanks to the Hölder inequality, considering a random variable Y such that $Y = j$ with probability X_j:

$$M_1 = \mathbb{E}(Y) \leq \mathbb{E}(Y^{k+1})^{1/(k+1)}\ ;\ M_k = \mathbb{E}(Y^k) \leq \mathbb{E}(Y^{k+1})^{k/(1+k)} \text{ thus } M_1 \times M_k \leq M_{k+1}.$$

Exploiting also that $(\ell+1)^k \leq 2^k \times \ell^k$ for $\ell \geq 1$, we deduce, with $C = C(k) = \lambda(2^k - 1)$:

$$V_k \leq C\, M_k + \lambda. \tag{6.5}$$

To obtain an upper-bound on the probability that $\sup_{s \leq t} M_k(s)$ is large, we want to exploit the Doob inequality on a non-negative sub-martingale that is an upper-bound of M_k. It leads us to consider the solution of the following equation:

$$\widehat{M}_k(t) := m + \lambda t + C \textstyle\int_0^t \widehat{M}_k(s)\, ds + \mathcal{M}_k(t), \tag{6.6}$$

because classical results of comparison imply that for any $t \geq 0$, $\widehat{M}_k(t) \geq M_k(t)$, see for instance Proposition 3.12 in [21]. The fact that $\widehat{M}_k$ is non-negative comes from the fact that M_k is non-negative. As a solution to equation (6.6), $\widehat{M}_k$ is clearly a sub-martingale. Since it is upper-bounded by d^k, we can also apply the Gromwall Lemma to deduce that for any initial condition x such that $M_k(x) \leq m$:

$$\sup_{\{s\leq t\}} \mathbb{E}_x^{(d)} \left[\widehat{M}_k(s)\right] \leq (m + \lambda t)\, e^{Ct}. \tag{6.7}$$

By exploiting Doob's inequality on $\widehat{M}_k$, then inequality (6.7) with $C_M := (1+\lambda t)\, e^{Ct}$, we obtain:

$$\begin{aligned}\mathbb{P}_x^{(d)}(\sup_{\{s\leq t\}} M_k(s) > m') &\leq \mathbb{P}_x^{(d)}(\sup_{\{s\leq t\}} \widehat{M}_k^{(k)}(s) > m') \\ &\leq \frac{\mathbb{E}_x^{(d)}[\widehat{M}_k(t)]}{m'} \leq \frac{C_M\, m}{m'}.\end{aligned}$$

This concludes the proof of Proposition 6.0.1. □

6.3 Proof of Proposition 6.0.2

Let $t, \epsilon > 0$. From Lemma 6.0.3, we can choose certain $\delta > 0$ and $m_1^\vee \geq 2\,\lambda/\alpha$ such that:

$$\sup \left\{\mathbb{P}_x^{(d)}\,(t < \tau_\partial)\,\middle|\, d \in \mathbb{N}, x \in \mathcal{X}_d, M_1(x) \geq m_1^\vee, x_0\, M_1(x) \leq \delta\right\} \leq \epsilon. \tag{6.8}$$

Recall $T_1(m_1) := \inf\{t \geq 0\ ;\ M_1(X_t) \leq m_1\} \leq T_1(m_1^\vee)$ for any $m_1 \geq m_1^\vee$. The value of m_1 will be fixed in (6.12), but we first need to prove that with a probability close to 1, X_0 has escaped from the boundary $x_0 = 0$ provided that M_1 has not been small. Let:

$$T_B(\delta) := \inf\{t \geq 0\ ;\ X_0(t)\, M_1(X_t) \leq \delta\}, \tag{6.9}$$

By the Markov property, we deduce from (6.8) that for any $x \in \mathcal{X}_d$ and $d \geq 1$:

$$\mathbb{P}_x^{(d)}\,(T_B(\delta) \leq t \leq T_1(m_1^\vee)\,,\ 2\,t < \tau_\partial) \leq \epsilon. \tag{6.10}$$

Recalling $m_1^\vee \geq 2\,\lambda/\alpha$, $t \leq T_B(\delta) \wedge T_1(m_1^\vee) \wedge \tau_\partial$ implies that for any $s \leq t$: $(\alpha M_1(s) - \lambda) X_0(s) \geq \alpha\,\delta/2$. Thus, X_0 is lower-bounded, a.s. on the event $\{t \leq T_B(\delta) \wedge T_1(m_1^\vee) \wedge \tau_\partial\}$, by the solution Y to:

$$dY_s = \alpha\,\delta/2\, ds + \sqrt{Y_s\,(1 - Y_s)} dB_0(s)\,, \qquad Y(0) = 0. \tag{6.11}$$

Note that Y is independent of d and x and 0 is an entrance boundary for Y, cf e.g. Subsection 3.3.3 in [12]. So we can choose $y_0 > 0$ such that: $\mathbb{P}(Y(t) \leq y_0) \leq \epsilon$. This implies that for any d and x:

$$\begin{aligned}\mathbb{P}_x^{(d)}(X_0(t) \leq y_0\,,\ t \leq T_1(m_1^\vee)\,,\ 2\,t < \tau_\partial) &\leq \mathbb{P}_x^{(d)}(T_B(\delta) \leq t \leq T_1(m_1^\vee)\,,\ 2\,t < \tau_\partial) \\ &\quad + \mathbb{P}_x^{(d)}(Y(t) \leq y_0\,,\ t \leq T_B(\delta) \wedge T_1(m_1^\vee) \wedge \tau_\partial) \\ &\leq 2\,\epsilon.\end{aligned}$$

Thanks to Lemma 6.0.4, we can choose a certain $m_1 \geq m_1^\vee$ associated to y_0. We thus deduce, with the Markov property at time t:

$$\begin{aligned}\mathbb{P}_x^{(d)}(2\,t < T_1(m_1) \wedge \tau_\partial) &\leq \mathbb{P}_x^{(d)}(X_0(t) \leq y_0\,,\ t \leq T_1(m_1^\vee)\,,\ 2\,t < \tau_\partial) \\ &\quad + \mathbb{E}_x^{(d)}[\mathbb{P}_{X(t)}^{(d)}(t \leq T_1(m_1) \wedge \tau_\partial)\ ;\ X_0(t) \geq y_0] \\ &\leq 3\,\epsilon.\end{aligned} \tag{6.12}$$

Thanks to Lemma 6.0.5, we can choose a certain $m_2 > 0$ associated to m_1 (with $T_2(m_2) := \inf\{t \geq 0\ ;\ M_2(X_t) \leq m_2\}$). and a certain $m_3 > 0$ associated to m_2 such that:

$$\mathbb{P}_x^{(d)}(3\,t < T_2(m_2) \wedge \tau_\partial) \leq 4\,\epsilon\,, \quad \mathbb{P}_x^{(d)}(4\,t < T_3(m_3) \wedge \tau_\partial) \leq 5\,\epsilon.$$

More generally, we can prove inductively that there exists m_k such that:

$$\mathbb{P}_x^{(d)}((k+1)\,t < T_k(m_k) \wedge \tau_\partial) \le (k+2)\,\epsilon,$$

so as to treat any moment. Since t is arbitrary, this concludes the proof of Proposition 6.0.2, given Lemmas 6.0.3-5 whose proofs follow. □

6.4 Proof of Lemma 6.0.3

This proof is an extension of the one of Proposition 3.8 in [1]. W.l.o.g., we assume $\delta \le \delta_\wedge := 1/(16\,\alpha)$. Consider an initial condition x such that there exists $m_1 \ge M_1(x) \wedge 1$ and $x_0\, m_1 \le \delta$, where δ is to be fixed later. Thus, on the event $\{\sup_{s\le t} X_0(s)\, M_1(s) \le 2\,\delta_\wedge\} \cap \{\sup_{s\le t} X_0(s) \le 1/2\}$, we have for any $s \le t$, $\quad (\alpha\, M_1(s) - \lambda)\, X_0(s) \le 2\,\alpha\,\delta_\wedge \le 1/4\,(1 - X_0(s))$. X_0 is thus upper-bounded on this event by the solution Y to:

$$dY_s = (1 - Y_s)/4\, ds + \sqrt{Y_s\,(1-Y_s)} dB_0(s)\,, \qquad Y(0) = y_0 := \delta/m_1.$$

The main interest of this upper-bound is that it is explicitly given as:

$$Y_t := y_0\, \exp\left[- \int_0^t \frac{1-Y_s}{4\,Y_s}\, ds + \int_0^t \sqrt{\frac{1-Y_s}{Y_s}}\, dB_0(s) \right],$$

which is an immediate consequence of Itô's formula. We then define the time-change:

$$\rho_t := \int_0^t \frac{1-Y_s}{Y_s}\, ds\,, \qquad W_t := M(\rho^{-1}(t)), \text{where } M_t := \int_0^t \sqrt{\frac{1-Y_s}{Y_s}}\, dB_0(s).$$
$$Y(\rho^{-1}(t)) := y_0\, \exp\left[-t/4 + W_t\right]$$

We can easily check from the quadratic variations that the martingale W is indeed a Brownian Motion. Through conditions on the law of $\exp[-t/4 + W_t]$ (independent of the parameters), we shall thus constrain Y, then X_0.

$$(\rho^{-1})'(t) = (\rho' \circ \rho^{-1}(t))^{-1} = \frac{y_0\, \exp[-t/4 + W_t]}{1 - y_0\, \exp[-t/4 + W_t]}.$$

For any $y > 0$, let: $\quad \tau_y^Y := \inf\{t \ge 0\ ;\ Y_t = y\}$ and remark that for any $\mu > 0$:

$$\{t < \tau_0^Y\} = \{t < \rho^{-1}(\infty)\} = \left\{ t < \int_0^\infty \frac{y_0\, \exp[-r/4 + W_r]}{1 - y_0\, \exp[-r/4 + W_r]} dr \right\}$$

On the event $\{\tau_0^Y < \tau_{y_0+\mu}^Y\}$, for any $t \ge 0$: $y_0 e^{-t/4 + W_t} < y_0 + \mu$, so that one can have an explicit upper-bound of $(1 - y_0 e^{-t/4+W_t})^{-1}$. On the event $\{\tau_{y_0+\mu}^Y < \tau_0^Y\}$, there must exist $t \ge 0$ such that $y_0 e^{-t/4+W_t} = y_0 + \mu$. From these, we deduce:

$$\mathbb{P}_{y_0}(t < \tau_0^Y < \tau_{y_0+\mu}^Y) \le \mathbb{P}\left(\frac{t\,(1 - y_0 - \mu)}{y_0} < \int_0^\infty \exp[-r/4 + W_r]\, dr \right), \tag{6.13}$$

$$\mathbb{P}_{y_0}(\tau_{y_0+\mu}^Y < \tau_0^Y) = \mathbb{P}\left((y_0 + \mu)/y_0 < \sup_{r\ge 0} \exp[-r/4 + W_r] \right). \tag{6.14}$$

Let $\epsilon > 0$. Since $W_t/t \underset{t\to\infty}{\longrightarrow} 0$, we can define $c_1, c_2 > 1$ such that

$$\mathbb{P}\left(c_1 < \int_0^\infty \exp[-r/4 + W_r]\ dr \right) \le \epsilon\,,\ \mathbb{P}\left(c_2 < \sup_{r\ge 0} \exp[-r/4 + W_r] \right) \le \epsilon.$$

Likewise, thanks to Lemma 3.2 in [1], we can choose a certain $c_3 > 0$ such that for any x:

$$\mathbb{P}_x(\sup_{\{s\leq t\}} M_1(s) - M_1(0) \geq \lambda\, t + c_3) \leq \epsilon.$$

This motivates: $m_1' := m_1 + \lambda\, t + c_3$, $\mu := \delta_\wedge / m_1'$.

We choose also $\delta \leq \delta_\wedge$ sufficiently small to ensure, with $m_1 \geq 1$:

$$\frac{t\,(1-y_0-\mu)}{y_0} \geq \frac{m_1\, t}{\delta} \times (1 - 2\frac{\delta_\vee}{m_1'}) \geq c_1\,, \qquad \frac{y_0+\mu}{y_0} \geq \frac{\mu}{y_0} \geq \frac{\delta_\wedge}{\delta} \times (1+\lambda\, t + c_3)^{-1} \geq c_2.$$

Thus, from equations (6.13) and (6.14) and the above definitions:

$$\mathbb{P}_x(\mathcal{A}) \geq 1 - 3\,\epsilon, \qquad \text{where } \mathcal{A} := \left\{\sup_{s\leq t} M_1(s) \leq m_1'\right\} \cap \left\{\tau_0^Y < t \wedge \tau_{y_0+\mu}^Y\right\}.$$

To check the upper-bound by Y, let $T_B := \inf\{s \geq 0\ ;\ X_0(s)\, M_1(s) \geq 2\,\delta_\wedge\}$. Then, on the event $\mathcal{A}$, for any $s \leq t \wedge T_B$ (recalling $\mu/y_0 \geq c_2 > 1$, where $\mu := \delta_\wedge/m_1'$) :

$$X_0(s) \leq Y_s \leq y_0 + \mu \leq 2\delta_\wedge/m_1'\ ,\ M_1(s) \leq m_1' \qquad \text{so } X_0(s)\, M_1(s) < 2\,\delta_\wedge.$$

By continuity of $X_0\, M_1$, $T_B < t$ is incompatible with $\mathcal{A}$, so that we indeed have $\forall s \leq t,\ X_0(s) \leq Y_s$, thus $\tau_\partial \leq t$. In conclusion, for any x such that $m_1 \geq 1$ and $x_0\, m_1 \leq \delta$:

$$\mathbb{P}_x(\tau_\partial \leq t) \geq \mathbb{P}_x(\mathcal{A}) \geq 1 - 3\,\epsilon.$$

□

6.5 Proof of Lemma 6.0.4

On the event $\{\inf_{s\leq t} M_1(X_s^{(d)}) \geq m_1\}$, X_0 is lower-bounded on $[0, t]$ by the solution Y to:

$$dY_s = r(m_1)\, Y_s\, ds + \sqrt{Y_s\ (1-Y_s)} dB_0(s)\ , \qquad Y(0) = y_0,$$
$$\text{where} \quad r(m_1) := \alpha\, m_1 - \lambda \underset{m_1\to\infty}{\longrightarrow} \infty.$$

Since $M_1(s) = 0$ as soon as $X_0 = 1$, this lower-bound cannot hold until $T_1^Y := \inf\{t \geq 0\ ;\ Y_t \geq 1\}$. We thus only have to prove that $\mathbb{P}(t < T_1^Y) \to 0$ as $m_1 \to \infty$.

Let $\epsilon, t_1 > 0$. The quadratic variation of the martingale part M_s until time $t_1 \leq t$ is upper-bounded by t_1, so that the Doob inequality implies:

$$\mathbb{P}_{y_0}(\sup_{s\leq t_1} |M_s| > y_0/2) \leq 8t_1/y_0^2. \tag{6.15}$$

By choosing t_1 sufficiently small, we can assume $8t_1/y_0^2 \leq \epsilon$. On the event $\left\{\sup_{s\leq t_1} |M_s| \leq y_0/2\right\}$, it is clear that Y stays above $y_0/2$ on the time-interval $[0, t_1]$. The drift term can thus be lower-bounded by $r(m_1)\, s\, y_0/2$ for any $s \leq t_1 \wedge T_1^Y$. Since it cannot exceed $1 - y_0/2$ before T_1^Y, it necessarily implies that for $r(m_1)$ sufficiently large (that is m_1 sufficiently large), we must have $T_1^Y < t_1$ on the event $\left\{\sup_{s\leq t_1} |M_s| \leq y_0/2\right\}$. With (6.15) and $t_1 \leq t$, this clearly implies $\mathbb{P}(t < T_1^Y) \to 0$ as $m_1 \to \infty$ and concludes the proof of Lemma 6.0.4. □

6.6 Proof of Lemma 6.0.5

For any $k \geq 1$:

$$dM_k(t) = \alpha\,(M_1(t) M_k(t) - M_{k+1}(t))\, dt + \lambda \sum_{j\leq k-1} c_j^k\, M_j(t) dt + d\mathcal{M}_k(t),$$

$$\text{where } \mathcal{M}_k(t) \text{ is a continuous martingale, and } \langle \mathcal{M}_k \rangle_t = \int_0^t (M_{2\,k}(s) - M_k(s))^2\, ds.$$

For a certain $m_1^{(k)}$ to be defined later, depending on ϵ, let $\tau_1^{(k)} := \inf\{t \geq 0 \ ; \ M_1(t) \geq m_1^{(k)}\}$ so that, since M_k is increasing with $k \geq 1$:

$$0 \leq \mathbb{E}_x(M_k(t \wedge \tau_1^{(k)}) \leq M_k(x) - \alpha\, \mathbb{E}_x \left(\int_0^{t\wedge\tau_1^{(k)}} M_{k+1}(s)\, ds \right) + C_k\, \mathbb{E}_x \left(\int_0^{t\wedge\tau_1^{(k)}} M_k(s)\, ds \right)$$

$$\mathbb{E}_x \left(\int_0^{t\wedge\tau_1^{(k)}} M_{k+1}(s)\, ds \right) \leq m_k/\alpha + C'_k\, \mathbb{E}_x \left(\int_0^{t\wedge\tau_1^{(k)}} M_k(s)\, ds \right).$$

By immediate induction:

$$\begin{aligned} \mathbb{E}_x \left(\int_0^{t\wedge\tau_1^{(k)}} M_{k+1}(s)\, ds \right) &\leq (k-1)\, m_k/\alpha + C''_k\, \mathbb{E}_x \left(\int_0^{t\wedge\tau_1^{(k)}} M_1(s)\, ds \right) \\ &\leq (k-1)\, m_k/\alpha + C''_k\, t\, m_1^{(k)}. \end{aligned}$$

For any $\epsilon > 0$, we then exploit Lemma 3.2 in [1] together with $M_1(x) \leq m_k$ to choose $m_1^{(k)}$ such that: $\mathbb{P}_x(\tau_1^{(k)} < t) \leq \epsilon$ for any $x \in \mathcal{X}_d$ such that $M_k(x) \leq m_k$. Now that $m_1^{(k)}, \tau_1^{(k)}$ is clearly defined, we can choose, by the Markov inequality, m_{k+1} such that:

$$\mathbb{P}_x \left(\int_0^{t\wedge\tau_1^{(k)}} M_{k+1}(s)\, ds \geq t\, m_{k+1} \right) \leq \epsilon.$$

This concludes the proof of Lemma 6.0.5:

$$\left\{ \inf_{s\leq t} M_{k+1}(X_s^{(d)}) \geq m_{k+1} \right\} \cap \left\{ t \leq \tau_1^{(k)} \right\} \subset \left\{ \int_0^{t\wedge\tau_1^{(k)}} M_{k+1}(s)\, ds \geq t\, m_{k+1} \right\}. \qquad \square$$

With this, the proof of Proposition 6.0.2 is completed and by extension the one of Proposition 2.3.2.

7 Proof of Theorem 2.3: the infinite dimensional case

As one can imagine, this final Section is much more technical than the previous ones. For instance, there is no explicit reference measure that seems to be exploitable as ζ: the Lebesgue measure cannot be extended on an infinite dimensional space ! Nonetheless, the core idea behind the proof is still that the individuals carrying many mutations are actually wiped out very rapidly, implying rapid shuffle of the last coordinates. Quite unexpectedly, the criteria we developed to deal with jump events has proved to be very effective in this context. Notably, we could deal with moments increasing too largely as exceptional events.

7.1 Main properties leading to the proof

The proof of Theorem 2.3 relies on the criteria stated in Subsection 2.2.1 of [26], as an extension to those in Section 2.1 of [25]. We will treat both the case of large yet finite values of d and $d = \infty$, for which we recall that any $x \in \mathcal{X}_\infty := \mathcal{X}^6$ has a finite sixth moment.

For the purpose of Theorem 2.3, we replace the notation given in (5.1) by the following one:

$$\mathcal{D}_\ell := \{x \in \mathcal{X}_d \ ; \ M_3(x) \leq \ell \, , \, x_0 \in [(3\ell)^{-1}, 1 - (3\ell)^{-1}]\}.$$

We prove Theorem 2.3 thanks to the following Theorems 7.1-3, ordered by difficulty, and Lemmas 7.1.2 and 7.1.3. We will see in Section 7.1 how these Theorems together with Theorems 2.2 in [26] imply Theorem 2.3. In the next subsections, we then prove Theorems 7.1-3 by order of appearance.

Escape from the Transitory domain. The sets E that we shall consider next are defined through three parameters $m_3,\, y > 0$ as follow:

$$E := \{x \in \mathcal{X}_d \ ; \ M_3(x) \le m_3 \ , \ \forall\, j \le \lfloor m_3/\eta \rfloor + 1,\ x_j \ge y\}. \tag{7.1}$$

We recall the notation τ_E as the entry time of E.

Theorem 7.1. *For any $t, \epsilon, \eta > 0$, there exists $m_3,\, y > 0$ such that, for any $d \in \mathbb{N} \cup \{\infty\}$:*

$$\sup_{d \in \mathbb{N}\cup\{\infty\}} \sup_{x \in \mathcal{X}_d} \mathbb{P}_x\left(t < \tau_\partial \wedge \tau_E\right) \le \epsilon.$$

In particular, for any $\rho, \eta > 0$ we can choose such m_3 and y such that:

$$\sup_{d \in \mathbb{N}\cup\{\infty\}} \sup_{x \in \mathcal{X}_d} \mathbb{E}_x\left(\exp[\rho\,(\tau_\partial \wedge \tau_E)]\right) \le 2.$$

Mixing property and accessibility. The two following theorems exploit a reference probability measure ζ on $\mathcal{X}_d$. This measure is chosen to be especially adapted for our arguments, in a way that makes it actually complex to express. Its specific definition is given in (7.22) by relying on the notations introduced in Subsections 7.4 and 7.5.

Theorem 7.2. *For any $d \in \mathbb{N} \cup \{\infty\}$, the probability measure ζ defined below in (7.22) satisfies the following uniform mixing condition. For any $\ell \ge 1$ and $t > 0$, there exists $L > \ell$ and $c > 0$ such that for any $d \in \mathbb{N} \cup \{\infty\}$ and $x \in \mathcal{D}_\ell$, with $T_{\mathcal{D}_L}$ the exit time out of $\mathcal{D}_L$:*

$$\mathbb{P}^{(d)}_x(X(t) \in dx \ ; \ t < T_{\mathcal{D}_L} \wedge \tau_\partial) \ge c\,\zeta(dx).$$

Remark 7.1.1. *It can be noted that for any $x \in \mathcal{D}_\ell$, $T_{\mathcal{D}_L} < \tau_\partial$, so that the latter may be omitted in the expression.*

Absorption with failures. We will consider any E of the form prescribed by Theorem 7.1 and again the same probability measure ζ.

Theorem 7.3. *Given any $\rho, m_3, \eta, y > 0$ and any $\epsilon \in (0, 1)$, there exists $t_\veebar, t_A, c_A > 0$ such that for any $d \in \mathbb{N} \cup \{\infty\}$ and $x \in E$, there exists a stopping time U_A such that for any $x_\zeta \in E$:*

$$\begin{aligned} &\{\tau_\partial \wedge t_\veebar \le U_A\} = \{U_A = \infty\} \qquad \textit{and} \qquad \mathbb{P}_x(U_A = \infty,\, t_\veebar < \tau_\partial) \le \epsilon\, \exp(-\rho\, t_\veebar),\\ &\textit{while} \quad \mathbb{P}_x\left(X(U_A) \in dy \ ; \ U_A < \tau_\partial\right) \le c_A\, \mathbb{P}_{x_\zeta}\left(X(t_A) \in dy \ ; \ t_A < \tau_\partial\right). \end{aligned} \tag{7.2}$$

Moreover, there exists a stopping time U_A^∞ satisfying the following properties:

- *$U_A^\infty := U_A$ on the event $\left\{\tau_\partial \wedge U_A < \tau_E^1\right\}$, where $\tau_E^1 := \inf\{s \ge t_\veebar : X_s \in E\}$.*
- *On the event $\left\{\tau_E^1 < \tau_\partial\right\} \cap \{U_A = \infty\}$, and conditionally on $\mathcal{F}_{\tau_E^1}$, the law of $U_A^\infty - \tau_E^1$ coincides with the one of $\widetilde{U}_A^\infty$ for the shifted process $(\widetilde{X}_t)_{t \ge 0} := (X_{\tau_E^1 + t})_{t \ge 0}$.*

For the proof of this Theorem, it is helpful to replace the initial condition ζ by $x_\zeta \in E$. Without this issue of having an estimate uniform in d sufficiently large, we could simply impose additionally in our choice of $m_3 \ge 1$ and $y > 0$ that $\zeta(E) \ge 1/2$. A priori, there is no reason to expect that it could not hold globally in d, yet we do not see how to justify it clearly. Instead, we exploit the following Lemma, which is to be combined with the fact that the definition of ζ in (7.22) makes it supported on:

$$\mathcal{X}^d_* := \{x \in \mathcal{X}^d \ ; \ x_0 \ge 1/10\}.$$

Lemma 7.1.2. *There exists $m_3, y, c, t > 0$ such that for any d and ζ supported on $\mathcal{X}^d_*$, we have $\mathbb{P}^{(d)}_\zeta(X_t \in E) \ge c$. In this expression, E is defined as (7.1) in terms of m_3, y with the arbitrary choice of $\eta := 1$.*

Proof of Lemma 7.1.2. We have thanks to Theorem 7.1 a uniform control on the time of coming back to E, provided we can handle the survival starting from ζ. From (7.22) and the definition of τ^0, Since ζ is supported on $\mathcal{X}_*^d$, we deduce that under $\mathbb{P}_\zeta^{(d)}$, whatever d, $(X_0(s))$ is lower-bounded by the solution Y to the equation:

$$dY_s = -\lambda\, ds + \sqrt{Y_s\ (1 - Y_s)} dB_0(s)\ , \qquad Y(0) = 1/10.$$

Thus, denoting $c := \mathbb{P}_{1/10}(Y_{t/2} > 0)/2 > 0$ independent of d, we have uniformly:

$$\mathbb{P}_\zeta^{(d)}(t/2 < \tau_\partial) \geq 2c. \tag{7.3}$$

Thanks to Theorem 7.1, with the arbitrary choice of $\eta = 1$, we then deduce $m_3, y > 0$ such that for any d and $x \in \mathcal{X}_d$:

$$\mathbb{P}_\zeta^{(d)}(t/2 < \tau_\partial \wedge \tau_E) \leq c.$$

Combined with (7.3), this concludes the proof of Lemma 7.1.2. □

Lower-bound of the survival capacity. The sets $(\mathcal{D}_\ell)$ do not completely cover the state space, yet:

$$\bigcup_{\ell \geq 1} \mathcal{D}_\ell = \mathcal{X}_d \setminus \{\delta_0\}, \qquad \text{where } \{\delta_0\} := (1, 0, 0, ...).$$

To deduce that h is positive on δ_0, we exploit the following Lemma:

Lemma 7.1.3. *For any $y_0 > 0$ and uniformly in $d \in \mathbb{N} \cup \{\infty\}$, there exists $t > 0$, $\ell \geq 1$ such that the sets $H_d := \{x \in \mathcal{X}_d : x_0 \geq y_0\}$ satisfy:*

$$\inf_{\{d \geq 1, x \in H_d\}} \mathbb{P}_x(\tau_{\mathcal{D}_\ell} \leq t \wedge \tau_\partial) > 0.$$

In particular, there exists ℓ such that $\mathbb{P}_{\delta_0}(\tau_{\mathcal{D}_\ell} < \tau_\partial) > 0$. The proof of Lemma 7.1.3 is a straightforward adaptation of the one of Lemma 5.1.4 by replacing Proposition 5.1.3 by Theorem 7.1. The reader will be spared further details. We note that the specific definition of the sets $\mathcal{D}_\ell$, which are different for the two lemmas, is actually not involved.

7.2 Proof of Theorem 2.3 with Theorems 7.1-3 and Lemmas 7.1.2-3.

For this proof, we plan again to exploit Theorem 2.4 and ensure this time Assumption (**AF**).

First, it is clear that the sets $\mathcal{D}_\ell$ satisfy assumption $(A0_S)$. From Theorem 7.2, assumption $(A1)$ holds true for the reference measure ζ defined by (7.22). This measure ζ defines a value for ρ_S From Lemma 3.0.2 in [25] and $(A1)$, we know that ρ_S is upper-bounded by a certain value $\widetilde{\rho}_S$ that only depends on the constants involved in $(A1)$. In order to satisfy $\rho > \rho_S$, we set $\rho := 2\widetilde{\rho}_S$.

From Theorem 7.1 and arbitrary imposing $\eta = 1$, we deduce $m_3, y > 0$ such that assumption $(A2)$ holds for this value of ρ and E defined as (7.1).

Since increasing m_3 and reducing y make the corresponding subset E increase, we assume without restriction that these values are larger than the ones specified in Lemma 7.1.2. Note that the definition of ζ given in (7.22) ensures that it is supported on $\mathcal{X}_*^d$ (because of the constraint $t_M < \tau_0$). It means that there exists $t_1, c_1 > 0$ such that for any d, we have $\mathbb{P}_\zeta^{(d)}(X_{t_1} \in E) \geq c_1$.

For the proof of $(A3_F)$, we use this choice of m_3, y, η and ρ and exploit Theorem 7.3 for any given $\epsilon > 0$ to define $t_\vee, t_A, c_A > 0$ and to express U_A as a function of x. Note that we exploit the Markov property to get the following inequality:

$$\mathbb{P}_x^{(d)}(X_{U_A} \in dx'\ ,\ U_A < \infty) \leq c_1.c_A\, \mathbb{P}_\zeta^{(d)}(X(t_1 + t_A) \in dx'\ ;\ t_1 + t_A < \tau_\partial).$$

This concludes that Assumption (**AF**) holds true. Exploiting Theorem 2.4, this concludes the proof that the semi-group displays QSC and that the Q-process exists on $\mathcal{H}$. Thanks to Proposition 2.5.3 and Lemma 7.1.3, we deduce the required positive lower-bounds of h_∞, notably implying $\mathcal{H} = \mathcal{X}$.

Since all the parameters can be chosen independently of d sufficiently large, this is also the case of any parameter involved in the convergences. One can indeed check that there are intricate yet explicit relations between all the parameters introduced in [26]. This concludes the proof of Theorem 2.3. □

7.3 Proof of Theorem 7.1

The proof of Theorem 7.1 relies on the following Lemmas, easily adapted from the uniform escape and the uniform descent of the moments in the finite dimensional systems.

Lemma 7.3.1. *For any $t > 0$:*

$$\sup\Big\{\mathbb{P}_x\left(t < \tau_\partial\right) \Big| d \in \mathbb{N} \cup \{\infty\}\ ,\ x \in \mathcal{X}_d\ ,\ M_1(x) \in (1, \infty)\ ,\ x_0\, M_1(x) \leq \delta\Big\} \text{ tends to 0}$$

as δ tends to 0.

Lemma 7.3.2. *Given any $t, y_0 > 0$, and recalling $T_1(m_1) := \inf\{t \geq 0\ ;\ M_1(X_t) \leq m_1\}$:*

$$\sup\Big\{\mathbb{P}_x\left(t \leq T_1(m_1) \wedge \tau_\partial\right) \Big| d \in \mathbb{N} \cup \{\infty\}\ ,\ x \in \mathcal{X}_d\ ,\ x_0 \geq y_0\Big\} \text{ tends to 0 as } m_1 \text{ tends to } \infty.$$

Lemma 7.3.3. *For any $t, m_1 > 0$:*

$$\sup\Big\{\mathbb{P}_x\left(t < \tau_\partial\right) \Big| d \in \mathbb{N} \cup \{\infty\}\ ,\ x \in \mathcal{X}_d\ ,\ M_1(x) \leq m_1\ ,\ x_0 \leq \delta\Big\} \text{ tends to 0}$$

as δ tends to 0.

From the finite dimensional case, we adapt the definition, for any $y > 0$ and $k \geq 0$, of:

$$T_y^k := \inf\{s \geq 0\ ;\ \forall j \leq k,\ X_j(s) \geq y\}.$$

Lemma 7.3.4. *Given any $k \in \mathbb{N}$, and t, $m_1, y_0 > 0$:*

$$\sup\Big\{\mathbb{P}_x\left(t < T_y^k \wedge \tau_\partial\right) \Big| d \in [\![k, \infty[\![\cup \{\infty\}\ ,\ x \in \mathcal{X}_d\ ,\ x_0 \geq y_0\ ,\ M_1(x) \leq m_1\Big\} \text{ tends to 0}$$

as y' tends to 0.

Lemma 7.3.5. *Given any $k \in \mathbb{N}$ and t, $m_1, y > 0$:*

$$\sup\Big\{\mathbb{P}_x\left(\exists\, j \leq k,\ \exists\, s \leq 3\,t,\ X_j(s) \leq y'\ ,\ 4\,t < \tau_\partial\right) \Big| d \in [\![k, \infty[\![\cup \{\infty\}\ ,\ x \in \mathcal{X}_d,$$
$$M_1(x) \leq m_1\ ;\ \forall j \leq k,\ x_j \geq y\Big\}$$

tends to 0 as y' tends to 0.

Lemma 7.3.6. *Given any $k \in \mathbb{N}$ and t, $m_k > 0$:*

$$\sup\Big\{\mathbb{P}_x\left(\inf_{s \leq t} M_{k+1}(X_s) \geq m_{k+1}\right) \Big| d \in [\![k, \infty[\![\cup \{\infty\}\ ,\ x \in \mathcal{X}_d\ ,\ M_{2\,k}(x) < \infty\ ,\ M_k(x) \leq m_k\Big\}$$

tends to 0 as m_{k+1} tends to ∞.

We let the reader check that the proofs of Lemmas 7.3.1, 7.3.2, 7.3.4 and 7.3.6 can be adapted mutatis mutandis from the ones of resp. Lemmas 6.0.3, 6.0.4, 5.5.2 and 6.0.5.

As a generalization of Lemma 5.5.1, Lemma 7.3.3 is a consequence of the fact that X_0 is upper-bounded on the event $\{\sup_{s\leq t} M_1(s) \leq m_1'\}$ by the solution Y of:

$$dY_t = \alpha\, m_1'\, Y_t\, dt + \sqrt{Y_t(1-Y_t)}dB_0(t)\ ,\ Y_0 = \delta,$$

for which it is known that 0 is an absorbing value. Thanks to Lemma 3.2 in [1], we know an upper-bound going to 0 as m_1' goes to ∞ of $\mathbb{P}_x(\sup_{s\leq t} M_1(s) \geq m_1')$, uniform in $x \in \mathcal{X}_d$ such that $M_1(x) \leq m_1$.

We begin with the remaining proof of Lemma 7.3.5, and then conclude the proof of Theorem 7.1.

7.3.1 Step 1: proof of Lemma 7.3.5

Let $k \geq 1$, $t, m_1, y > 0$. For $\delta > 0$, that we shall choose sufficiently small, let: $\tau_\delta := \inf\{t \geq 0\ ;\ X_0(t) \leq \delta\}$. By Lemma 3.2 in [1], $M_1((3t)\wedge\tau_\delta) \leq m_1'$ with probability close to one for m_1' sufficiently large and independently of δ. Thus, thanks to Lemma 7.3.3, choosing δ sufficiently small ensures that on the event $\{\tau_\delta \leq 3t\}$ extinction before $4\,t$ happens with a probability close to one.

We restrict ourselves in the following to the event $\{3\,t < \tau_\delta\}$. Now, for $1 \leq i \leq k$, and $y_{i-1} > 0$, let:

$$\tau^{i-1}(y_{i-1}) := \inf\{t \geq 0\ ;\ \exists\, j \leq i,\ X_j(t) \leq y_{i-1}\}.$$

The proof relies on an induction over the coordinates $1 \leq i \leq k$ that there exits $0 < y_i \leq y_{i-1}$ such that $3\,t < \tau^i(y_i)$ with a probability close to 0 conditionally on the event $\{3\,t < \tau^{i-1}(y_{i-1})\}$.

On the event $\{3t < \tau^{i-1}(y_{i-1})\}$, we can observe:

$$dX_i(t) \geq \lambda\, y_{i-1}\, dt - (i\,\alpha + \lambda)\, X_i(t)\, dt + \sqrt{X_i(t)\,(1 - X_i(t))}\, dB_i(t),$$

for some Brownian Motion B_i (these are clearly not independent for different values of i). By some comparison principle, for instance Proposition 3.12 in [21], X_i is lower-bounded (uniformly in x) by the solution to the SDE:

$$dY_i(t) = \lambda\, y_{i-1}\, dt - (i\,\alpha + \lambda)\, Y_i(t)\, dt + \sqrt{Y_i(t)\,(1 - Y_i(t))}\, dB_i(t),$$

with $Y_i(0) = y$ and absorption at 1. Note that Y_i cannot be absorbed at 1 before $\tau^{i-1}(y_{i-1})$ by the definition of the latter and that Y_i does not depend on d.

Now, for any i, 0 is an entrance boundary for Y_i, cf e.g. Subsection 3.3.3 in [12], so that there exists $0 < y_i \leq y_{i-1}$ such that with a probability close to 1 conditionally on the event $\{3t \leq \tau^{i-1}(y_{i-1})\}$:

$$\inf_{s\leq t} Y_i(s) \geq y_i \qquad \text{thus } t \leq \tau^i(y_i).$$

More precisely, for any ϵ, the above arguments shows by induction that there exists a decreasing sequence $(y_i)_{1\leq i\leq k} \in (\mathbb{R}_+^*)^k$, with $y_0 = \delta$, such that:

$$\sup\left\{\mathbb{P}_x\left(4t < \tau_\partial \,\middle|\, 3t \leq \tau^0_{y_0}\right) \,\middle|\, x \in \mathcal{X}_d, M_1(x) \leq m_1\ ;\ \forall\, j \leq k,\ x_j \geq y\right\} \leq \epsilon/2,$$
$$\sup\left\{\mathbb{P}_x\left(\tau^i_{y_i} < 3t \,\middle|\, 3t \leq \tau^{i-1}_{y_{i-1}}\right) \,\middle|\, x \in \mathcal{X}_d, M_1(x) \leq m_1\ ;\ \forall\, j \leq k,\ x_j \geq y\right\} \leq \epsilon/2^{i+1}.$$

Now, since an immediate induction ensures that:

$$\mathbb{P}_x(\tau^k(y_k) \leq 3\,t\ ,\ 4\,t < \tau_\partial) \leq \mathbb{P}_x(4\,t < \tau_\partial \,|\, \tau_\delta \leq 3t) + \sum_{i=1}^k \mathbb{P}_x\left(\tau^i_{y_i} \leq 3t \,\middle|\, 3t < \tau^{i-1}_{y_{i-1}}\right),$$

we can indeed conclude that the probability of $\{\tau^k(y_k) \leq 3\,t\} \cap \{4\,t < \tau_\partial\}$ is uniformly upper-bounded by ϵ. This ends the proof of Lemma 7.3.5. □

7.3.2 Step 2: final proof of Theorem 7.1

With exactly the same reasoning as for Proposition 6.0.2, with Lemmas 7.3.1 and 7.3.2 instead of 6.0.3 and 6.0.4, we deduce that for any $x \in \mathcal{X}_d$, $t > 0$ and ϵ, we can find m_1 such that:

$$\mathbb{P}_x(\mathcal{E}_1) \leq 3\,\epsilon, \quad \text{with } \mathcal{E}_1 := \{2\,t < T_1(m_1) \wedge \tau_\partial\}$$

From Lemma 7.3.3, we can choose y_0 such that:

$$\mathbb{P}_x(\mathcal{E}_2) \leq \epsilon, \quad \text{with } \mathcal{E}_2 := \{X_0(T_1(m_1)) \leq y_0\} \cap \{T_1(m_1) \leq 2\,t\} \cap \{3\,t < \tau_\partial\}.$$

Again with Lemma 3.2 in [1] and the Markov property, we choose $m_1' > 0$ such that:

$$\begin{aligned} \mathbb{P}_x(\mathcal{E}_4) \leq \epsilon, \quad \text{with } \mathcal{E}_4 := &\{T_1(m_1) + t < \tau_\partial\} \\ &\cap \{\exists\, s \in [T_1(m_1), T_1(m_1) + t],\ M_1(s) \geq m_1'\} \end{aligned}$$

Before we ensure that many components escape 0 during this time-interval $[T_1(m_1), T_1(m_1) + t]$, we need to know the number of components needed, which is given by the next step. Thus, thanks to Lemma 7.3.6 and again the Markov property at $T_1(m_1) + t$, we choose first a certain m_3 (with an implicit step for the second moment) such that:

$$\begin{aligned} \mathbb{P}_x\left(\mathcal{E}_6\right) \leq \epsilon, \quad \text{with } \mathcal{E}_6 := &\{T_1(m_1) + t < \tau_\partial\} \cap \{M_1(T_1(m_1) + t) \leq m_1'\} \\ &\cap \{T_1(m_1) + 3\,t \leq \widetilde{T}_3(m_3)\} \\ \text{and } \widetilde{T}_3(m_3) := \inf\{s \geq T_1(m_1) + t\ ;\ &M_3(s) \leq m_3\} \end{aligned}$$

Now, we can define $k := \lfloor m_3/\eta \rfloor + 1$ (η being an imposed parameter in the statement of Theorem 7.1) and choose thanks to Lemma 7.3.4 a certain y such that:

$$\begin{aligned} \mathbb{P}_x\left(\mathcal{E}_3\right) \leq \epsilon, \quad \text{with } \mathcal{E}_3 := &\{T_1(m_1) < \tau_\partial\} \cap \{X_0(T_1(m_1)) \geq y_0\} \\ &\cap \{T_1(m_1) + t \leq \widetilde{T}_y^k\} \\ \text{where } \widetilde{T}_y^k := \inf\{s \geq T_1(m_1)\ ;\ &\forall\, j \leq k,\ X_j(s) \geq y\}. \end{aligned}$$

Finally, we choose thanks to Lemma 7.3.5 a certain y' such that:

$$\begin{aligned} \mathbb{P}_x\left(\mathcal{E}_5\right) \leq \epsilon, \quad \text{with } \mathcal{E}_5 := &\{\widetilde{T}_y^k + 4\,t < \tau_\partial\} \cap \{M_1(\widetilde{T}_y^k) \leq m_1'\} \\ &\cap \{\exists\, s \in [\widetilde{T}_y^k, \widetilde{T}_y^k + 3\,t],\ \exists\, j \leq k,\ X_j(s) \leq y'\} \end{aligned}$$

Provided that we prove that the event $\mathcal{E} := \{6\,t < \tau_\partial \wedge \tau_E\}$ (with y' instead of y in the definition of τ_E) is necessarily included in the union of the exceptional events we have just defined, this ensures: $\forall\, x \in \mathcal{X}_d, \quad \mathbb{P}_x(6\,t < \tau_\partial \wedge \tau_E) \leq 8\,\epsilon$ and concludes the proof since t and ϵ have been arbitrary chosen.

On $\mathcal{E} \setminus \mathcal{E}_1$, we know $T_1(m_1) \leq 2\,t$.

On $\mathcal{E} \setminus \cup_{i=1}^2 \mathcal{E}_i$, we deduce also $X_0(T_1(m_1)) \geq y_0$.

On $\mathcal{E} \setminus \cup_{i=1}^3 \mathcal{E}_i$: $\quad \widetilde{T}_y^k \leq T_1(m_1) + t \leq 3\,t$.

On $\mathcal{E} \setminus \cup_{i=1}^4 \mathcal{E}_i$: $\quad M_1(\widetilde{T}_y^k) \vee M_1(T_1(m_1) + t) \leq m_1'$.

On $\mathcal{E} \setminus \cup_{i=1}^5 \mathcal{E}_i$: $\quad \forall\, s \in [\widetilde{T}_y^k, \widetilde{T}_y^k + 3\,t],\ \forall\, j \leq k,\ X_j(s) \geq y'$.

On $\mathcal{E} \setminus \cup_{i=1}^6 \mathcal{E}_i$: $\quad \widetilde{T}_3(m_3) \leq T_1(m_1) + 3\,t \leq 5\,t$.

Since moreover $\widetilde{T}_y^k \leq T_1(m_1) + t$, while, by definition of $\widetilde{T}_3(m_3)$, $\widetilde{T}_3(m_3) \geq T_1(m_1) + t$, we deduce: $\widetilde{T}_3(m_3) \in [\widetilde{T}_y^k, \widetilde{T}_y^k + 3\,t]$. As a consequence: $\forall\, j \leq k,\ X_j(\widetilde{T}_3(m_3)) \geq y'$. Then, it would imply $\tau_E \leq \widetilde{T}_3(m_3) \leq 5\,t$, which contradicts the definition of $\mathcal{E}$. Thus: $\mathcal{E} \subset \cup_{i=1}^6 \mathcal{E}_i$, and the conclusion of Theorem 7.1 is proved. □

Transformation of the system of SDEs

The changes in the description of the system specified in the next two sections will be crucial for both the proofs of Theorems 7.3 and 7.2. Up to a multiplicative constant in the probabilities, they makes it possible to gather the last coordinates in one specific block while keeping a Markovian description. Our aim is then to prove that the dependence in the initial values of these last coordinates vanishes very quickly. We split the system of SDEs to distinguish the "descendants" of these doomed lineages from the unlucky newcomers that have acquired additional mutations (whose traits are predictable).

7.4 Aggregation of the last coordinates

In the following, for any $d \in \mathbb{N} \cup \{\infty\}$ and $k \leq d$, we denote by $\mathbb{P}^{(k,d)}$ the law of the solution to:

$$\begin{aligned}\forall\, i \leq d, \quad dX_i(t) &= \alpha(M_1^{(k)}(t) - i \wedge k)\, X_i(t)\, dt + \lambda(X_{i-1}(t) - \mathbf{1}_{\{i<d\}}\, X_i(t))\, dt \\ &\quad + \sqrt{X_i(t)}\, dW_t^i - X_i(t)\, dW_t \qquad (7.4):(S^{(k,d)})\end{aligned}$$

$$\text{where } W_t := \sum_j \int_0^t \sqrt{X_j(s)} dW_s^j,$$

$$M_1^{(k)}(t) := \sum_i (i \wedge k)\, X_i(t) \quad = \sum_{i \leq k-1} i\, X_i(t) + k \sum_{i \geq k} X_i(t),$$

with $(W^i)_{i \geq 0}$ a family of mutually independent Brownian Motions.

For the following proposition, we shall exploit a control on moments of order δ relying on the stopping time:

$$\tau_m^\delta := \inf\{s \geq 0\ ;\ M_\delta(s) \geq m\}.$$

Proposition 7.4.1. *Given any $t, \epsilon > 0$, $\delta > 2$, there exists $C_M, C_G > 0$ for which the following holds. For any $m \geq 1$, with $m' := C_M \times m$, for any $d \in \mathbb{N} \cup \{\infty\}$, $k \leq d$, and any $x \in \mathcal{X}_d \cap \mathcal{X}^{2\delta}$ such that $M_\delta(x) \leq m$, there exists a coupling between $\mathbb{P}^{(k,d)}$ and $\mathbb{P}^{(d)}$ such that:*

$$\text{on the event } \{t < \tau_{m'}^\delta\} : \quad \left| \log \left(\frac{d\mathbb{P}_x^{(k,d)}}{d\mathbb{P}_x^{(d)}}_{|[0,t]} \right) \right| \leq C_G \frac{m}{k^{\delta-2}},$$

$$\text{where } m' \text{ is such that } \mathbb{P}_x^{(d)} \left(\tau_{m'}^\delta \leq t \right) \leq \epsilon.$$

An analogous result holds for $\delta := 2$, except that the upper-bound on the log-ratio of densities is then $C_G \times m$.

Remark 7.4.2. *In practice, we shall exploit Proposition 7.4.1 only for $\delta = 3$. Yet, the proof is almost the same for any moment provided $\delta > 2$, while we mentioned earlier that the requirement that $M_6(x) < \infty$ could be replaced by the requirement $M_{2\delta}(x) < \infty$. So we treat Proposition 7.4.1 in this generality and let the interested reader extend the argument.*

This transform is naturally associated to the following projection π_k from $\mathcal{X}_d$ to $\mathcal{X}_k$, given by:

$$\pi_k(x)_i \begin{cases} = x_i, \ \text{if } i \leq k-1, \\ = \displaystyle\sum_{j=k}^{d} x_j = 1 - \sum_{j=0}^{k-1} x_j, \ \text{if } i = k, \end{cases} \qquad (7.5)$$

For the following proposition, we also define $(X_i^{[F]})_{i\in[\![k,d]\!]}$ as the solution to:

$$\begin{aligned} dX_i^{[F]}(t) &:= \lambda\frac{X_{k-1}(t)}{X_{(k)}(t)}(\mathbf{1}_{\{i=k\}} - X_i^{[F]}(t))\,dt + \lambda(X_{i-1}^{[F]}(t)\mathbf{1}_{\{i\geq k+1\}} - \mathbf{1}_{\{i<d\}}\,X_i^{[F]}(t))\,dt \\ &\quad + \sqrt{\frac{X_i^{[F]}(t)}{X_{(k)}(t)}}dW_t^{[F],i} - \frac{X_i^{[F]}(t)}{\sqrt{X_{(k)}(t)}}dW_t^{[F]} \end{aligned} \tag{7.6}$$

$$\text{where } X_{(k)}(t) := 1 - \sum_{i\leq k-1} X_i(t)\,,\; W_t^{[F]} := \sum_{i\geq k}\int_0^t \sqrt{\frac{X_i^{[F]}(t)}{X_{(k)}(t)}}dW_t^{[F],i},$$

where $(W^{[F],i})_{i\in[\![k,d]\!]}$ is a sequence of independent Brownian Motion that are mutually independent from the $(W^i)_{i\in[\![0,d]\!]}$. $X^{[F]}$ shall play the role of the renormalized sequence of the last coordinates, and F stands for "Final".

Proposition 7.4.3. *For any $k \geq 1$, $\pi_k(X)$ is by itself a Markov process under any $\mathbb{P}_x^{(k,d)}$, whose law is independent of $d \in [\![k,\infty[\![\cup\{\infty\}$. The vector X under the law $\mathbb{P}_x^{(k,d)}$ has the same law as the vector $(X_i : 0 \leq i \leq k-1\ ;\ X_{(k)} \times X_i^{[F]} : i \geq k)$, where the $X_i^{[F]}$ are defined in* (7.6).

The control of the increase in the moments for Proposition 7.4.1 is obtained with the following proposition as the first step of proof.

Proposition 7.4.4. *For any $t > 0$, there exists $C \geq 1$ such that for any m, m', $d \in \mathbb{N}\cup\{\infty\}$ and $x \in \mathcal{X}_d$ such that $M_3(x) \leq m$,*

$$\mathbb{P}_x^{(d)}\left(\tau_{m'}^3 \leq t\right) \leq \frac{Cm}{m'}.$$

Similarly, exploiting the decomposition in Proposition 7.4.3, we define:

$$M_3^{[F]}(t) := \sum_{i\geq k} i^3\,X_i^{[F]}(t) \in [k^3,\infty) \tag{7.7}$$

$$\tau_m^{[F],3} := \inf\{s \geq 0\ ;\ M_3^{[F]}(s) \geq m\},\quad m > 0. \tag{7.8}$$

For clarity, we define $\mathcal{F}^{(k)} = \sigma(W^i : i \leq k-1\ ;\ W)$. Recall that the process $X_i^{[F]}$ is driven by Brownian Motions $(W^{[F],i} : i \geq k)$ that are independent of $\mathcal{F}^{(k)}$. The inclusion $\sigma(\pi_k(X)) \subset \mathcal{F}^{(k)}$ is directly obtained through the autonomous set of equation verified by $\pi_k(X)$. The following control on $M_3^{[F]}$ exploits the filtration $\mathcal{F}_t^{(k)} := \mathcal{F}^{(k)} \vee \mathcal{F}_t$.

Proposition 7.4.5. *For any $t > 0$, there exists $C \geq 1$ such that for any m, m', $d \in \mathbb{N}\cup\{\infty\}$ and $x \in \mathcal{X}_d$ such that $M_3^{[F]}(x) \leq m$,*

$$\mathbb{P}_x^{(k,d)}\left(\tau_{m'}^{[F],3} \leq t \,\middle|\, \mathcal{F}^{(k)}\right) \leq \frac{Cm}{m'}.$$

7.4.1 Proof of Proposition 7.4.1

Step 1. We first express the Girsanov transform that makes it possible to relate $\mathbb{P}^{(k,d)}$ and $\mathbb{P}^{(d)}$. It is expressed in the following Lemma 7.4.6 in terms of the following processes:

$$R_1^{(k)}(t) := \sum_{i\geq k+1}(i-k)\,X_i(t)\,,\quad R_2^{(k)}(t) := \sum_{i\geq k+1}(i-k)^2\,X_i(t)$$

One can notice that they correspond to the expectation and variance of the vector $(\sum_{i=0}^k X_i;\ X_{k+1}; X_{k+2}; ...)$.

Lemma 7.4.6. *There exists a coupling between $\mathbb{P}^{(k,d)}$ and $\mathbb{P}^{(d)}$ such that:*

$$\log \frac{d\mathbb{P}_x^{(k,d)}}{d\mathbb{P}_x^{(d)}}_{|[0,t]} = \alpha\, R_1^{(k)}(0) - \alpha\, R_1^{(k)}(t) + \alpha^2 \int_0^t (M_1(s) - k)\, R_1^{(k)}(s)\, ds$$
$$- \alpha^2 \int_0^t R_2^{(k)}(s)\, ds + \alpha\,\lambda \int_0^t X_{(k)}(s)\, ds - \frac{\alpha^2}{2} \int_0^t \left[R_2^{(k)}(s) - R_1^{(k)}(s)^2\right] ds.$$

Proof: While applying the Girsanov formula, we consider the exponential martingale of:

$$L_t^{(k)} := -\alpha \sum_{i\geq k+1} (i-k) \int_0^t \sqrt{X_i(s)}\, dW_s^i + \alpha \int_0^t R_1^{(k)}(s)\, dW_s$$

By this choice, we obtain the following equalities:

$$d\langle L^{(k)}, W^i\rangle_s = \alpha \left[M_1(s) - M_1^{(k)}(s) - (i-k)_+\right] \sqrt{X_i(s)}\, ds,$$
$$d\langle L^{(k)}, W\rangle_s = -\alpha \sum_{i\geq k+1} (i-k)\, X_i(s)\, ds + \alpha\, R_1^{(k)}(s)\, ds = 0,$$
$$d\langle L^{(k)}\rangle_s = -\alpha \sum_{i\geq k+1} (i-k)\, \sqrt{X_i(s)} d\langle L^{(k)}, W^i\rangle_s = \alpha^2 \left[R_2^{(k)}(t) - R_1^{(k)}(t)^2\right]$$

The coupling can thus indeed be given by the exponential martingale associated to $L^{(k)}$, i.e.:

$$\log \frac{d\mathbb{P}_x^{(k,d)}}{d\mathbb{P}_x^{(d)}}_{|[0,t]} = -\alpha \sum_{i\geq k+1} (i-k) \left[\int_0^t \sqrt{X_i(s)}\, dW_s^i - \int_0^t X_i(s)\, dW_s\right]$$
$$- \frac{\alpha^2}{2} \int_0^t \left[R_2^{(k)}(s) - R_1^{(k)}(s)^2\right] ds.$$

By noting the following equality

$$dR_1^{(k)}(s) = \alpha(M_1(s) - k)\, R_1^{(k)}(s)\, ds + \alpha\, R_2^{(k)}(s) + \lambda\, X_{(k)}(s)\, ds$$
$$+ \sum_{i\geq k+1} (i-k) \left[\int_0^t \sqrt{X_i(s)}\, dW_s^i - \int_0^t X_i(s)\, dW_s\right].$$

we can state the above expression as stated in Lemma 7.4.6 in term of the solution to $\mathbb{P}_x^{(d)}$.

Step 2: The aim is now to get uniform upper-bound on the expression given in Lemma 7.4.6. We note that:

$$0 \leq R_1^{(k)} \leq k^{-(\delta-1)} \sum_{i\geq k+1} i^\delta\, X_i(t) \leq k^{-(\delta-1)} M_\delta, \tag{7.9}$$

$$\text{Likewise} \quad 0 \leq M_1 \times R_1^{(k)} \leq k^{-(\delta-2)} M_1 \times M_{\delta-1} \leq k^{-(\delta-2)} M_\delta, \tag{7.10}$$

where we exploited Hölder's inequality to deduce: $M_1 \leq M_\delta^{1/\delta}$ and $M_{\delta-1} \leq M_\delta^{(\delta-1)/\delta}$. Similarly, $0 \leq X_{(k)} \leq k^{-\delta} M_\delta$, $0 \leq R_2^{(k)} \leq k^{-(\delta-2)} M_\delta$ while $(R_1^{(k)})^2 \leq R_2^{(k)}$ by the Cauchy-Schwarz inequality. Thanks to Lemma 7.4.6, this proves that on the event $\{t < \tau_{m'}^\delta\}$, for a certain constant C_1 only depending on δ and t (also on α and λ):

$$\Big|\log \frac{d\mathbb{P}_x^{(k,d)}}{d\mathbb{P}_x^{(d)}}_{|[0,t]}\Big| \leq C_1\, \frac{m'}{k^{\delta-2}}. \tag{7.11}$$

For ϵ, with the constant C_2 coming from Proposition 7.4.4, to prove $\mathbb{P}_x^{(d)}\left(\tau_{m'}^{\delta} \leq t\right) \leq \epsilon$, it suffices to choose $C_M := C_2/\epsilon$, where we recall $m' = C_M \times m$. With $C_G = C_1 \times C_2/\epsilon$, we deduce then from inequality (7.11) that on the event $\{t < \tau_{m'}^{\delta}\}$:

$$\Big| \log \frac{d\mathbb{P}_x^{(k,d)}}{d\mathbb{P}_x^{(d)}}_{|[0,t]} \Big| \leq C_G \, \frac{m}{k^{\delta-2}}.$$

The proof is quite the same for the case $\delta = 2$, except that we only require $(k \vee M_1) R_1^{(k)} \vee R_2^{(k)} \leq M_2$. This concludes the proof of Proposition 7.4.1. □

7.4.2 Proof of Proposition 7.4.3

By defining $X_{(k)}(t) := \sum_{i \geq k} X_i(t)$, we can remark that: $M_1^{(k)}(t) := \sum_{i \leq k-1} i \, X_i(t) + k \, X_{(k)}(t)$. Moreover, under $\mathbb{P}^{(k,d)}$:

$$\begin{aligned}
&dX_{(k)}(t) = \alpha(M_1^{(k)}(t) - k) \, X_{(k)}(t) \, dt + \lambda X_{k-1}(t) \, dt + \sqrt{X_{(k)}(t)} \, dW_t^{(k)} - X_{(k)}(t) \, dW_t, \\
&\qquad \text{where } W_t^{(k)} := \sum_{j \geq k} \int_0^t \sqrt{\frac{X_j(s)}{X_{(k)}(s)}} \, dW_s^j, \\
&W_t := \sum_i \int_0^t \sqrt{X_i(s)} dW_s^i = \sum_{i \leq k-1} \int_0^t \sqrt{X_i(s)} dW_s^i + \int_0^t \sqrt{X_{(k)}(s)} dW_s^{(k)}.
\end{aligned}$$

Here, $W^{(k)}$ indeed defines a Brownian Motion at least until τ_∂, since the mutation term ensures that $X_{(k)}$ stays positive, as it has been stated in Lemmas 7.3.4 and 7.3.5. (the way we may extend it afterwards plays no role). The correlation between $W^{(k)}$ and the $(W^i)_{i \leq k-1}$ remains zero, while they constitute a system of Brownian Motions under the same filtration. $W^{(k)}$ is thus independent from $\sigma(W^i; i \leq k-1)$, so that the system of equations satisfied by $\pi_k(X)$ is equivalent for any $\mathbb{P}_x^{(k,d)}$.

Concerning $X^{[F]}$, we define:

$$\begin{aligned}
d\bar{X}_i^{[F]}(t) := \lambda \frac{X_{k-1}(t)}{X_{(k)}(t)} (\mathbf{1}_{\{i=k\}} - \bar{X}_i^{[F]}(t)) \, dt + \lambda(\bar{X}_{i-1}^{[F]}(t) \mathbf{1}_{\{i \geq k+1\}} - \bar{X}_i^{[F]}(t)) \, dt \\
+ \sqrt{\frac{\bar{X}_i^{[F]}(t)}{X_{(k)}(t)}} dW_t^i - \frac{\bar{X}_i^{[F]}(t)}{\sqrt{X_{(k)}(t)}} dW_t^{(k)}.
\end{aligned}$$

Then, we define $\bar{X}_i$ as X_i for $i \leq k-1$ and as $X_{(k)} \times \bar{X}_i^{[F]}$ for $i \geq k$. Next, we check by exploiting the Itô formula that X_i coincide with $\bar{X}_i$ also for $i \geq k$, in which case:

$$\begin{aligned}
d\bar{X}_i(t) = \lambda(\bar{X}_{i-1}(t) - \bar{X}_i(t)) \, dt - \lambda \, X_{k-1}(t) \, \bar{X}_i(t) \, dt + \alpha(M_1^{(k)} - k) \, \bar{X}_i(t) \, dt + \lambda \, \bar{X}_{k-1}(t) \, \bar{X}_i(t) \, dt \\
+ \sqrt{\bar{X}_i(t)} dW_t^i - \sqrt{X_{(k)}(t)} \sqrt{\bar{X}_i^{[F]}(t)} dW_t^{(k)} + \sqrt{\bar{X}_i^{[F]}(t)} \sqrt{X_{(k)}(t)} dW_t^{(k)} \\
- \bar{X}_i(t) dW_t^i + \tfrac{1}{2} \, d\langle \mathcal{M}^{(k)}, \mathcal{M}^{[F],i} \rangle_t, \qquad (7.12)
\end{aligned}$$

where $\mathcal{M}_t^{(k)}$ and $\mathcal{M}_t^{[F],i}$ denote the martingale parts of resp. $X_{(k)}(t)$ and $\bar{X}_i^{[F]}(t)$. In fact, this covariation vanishes because, from the definitions of $(W^j, j \in \mathbb{Z}_+)$ and $W^{(k)}$, we first deduce:

$$\sqrt{\bar{X}_i^{[F]}(t)} d\langle W^i, W^{(k)} \rangle_t - \bar{X}_i^{[F]}(t) d\langle W^{(k)}, W^{(k)} \rangle_t \equiv 0. \qquad (7.13)$$

Concerning then the covariation with W and recalling $dW_t = \sum_{j \leq k-1} \sqrt{X_j(t)} dW^j(t) + \sqrt{X_{(k)}(t)} dW^{(k)}(t)$, since for any $j \in [\![0, k-1]\!]$, $\langle W^i, W^j \rangle \equiv 0$ and $\langle W^{(k)}, W^j \rangle \equiv 0$, we can conclude that $d\langle \mathcal{M}^{(k)}, \mathcal{M}^{[F],i} \rangle_t \equiv$

0. After simplification and replacing $dW_t^{(k)}$ by its expression involving $(\bar{X}_i)_{i\geq k}$, the system of equations (7.12) satisfied by $(\bar{X}_i)_{i\in[\![0,d]\!]}$ coincide with the system satisfied by $(X_i)_{i\in[\![0,d]\!]}$. By the uniqueness of the whole system, X_i coincide with $\bar{X}_i$.

Now, we have exploited in the previous calculation that the martingale component of $\bar{X}_i^{[F]}$, for $i \geq k$, has a quadratic co-variation with $W^{(k)}$ that stays null. Since these semi-martingales are also adapted to a common filtration $(\mathcal{F}_t)$ and their increments after time t is independent of $\mathcal{F}_t$, we deduce from Theorem 2.1.8 in [9] that $W^{(k)}$ is actually independent of the martingale components driving $\bar{X}^{[F]}$. Moreover, $\sigma(W^i : i \leq k-1)$ is by construction independent of $\sigma(W^i : i \geq k)$. Thus, considering $\sigma(W^i : i \leq k-1; W^{(k)})$, for which W is measurable, we deduce that it is independent of the family of martingale driving $\bar{X}^{[F]}$. We can thus replace the latter by the expression with a system of independent copies $W^{[F],i}$ as in Proposition 7.4.3 without changing the law of the vector. This ends the proof of Proposition 7.4.3.

□

7.4.3 Proof of Proposition 7.4.4

Let $t > 0$, $\delta > 1$, $m' \geq m$, $d \in \mathbb{N} \cup \{\infty\}$ and $x \in \mathcal{X}_d$ such that $M_\delta(x) \leq m$ be fixed. The proof generalizes the one of Lemma 6.0.1 in the case where the martingale part is a priori only local.

We again consider the semi-martingale decomposition of M_δ:

$$dM_\delta(t) = V_\delta(t)\,dt + d\mathcal{M}_\delta(t), \tag{7.14}$$

where $\mathcal{M}_\delta$ is a continuous local martingale starting from 0, whose quadratic variation is

$$\langle \mathcal{M}_\delta \rangle_t = \int_0^t (M_{2\,\delta}(s) - M_\delta(s)^2)\,ds,$$

and V_δ is a bounded variation process. Thanks to Theorem 3 in [1], since $M_{2\delta}(x) < \infty$, we know that $(M_{2\delta}(t))_{t\geq 0}$ is a.s. finite. This expression for the quadratic variation is thus well-defined.

Step 1: We first control the expectation of M_δ, by proving the following lemma.

Lemma 7.4.7. *Let $\delta, t > 0$ be given. There exists $C_M = C > 0$ such that for any $m \geq 1$, $x \in \mathcal{X}^{2\delta}$ such that $M_\delta(x) \leq m$, there exists a sequence of positive sub-martingale $\widehat{M}_\delta^{(k)}$ stopped at times T_k, where $T_k \to \infty$, such that for any $s \leq t \wedge T_k$, $M_\delta(s) \leq \widehat{M}_\delta^{(k)}(s)$, and such that:*

$$\mathbb{E}_x[\widehat{M}_\delta^{(k)}(t)] \leq C_M\, m.$$

Proof: The sequence T_k, for $k \geq 1$, is introduced to localize $\mathcal{M}_\delta$ and have an upper-bound on $M_\delta^{(k)}$.

$$T_k := \inf\{s \geq 0 \;;\; \langle \mathcal{M}_\delta \rangle_s \geq k \;,\; M_\delta^{(k)}(s) \geq k\}.$$

Recalling Theorem 3 in [1], for the fact that $(M_{2\delta}(t))_{t\geq 0}$ is a.s. finite, we easily deduce that $T_k \to \infty$ as $k \to \infty$.

We wish to characterize $\widehat{M}_\delta^{(k)}$ as the solution to the following equation, valid for $s \leq t \wedge T_k$:

$$\widehat{M}_\delta^{(k)}(s) = m + \int_0^s (C\widehat{M}_\delta^{(k)}(r) + \lambda)dr + \mathcal{M}_\delta^{(k)}(s). \tag{7.15}$$

We exploit Duhamel's formula to first define what will be $\widehat{M}_\delta^{(k)}(s) - M_\delta^{(k)}(s)$ on the event $\{s \leq T_k\}$:

$$E(s) := (m - M_\delta(x))\, e^{Cs} + e^{Cs} \int_0^s e^{-Cr}(CM_\delta(r) + \lambda - V_\delta(r))dr.$$

For any $s < T_k$, $E(s)$ is upper-bounded by a deterministic constant. Similarly as in Proposition 6.0.1 and thanks to the Hölder inequality, there exists $C = C(\delta) = \lambda(2^\delta - 1)$ such that: $V_\delta \leq C\, M_\delta + \lambda$. $E(s)$ is thus positive.

Let us check that $\widehat{M}_\delta^{(k)}(s) := M_\delta^{(k)}(s \wedge T_k) + E(s \wedge T_k)$ is indeed solution to equation (7.15). Let $s \leq T_k$ and compute:

$$\begin{aligned}
E(s) = e^{Cs} \times (e^{-Cs}E(s)) &= (m - M_\delta(x)) + \int_0^s e^{Cr} \times e^{-Cr}(CM_\delta(r) + \lambda - V_\delta(r))dr \\
&\quad + \int_0^s Ce^{Cr} \times (e^{-Cr}E(r))dr \\
&= (m - M_\delta(x)) + \int_0^s [C(E(r) + M_\delta(r)) + \lambda - V_\delta(r)]dr \\
&= \widehat{M}_\delta^{(k)}(0) - M_\delta(0) - \int_0^s V_\delta(r)dr + \int_0^s [C\widehat{M}_\delta^{(k)}(r) + \lambda]dr,
\end{aligned}$$

from which it is clear that $\widehat{M}_\delta^{(k)}(s)$ is indeed solution.

Recalling that E is positive, we immediately deduce that $\widehat{M}_\delta^{(k)}(s) > M_\delta(s)$ for any $s \leq T_k$. Since M_δ is non-negative by definition, it is also the case for $\widehat{M}_\delta^{(k)}$. Because of (7.15), with the fact that $\widehat{M}_\delta^{(k)}$ stays fixed after T_k, this proves that $\widehat{M}_\delta^{(k)}$ is a positive sub-martingale and that for any $s \leq t$:

$$\mathbb{E}_x[\widehat{M}_\delta^{(k)}(s)] \leq (m + \lambda t) + C \int_0^s \mathbb{E}_x[\widehat{M}_\delta^{(k)}(r)]dr.$$

Recall that both $M_\delta(s)$ and $E(s)$, are upper-bounded for any $s \leq t \wedge T_k$, by a uniform constant depending on t and k. This implies a similar upper-bound on $\mathbb{E}_x[\widehat{M}_\delta^{(k)}(r)]$, that guaranties that we are in conditions to apply Gromwall's Lemma, see for instance Proposition 6.59 in [21]. From this we deduce:

$$\mathbb{E}_x[\widehat{M}_\delta^{(k)}(t)] \leq (m + \lambda t)\, e^{Ct}.$$

This concludes the proof of Lemma 7.4.7 with $C_M := (1 + \lambda t)\, e^{Ct}$ (recalling $m \geq 1$).

Step 2: By exploiting Doob's inequality on $\widehat{M}_\delta^{(k)}$, we obtain:

$$\begin{aligned}
\mathbb{P}_x(\sup_{s \leq t \wedge T_k} M_\delta(s) > m') &\leq \mathbb{P}_x(\sup_{s \leq t} \widehat{M}_\delta^{(k)}(s) > m') \\
&\leq \frac{\mathbb{E}_x[\widehat{M}_\delta^{(k)}(t)]}{m'} \leq \frac{C_M\, m}{m'}.
\end{aligned}$$

We know let $T_k \to \infty$ and conclude the proof of Proposition 7.4.4 by showing that:

$$\mathbb{P}_x(\tau_{m'}^\delta \leq t) \leq \frac{C_M\, m}{m'}.$$

□

7.4.4 Proof of Proposition 7.4.5

Under $\mathbb{P}^{(k,d)}$, we exploit the Itô formula and distinguish the part involving the Brownian Motions. $M_3^{[F]}$ is solution of:

$$dM_3^{[F]}(t) := V_3^{[F]}(t)dt + d\mathcal{M}_3^{[F]}(t), \tag{7.16}$$

where $V_3^{[F]}$ is a bounded variation process defined as:

$$V_3^{[F]} := \lambda \frac{X_{k-1}}{X_{(k)}}(k^3 - M_3^{[F]}(t)) + \lambda \sum_{\ell \geq k} (\ell+1)^3 X_\ell^{[F]} - \lambda M_3^{[F]}. \tag{7.17}$$

Note that whatever the values of $\frac{X_{k-1}}{X_{(k)}}$, with the rough estimate $(\ell+1)^3 \leq 8\ell^3$ for $\ell \geq 1$, we always have $V_3^{[F]} \leq 7\lambda\, M_3^{[F]}$. On the other hand, $\mathcal{M}_3^{[F]}$ is defined as:

$$\mathcal{M}_3^{[F]} := \sum_{i \geq k} i^3 \left[\sqrt{\frac{X_i^{[F]}(t)}{X_{(k)}(t)}} dW_t^{[F],i} - \frac{X_i^{[F]}(t)}{\sqrt{X_{(k)}(t)}} dW_t^{[F]} \right]. \tag{7.18}$$

Relying on the same calculations as for M_3, $\mathcal{M}_3^{[F]}$ is a continuous local martingale starting from 0 for the filtration $\mathcal{F}_t^{(k)}$ whose quadratic variation is

$$\langle \mathcal{M}_3^{[F]} \rangle_t = \int_0^t \frac{\mathcal{M}_6^{[F]}(s) - (\mathcal{M}_3^{[F]}(s))^2}{X_{(k)}(s)}\, ds.$$

The rest of the proof of Proposition 7.4.5 can be easily adapted from the one of Proposition 7.4.4 (with $C = \exp[7\lambda t_H] \vee 1$). □

Now that we have all the tools we need for the first transform, we can turn in the next subsection to the second one.

7.5 Splitting of the solution

For any $d \in \mathbb{N} \cup \{\infty\}$ and $k \in [\![1, d]\!]$, consider the solution to: $\forall\, i \leq d$,

$$\begin{aligned} dX_i^{[G]}(t) &= \alpha(M_1^{(k)}(t) - i \wedge k)\, X_i^{[G]}(t)\, dt + \lambda(X_{i-1}^{[G]}(t) - \mathbf{1}_{\{i \neq d\}} X_i^{[G]}(t))\, dt \\ &\quad + \sqrt{X_i^{[G]}(t)}\, dW_t^{[G],i} - X_i^{[G]}(t)\, dW_t, \end{aligned} \tag{$S^{[G]}$}$$

$$\text{where } X_i^{[G]}(0) := x_i\, \mathbf{1}_{\{i \leq k-1\}}\ ; \quad X_{[R]}(t) := 1 - \sum_{i=0}^{d} X_i^{[G]}(t)\,;$$

$$W_t := \sum_i \int_0^t \sqrt{X_i^{[G]}(s)} dW_s^{[G],i} + \int_0^t \sqrt{X_{[R]}(s)} dW_s^{[R]},$$

$$M_1^{(k)}(t) := \textstyle\sum_{i \leq d} (i \wedge k)\, X_i^{[G]}(t) + k\, X_{[R]}(t).$$

Here, the $(W^{[G],i}, i \geq 0\,; W^{[R]})$ defines a mutually independent family of standard Brownian Motions. $[G]$ stands for "Generative" while $[R]$ stands for "Rest", with the idea that the $[R]$ component shall quickly get extinct. Notably, 0 is an absorbing state for $X_{[R]}$, whose absorption time is denoted $\tau_\partial^{[R]}$.

The following solutions are well-defined in the time interval $[0, \tau_\partial^{[R]})$.

$$\forall\, i \leq d, \quad dX_i^{[R]}(t) = \lambda\, (X_{i-1}^{[R]}(t) - \mathbf{1}_{\{i \neq d\}}\, X_i^{[R]}(t))\, dt + \sqrt{\frac{X_i^{[R]}(t)}{X_{[R]}(t)}}\, dW_t^{[R],i} - \frac{X_i^{[R]}(t)}{\sqrt{X_{[R]}(t)}}\, dW_t^{[R]}$$

$$\text{where } W_t^{[R]} := \sum_{i \geq 0} \int_0^t \sqrt{X_i^{[R]}(s)} dW_s^{[R],i}, \quad X_i^{[R]}(0) = \frac{x_i\, \mathbf{1}_{\{i \geq k\}}}{\sum_{j \geq k} x_j}. \tag{$S^{[R]}$}$$

Again, the $(W^{[R],i})$ define a mutually independent family of Brownian Motions, also mutually independent from the family $(W^{[G],i})$ (and not from $W^{[R]}$ of course). Let $\mathcal{F}^{[G]}$ be the filtration generated by the family $(W^{[G],i}, i \geq 0 \;;\; W^{[R]})$.

Looking at the equations for $X^{[G]}$, we see that it describes an autonomous system. We thus deduce the following fact.

Fact 7.5.1. *Considering two initial conditions x and x' such that $\forall\, i \leq k-1,\ x_i = x'_i$, $X^{[G]}$ under $\mathbb{P}_x$ has the same law as $\bar{X}^{[G]}$ under $\mathbb{P}_{x'}$.*

Our model is implicitly exploiting the following lemma:

Lemma 7.5.2. *For any t, $\sum_j X_j^{[R]}(t) = 1$. $W^{[R]}$ as defined in $S^{[R]}$ is a Brownian Motion, by construction independent from the family $(W^{[G],i})$. We are thus allowed to choose this same Brownian Motion in the coupling between the dynamics of $X^{[R]}$ and $X^{[G]}$.*

Lemma 7.5.3. *$X_{[R]}$ is solution to:*

$$dX_{[R]}(t) = \alpha\,(M_1^{(k)}(t) - k)\,X_{[R]}(t)\,dt + \sqrt{X_{[R]}(t)\,(1 - X_{[R]}(t))}\,d\widehat{W}_t^{[R]},$$

for a certain Brownian Motion $\widehat{W}^{[R]}$. Looking more precisely at the interactions with $X^{[G]}$, it is actually solution to:

$$dX_{[R]}(t) = \alpha\,(M_1^{(k)}(t) - k)\,X_{[R]}(t)\,dt + \sqrt{X_{[R]}(t)}\,dW_t^{[R]} - X_{[R]}(t)\,dW_t. \tag{7.19}$$

We see in the next lemma how these solutions are related to our initial problem.

Lemma 7.5.4. *The $(X_i^{[G]} + X_{[R]} \times X_i^{[R]})_{i\leq d}$ defines a solution to (7.4): $(S^{(k,d)})$.*

Moreover, an analogue of Proposition 7.4.4 can also be obtained in this setting. Let:

$$M_3^{[R]}(t) := \textstyle\sum_{\{i\geq k\}} i^3\, X_i^{[R]}(t), \tag{7.20}$$

$$\tau_m^{[R],3} := \inf\{s \geq 0 \;;\; M_3^{[R]}(s) \geq m\}, \quad m > 0. \tag{7.21}$$

Lemma 7.5.5. *For any $t > 0$, there exists $C \geq 1$ such that for any m, m', $d \in \mathbb{N} \cup \{\infty\}$ and $x \in \mathcal{X}_d$ such that $x_{(k)} > 0$ and $M_3^{[F]}(x) \leq m$,*

$$\mathbb{P}_x^{(k,d)}\left(\tau_{m'}^{[R],3} \leq t \,\middle|\, \mathcal{F}^{[G]}\right) \leq \frac{Cm}{m'}.$$

The proof of this lemma relies on the same ingredient as the ones of Propositions 7.4.4 and 7.4.5. The main difference is that the sequence $(T_k^{[R]})$, i.e. the analogue to the sequence (T_k) of Lemma 7.4.7, now satisfies $\lim_{k\to\infty} T_k^{[R]} \geq \tau_\partial^{[R]} := \inf\{t \geq 0 \;;\; X_{[R]}(t) = 0\}$. The reader will be spared further details. Next we prove Lemmas 7.5.2 to 7.5.4.

Remark 7.5.6. • *To define $(X_i^{[R]})_{i\geq 0}$, one can extend the well-defined solutions on $[0, \tau_n^{[R]}]$ where $\tau_n^{[R]} := \inf\{t \geq 0 \;;\; X_{[R]}(t) \leq 2^{-n}\}$.*

• *By construction, $X_i^{[R]}(t) = 0$ for any $t < \tau_\partial^{[R]}$ and $i \leq k-1$.*

• *$X^{[R]}$ gathers all the information related to dependence in the initially large components. It is doomed to disappear quickly when k is large, because it concerns only a very small fraction of the population, and the source term has been driven towards $X_k^{[G]}$. Note that $M_1^{(k)}(t) \leq k$ so that even the drift term will not help.*

7.5.1 Proof of Lemma 7.5.2:

The independence is clear and to prove that $W^{[R]}$ is actually a Brownian Motion, we simply have to check that for any $t \in [0, \tau_\partial^{[R]})$, $\sum_j X_j^{[R]}(t) = 1$. Let $Y_t := 1 - \sum_j X_j^{[R]}(t)$. By definition $Y_0 = 0$. This process satisfies on the other hand:

$$\begin{aligned} dY_t = -\textstyle\sum_{\{j\le d\}} dX_j^{[R]}(t) &= -\frac{1 - \sum_{\{j\le d\}} X_j^{[R]}(t)}{\sqrt{X_{[R]}(t)}} \cdot (\textstyle\sum_{\{j\le d\}} \sqrt{X_j^{[R]}(t)} dW_t^{[R],j}). \\ &= -\frac{Y_t}{\sqrt{X_{[R]}(t)\cdot(1-Y_t)}} d\widetilde{W}_t^{[R]}, \end{aligned}$$

where $\widetilde{W}_t^{[R]} := \sum_{\{i\le d\}} \int_0^t \frac{\sqrt{X_i^{[R]}(s)}}{\sqrt{Y_t}} dW_s^{[R],i}$ is a Brownian Motion. Classical methods as the ones exploited in [21] for comparison principles makes it possible to prove for any $m \ge 1$ that $Y_t = 0$ for any $t \le \inf\{s \ge 0\ ;\ X_{[R]}(s) \le 1/m\}$. This thus holds true for any $t < \tau_\partial^{[R]}$. □

7.5.2 Proof of Lemma 7.5.3

We deduce equation 7.19 from $(S^{[G]})$ since $\sum_{i\le d}(x_{i-1} - \mathbf{1}_{\{i\ne d\}} x_i(t)) = 0$ for any $x \in \mathcal{X}_d$ and:

$$\begin{aligned} \sum_i [M_1^{(k)}(t) - (i\wedge k)]\, X_i^{[G]}(t) &= M_1^{(k)}(t)\,[1 - X_{[R]}(t)] - \sum_i (i\wedge k)\, X_i^{[G]}(t) \\ &= k\, X_{[R]}(t) - M_1^{(k)}(t)\, X_{[R]}(t) \end{aligned}$$

From this, we deduce:

$$\begin{aligned} dX_{[R]}(t) &= \alpha(M_1^{(k)}(t) - k)\, X_{[R]}(t) + \sqrt{X_{[R]}(t)\,(1 - X_{[R]}(t))} d\widehat{W}_t^{[R]}, \\ \text{where } d\widehat{W}_t^{[R]} &:= \mathbf{1}_{\{t<\tau_\partial\}} \left(\sqrt{1 - X_{[R]}(t)}\, dW_t^{[R]} - \sqrt{X_{[R]}(t)} \sum_{i\le d} \sqrt{\frac{X_i^{[G]}(t)}{1 - X_{[R]}(t)}}\, dW_t^{[G],i} \right) \\ &\quad + \mathbf{1}_{\{t\ge\tau_\partial\}}\, dW_t^{(e)}, \end{aligned}$$

is indeed a Brownian Motion for $W^{(e)}$ a Brownian Motion, noting that $1 - X_{[R]} \ge X_0^{[G]} > 0$ as soon as $t < \tau_\partial := \inf\{t \ge 0\ ;\ X_0^{[G]}(t) = 0\}$. □

7.5.3 Proof of Lemma 7.5.4

For $i \le d$, denote $\widetilde{X}_i := X_i^{[G]} + X_{[R]} \times X_i^{[R]}$. We deduce the system of equations it satisfies from the Itô formula. Note that the martingale parts of $X_{[R]}$ and $X^{[R]}$ have a zero covariation. The bounded variation term in the equation of $\widetilde{X}_i$ is thus:

$$\begin{aligned} &\alpha(M_1^{(k)}(t) - i\wedge k)\, X_i^{[G]}(t) + \lambda\,(X_{i-1}^{[G]}(t) - \mathbf{1}_{\{i\ne d\}}\, X_i^{[G]}(t)) \\ &\quad + \alpha\,(M_1^{(k)}(t) - k)\, X_{[R]}(t)\, X_i^{[R]} + X_{[R]}(t)\,\lambda\,(X_{i-1}^{[R]}(t) - \mathbf{1}_{\{i\ne d\}}\, X_i^{[R]}(t)) \\ &= \alpha(M_1^{(k)}(t) - i\wedge k)\, \widetilde{X}_i(t) + +\lambda\,(\widetilde{X}_{i-1}(t) - \mathbf{1}_{\{i\ne d\}}\, \widetilde{X}_i(t)). \end{aligned}$$

For the martingale part, we find:

$$d\mathcal{M}_i(t) := \sqrt{X_i^{[G]}(t)}\, dW_t^{[G],i} - X_i^{[G]}(t)\, dW_t$$
$$+ \sqrt{X_{[R]}(t)} \left[\sqrt{X_i^{[R]}(t)}\, dW_t^{[R],i} - X_i^{[R]}(t)\, dW_t^{[R]}\right] + X_i^{[R]}(t)\ \left[\sqrt{X_{[R]}(t)}\, dW_t^{[R]} - X_{[R]}(t)\, dW_t\right]$$
$$= \sqrt{X_i^{[G]}(t)}\, dW_t^{[G],i} + \sqrt{X_{[R]}(t)\, X_i^{[R]}(t)}\, dW_t^{[R],i} - (X_i^{[G]}(t) + X_{[R]}(t)\, X_i^{[R]}(t))\, dW_t$$

Here, we thus define the family $\widetilde{W}_i$ by

$$W_t := \sum_i \int_0^t \sqrt{X_i^{[G]}(s)} dW_s^{[G],i} + \int_0^t \sqrt{X[R](s)} dW_s^{[R]}$$

Then, (X_i) indeed defines solutions to (7.4):$(S^{(k,d)})$ since for any $i \leq d$ and decomposition $y = y^r + y^n$, we have: $x_{i-1}^r - \mathbf{1}_{\{i \neq d\}} x_i^r(t) + x_{i-1}^n - \mathbf{1}_{\{i \neq d\}} x_i^n(t) = x_{i-1} - \mathbf{1}_{\{i \neq d\}} x_i(t)$. In particular, we have a.s. $\sum_j X_j^{[G]} + X_j^{[R]} = 1$. □

With the tools for the two transforms given in the last two subsections, we are now in position to prove Theorem 7.2 and Theorem 7.3 in the next two subsections.

7.6 Proof of Theorem 7.2

We recall the definition of $\mathcal{D}_\ell$:

$$\mathcal{D}_\ell := \{x \in \mathcal{X}_d\ ;\ M_3(x) \leq \ell\ ,\ x_0 \in (\,(3\ell)^{-1}, 1 - (3\ell)^{-1}\,)\}$$

The mixing will be achieved in two steps, with each step being completed after a time-interval of length t_M. t_M is arbitrarily taken as $t_M = 1$. We will exploit upper-bounds of the third moments given in Propositions 7.5.5 and 7.4.5 for the specific case where $k = 1$. In this proof, considering second moments instead of third ones would have been sufficient. Yet, estimates of the third moment shall be required for Theorem 7.3 and we wish to emphasize the similarity between these proofs.

The proof of Theorem 7.2 is concluded by exploiting the two following lemmas, that we shall prove as first steps. For certain $m > \ell$ and $y < 1/(2\ell)$ to be fixed later, and noting $M_3^{[F]} = M_3/(1 - X_0)$, let:

$$\tau_m^{[F],3} := \inf\{t \geq 0\ ;\ M_3(t) \geq m\,(1 - X_0(t))\}\ ,\ T_y^0 := \inf\{t \geq 0\ ;\ X_0(t) \notin (y, 1 - y)\} < \tau_\partial.$$

Lemma 7.6.1. *For any $\ell \geq 1$ and $t_M > 0$, there exists $m > \ell$ such that for any $y \in (0, 1/2\,\ell)$ there exists $c > 0$ such that for any $x \in \mathcal{D}_\ell$:*

$$\mathbb{P}_x^{(d)}(X_0(t_M) \in dx_0\ ;\ t_M < \tau_m^{[F],3} \wedge T_y^0) \geq c\,\mathbf{1}_{\{x_0 \in (2\,y, 1-2\,y)\}}\, dx_0.$$

For the following lemma, we base ourselves on the splitting presented in subsection 7.5 with $k = 1$. With these definitions of $X^{[G]}$ and $X^{[R]}$, we also denote:

$$\tau_\partial^{[R]} := \inf\{t \geq 0\ ;\ X_{[R]}(t) = 0\}$$
$$\tau_{m_G}^{[G],3} := \inf\{t \geq 0\ ;\ M_3^{[G]}(t) \geq m_G\}\ ,\qquad \tau_{m_R}^{[R],3} := \inf\{t \geq 0\ ;\ M_3^{[R]}(t) \geq m_R\},$$
$$\text{where } M_3^{[G]} := \textstyle\sum_{i \geq 0} i^3\, X_i^{[G]}\ ,\quad M_3^{[R]} := \textstyle\sum_{i \geq 0} i^3\, X_i^{[R]}.$$

Moreover, recalling that $\mathcal{F}^{[G]}$ is the filtration generated by the family $(W^{[G],i}, i \geq 0\ ;\ W^{[R]})$, the event $\{\tau_\partial^{[R]} \leq t_M < \tau_{m_G}^{[G],3} \wedge \tau^0\}$ is $\mathcal{F}^{[G]}$-measurable.

Lemma 7.6.2. *For any $t_M > 0$,there exists $m_G, c > 0$, $y \in (0, 1/10)$, $y' \in (0, y)$, such that for any $x \in \mathcal{X}_d$ such that $x_0 \in (1-3\,y, 1-2\,y)$, with $\tau^0 := \inf\{t \geq 0\ ;\ X_0(t) \notin (1/10, 1-y')\}$:*

$$\mathbb{P}^{(1,d)}_x(\tau^{[R]}_\partial \leq t_M < \tau^{[G],3}_{m_G} \wedge \tau^0) \geq c.$$

Thanks to these two Lemmas and Lemma 7.5.5, we shall be able to prove Theorem 7.2 with the following definition of ζ. In this formula, the values of y and m_G are deduced thanks to Lemma 7.6.2 with the (arbitrary) choice $t_M := 1$.

$$\zeta(dx) := \tfrac{1}{y}\int_{1-3\,y}^{1-2y} dx_0\, \mathbb{P}^{(1,d)}_{\bar{x_0}}(X(t_M) \in dx \,\big|\, \tau^{[R]}_\partial \leq t_M < \tau^{[G],3}_{m_G} \wedge \tau^0), \qquad (7.22)$$
$$\text{where } \bar{x_0} := x_0\delta_0 + (1-x_0)\,\delta_1.$$

Remark 7.6.3. • *As claimed in Subsection 7.2 while exploiting Lemma 7.1.2, the constraint $t_M < \tau_0$ indeed ensures that $X_0 \geq 1/10$, ζ-a.s. The constant $1/10$ is arbitrary chosen for convenience.*

• *For simplicity, we shall apply the cutting and the splitting for $k = 1$. The proof of Theorem 7.3 will exploit a generalization of this result for k large, with the first step ensuring a lower-bounded density of $(X_i(t_M)\ ;\ i \leq k-1)$ on any $\mathcal{Y}_k(y)$, where $\mathcal{Y}_k(y)$ is defined for $y \in (0, 1/k)$ by:*

$$\mathcal{Y}_k(y) := \{x \in \mathcal{X}_d\ ;\ (\textstyle\bigwedge_{i\leq k-1} x_i) \wedge (1 - \sum_{i\leq k-1} x_i) > y\}.$$

7.6.1 Step 1: proof of Lemma 7.6.1.

Under $\mathbb{P}^{(1,d)}$, X_0 is solution to the following autonomous equation:

$$dX_0(t) := \alpha X_0(t)\,(1 - X_0(t))\,dt - \lambda X_0(t)\,dt + \sqrt{X_0(t)\,(1 - X_0(t))}\,d\widehat{W}^0_t.$$

This property can be deduced as in Lemma 7.5.3, by identifying $\widehat{W}^0$ as a Brownian Motion that satisfies for any $t < \tau_\partial$:

$$\sqrt{X_0(t)\,(1 - X_0(t))}\,d\widehat{W}^0_t = \sqrt{X_0(t)}\,dW^0_t - X_0(t)\,dW_t.$$

It is then classical for such an elliptic diffusion that $X_0(t_M)$ has a lower-bounded density on $(2\,y, 1-2\,y)$ on the event $\{t_M < T^0_y\}$, uniformly over any initial condition such that $x_0 \in (1/\ell, 1-1/\ell)$.

We apply Lemma 7.5.5 with splitting generated for $k = 1$. Thus, we can choose a certain $m > 0$ such that for any $x \in \mathcal{D}_\ell$:

$$\mathbb{P}^{(1,d)}_x(t_M < \tau^{[R],3}_m \,\big|\, \mathcal{F}^{[G]}) \geq 1/2. \qquad (7.23)$$

Note that $\widehat{W}^0$ is clearly $\mathcal{F}^{[G]}$-measurable

On the event $\{t_M < \tau^{[R],3}_m \wedge T^0_y\}$, one has for any $s < t_M$:

$$\sup\{X_{(1)}(s)\ ;\ M_1(s)\ ;\ R^{(1)}_1(s)\ ;\ R^{(1)}_2(s)\} \leq M_3(s) \leq m.$$

Applying Lemma 7.4.6, we thus deduce that there exists $C > 0$ only depending on m and t_M such that:

$$\mathbb{P}^{(d)}_x(X_0(t_M) \in dx_0\ ;\ t_M < \tau^{[R],3}_m \wedge T^0_y) \geq C\,\mathbb{P}^{(1,d)}_x(X_0(t_M) \in dx_0\ ;\ t_M < \tau^{[R],3}_m \wedge T^0_y).$$

Because of (7.23), this concludes the proof of Lemma 7.6.1. □

7.6.2 Step 2: proof of Lemma 7.6.2.

Thanks to the Harnack inequalities, recalling that the equation for $X_0 = X_0^{[G]}$ is autonomous, we can choose a certain $c_0 > 0$ such that for any $x \in \mathcal{X}$ satisfying $x_0 \in [1/2, 1]$:

$$\mathbb{P}_x^{(1,d)}(t_M < \tau_{1/10}) \geq c_0, \quad \text{where } \tau_{1/10} := \inf\{t \geq 0 \ ; \ X_0(t) \leq 1/10\}. \tag{7.24}$$

Likewise, $X_{(1)}$ being governed by an autonomous equation, we can choose a certain y sufficiently small such that for any $x \in \mathcal{X}$ satisfying $x_0 \in [1 - 2\,y, 1 - y]$ (i.e. $X_{(1)}(0) \in [y, 2\,y]$):

$$\mathbb{P}_x^{(1,d)}(t_M < \tau_\partial^{[R]}) \leq c_0/4. \tag{7.25}$$

Since the border 1 is an entrance boundary for X_0, cf e.g. Subsection 3.3.3 in [12], there exists $y' \in (0, y)$ (again independent of d because X_0 is autonomous under $\mathbb{P}^{(1,d)}$), such that for any $x \in \mathcal{X}$ satisfying $x_0 \in [1/2, 1 - y]$:

$$\mathbb{P}_x^{(1,d)}(t_M < \tau^0) \geq 3c_0/4. \tag{7.26}$$

By Proposition 7.4.4, there exists m_G sufficiently large such that for any $x \in \mathcal{X}$:

$$\mathbb{P}_x^{(1,d)}(\tau_{m_G}^{[G],3} \leq t_M) = \mathbb{P}_{\bar{x}_0}^{(1,d)}(\tau_{m_G}^{[G],3} \leq t_M) \leq c_0/4, \tag{7.27}$$

where we exploited that $M_3(\bar{x}_0) \leq 1$ because $\bar{x}_0$ has support on $\{0, 1\}$.

Combining the inequalities (7.26), (7.25), (7.27), we obtain the following inequality:

$$\mathbb{P}_x^{(1,d)}(\tau_\partial^{[R]} \leq t_M < \tau_{m_G}^{[G],3} \wedge \tau^0) \geq c_0/2.$$

This concludes the proof of Lemma 7.6.2. □

7.6.3 Step 3: conclusion of the proof of Theorem 7.2.

We first define $m_G, c_G > 0$, $y \in (0, 1/10)$, $y' \in (0, y)$ thanks to Lemma 7.6.2 such that for any $x \in \mathcal{X}_d$ such that $x_0 \in (1 - 3\,y, 1 - 2\,y)$:

$$\mathbb{P}_x^{(1,d)}(\tau_\partial^{[R]} \leq t_M < \tau_{m_G}^{[G],3} \wedge \tau^0) \geq c_G. \tag{7.28}$$

Given $\ell \geq 1$, we then define $m_F, c_\ell > 0$ such that for any $x \in \mathcal{D}_\ell$:

$$\mathbb{P}_x^{(d)}(X_0(t_M) \in dx_0 \ ; \ t_M < \tau_{m_F}^{[F],3} \wedge T_y^0) \geq c_\ell \, \mathbf{1}_{\{x_0 \in (1-3\,y, 1-2\,y)\}} \, dx_0. \tag{7.29}$$

We define also $m_R > 0$ thanks to Lemma 7.5.5, so that:

$$\mathbb{P}_x^{(1,d)}(t_M < \tau_{m_R}^{[R],3} \,\big|\, \mathcal{F}^{[G]}) \geq 1/2, \tag{7.30}$$

provided x satisfies $M_3^{[R]}(0) \leq 2m_F$ (so in particular when $x_0 \geq 1/2$ and $M_3(x) \leq m_F$).

By choosing L sufficiently large, we ensure $L \geq m_F \vee (m_G + m_R) \vee (1/y')$. Recalling that $T_{\mathcal{D}_L}$ denotes the exit time out of $\mathcal{D}_L$, and that $M_3 \leq M_3^{[F]}$ (here $k = 1$), it proves that $\{t_M < \tau_{m_F}^{[F],3} \wedge T_y^0\} \subset \{t_M < T_{\mathcal{D}_L}\}$. Likewise, since $M_3 \leq M_3^{[G]} + M_3^{[R]}$, $\{t_M < \tau_{m_G}^{[G],3} \wedge \tau_{m_R}^{[R],3} \wedge \tau_0\} \subset \{t_M < T_{\mathcal{D}_L}\}$.

By the Markov property:

$$\mathbb{P}_x^{(d)}(X(2\,t_M) \in dx' \ ; \ 2\,t_M < T_{\mathcal{D}_L}) \geq \int_{\mathcal{X}^{(d)}} \nu_x(dz) \mathbb{P}_z^{(d)}(X(t_M) \in dx' \ ; \ t_M < T_{\mathcal{D}_L}),$$

where $\nu_x(dz) := \mathbb{P}_x^{(d)}(X(t_M) \in dz \ ; \ t_M < \tau_{m_F}^{[F],3} \wedge T_y^0)$.

The previous r.h.s. is itself lower-bounded by

$$\int_{\mathcal{X}^{(d)}} \nu_x(dz) \mathbf{1}_{\{z_0 \in (1-3y, 1-2y)\}} \mathbb{P}^{(d)}_z (X(t_M) \in dx' \ ; \ \tau^{[R]}_\partial \leq t_M < \tau^{[G],3}_{m_G} \wedge \tau^0 \wedge \tau^{[R],3}_{m_R}).$$

Note that on the event $\{\tau^{[R]}_\partial \leq t_M\}$, we know thanks to Proposition 7.4.3 that $X(t_M) = X^{[G]}(t_M)$. Both $X^{[G]}(t_M)$ and $\{\tau^{[R]}_\partial \leq t_M\}$ are $\mathcal{F}^{[G]}$-measurable. Moreover, on the event $\{t_M < \tau^{[G],3}_{m_G} \wedge \tau^0 \wedge \tau^{[R],3}_{m_R}\}$:

$$\sup_{s \leq t_M} M_3(s) \leq \sup_{s \leq t_M} M^{[G]}_3(s) + \sup_{s \leq t_M \wedge \tau^{[R]}_\partial} M^{[R]}_3(s) \leq m_G + m_R.$$

Using Lemma 7.4.6 with a uniform upper-bound on the exponential martingale (with $k = 1$) on the event $\{\tau^{[R]}_\partial \leq t_M < \tau^{[G],3}_{m_G} \wedge \tau^0 \wedge \tau^{[R],3}_{m_R}\}$, we deduce that there exists $c_E > 0$ such that:

$$\begin{aligned}
&\mathbb{P}^{(d)}_x (X(2\, t_M) \in dx' \ ; \ 2\, t_M < T_{\mathcal{D}_L}) \\
&\geq c_E \int_{\mathcal{X}^{(d)}} \nu_x(dz) \mathbf{1}_{\{z_0 \in (1-3y, 1-2y)\}} \mathbb{E}^{(1,d)}_z \big[\mathbb{P}^{(1,d)}_z (t_M < \tau^{[R],3}_{m_R} \,\big|\, \mathcal{F}^{[G]}) \ ; \ X^{[G]}(t_M) \in dx', \\
&\qquad\qquad \tau^{[R]}_\partial \leq t_M < \tau^{[G],3}_{m_G} \wedge \tau^0 \big] \\
&\geq (c_E/2) \int_{\mathcal{X}^{(d)}} \nu_x(dz) \mathbf{1}_{\{z_0 \in (1-3y, 1-2y)\}} \mathbb{E}^{(1,d)}_{\bar{z_0}} \big[X^{[G]}(t_M) \in dx' \, , \, \tau^{[R]}_\partial \leq t_M < \tau^{[G],3}_{m_G} \wedge \tau^0 \big]
\end{aligned}$$

where we exploited that both $X^{[G]}$ and the event $\{\tau^{[R]}_\partial \leq t_M < \tau^{[G],3}_{m_G} \wedge \tau^0\}$ are $\mathcal{F}^{[G]}$-measurable, that they depend on z only through z_0, and (7.30). We thus have the same laws for initial condition z and $\bar{z_0}$. Thanks to (7.28) and (7.29), this concludes the proof of Theorem 7.2.

□

7.7 Proof of Theorem 7.3

7.7.1 Choice of the parameters

We give ourselves $\rho, m_3^0, \eta^0, y^0 > 0$ and fix arbitrarily $t_\veebar = 1$.

In the first time-interval of length $t_H \leq t_\veebar$ to be fixed below, we couple the first k coordinates between two processes with different initial conditions. Then, we aggregate the last coordinates in $X_{[R]}$ and impose that $X_{[R]}$ get extinct in the next time-interval of length $t_\veebar$ while $X^{[N]}$ evolves independently.

In view of Lemma 7.5.3, we wish to control extinction of an upper-bound of $X_{[R]}$ of the form:

$$\text{for } t \geq t_\veebar, \quad dZ_t := \sqrt{Z_t\,(1 - Z_t)}\, dW_t \, , \quad Z_{t_\veebar} = z, \tag{7.31}$$

where W is a Brownian Motion. Namely, for any $\epsilon > 0$, we choose z in such a way that:

$$\mathbb{P}_z (t_\veebar/2 < \tau^Z_\partial) \leq \epsilon, \quad \text{ where } \tau^Z_\partial := \inf\{t \geq 0 \ ; \ Z_t = 0\}. \tag{7.32}$$

Now, with the constants C_G, C_M appearing in Proposition 7.4.1 for $\delta = 3$ and $t = t_\veebar$, we choose η such that:

$$\eta < (z/C_M) \wedge (\epsilon/(C_G)^2) \wedge \eta^0. \tag{7.33}$$

Given some $\rho > 0$, we can choose thanks to Theorem 7.1 certain $m_3 \geq m_3^0$ and $y < y^0$ such that:

$$\forall\, x \in \mathcal{X}_d, \quad \mathbb{E}_x \left(\exp[\rho\, (\tau_\partial \wedge \tau_E)] \right) \leq 2, \tag{7.34}$$

where we recall:

$$E := \{x \in \mathcal{X}_d \ ; \ M_3(x) \leq m_3 \, , \, \forall\, j \leq \lfloor m_3/\eta \rfloor + 1, \ x_j \geq y\}. \tag{7.35}$$

Remark 7.7.1. *The set E^0 defined through m_3^0, η^0, y^0 is included in this set E (depending on ϵ). Proving the inequalities for any $x, x_\zeta \in E$ clearly implies them for any $x, x_\zeta \in E^0$, as required.*

Exploiting Proposition 7.4.1 with $m' = C_M m_3$, we deduce that for any $x \in E$:

$$\mathbb{P}_x(\tau^3_{m'} \leq t_{\veebar}) \leq \frac{C_G m_3}{k} \leq \epsilon. \tag{7.36}$$

Recalling that $k := \lfloor m_3/\eta \rfloor + 1$ and $\eta \leq z/C_M$, we deduce that on the event $\{t_{\veebar} < \tau^3_{m'}\}$, for any $t_H < t_{\veebar}$:

$$X_{(k)}(t_H) \leq \frac{C_M \times m_3}{k^3} \leq z. \tag{7.37}$$

This provides the initialisation of $X_{[R]}$, that we couple to the original process from time t_H onward thanks to the Markov property. Notably, $X_{[R]}$ is upper-bounded by Z (see definition (7.31)).

Note that for any $x \in E$, recalling (7.5):

$$\pi_k(x) \in \mathcal{Y}(y) \qquad \text{where } \mathcal{Y}(y) := \{x \in \mathcal{X}_k \ ; \ (\bigwedge_{i[\![0,k]\!]} x_i) > y\}.$$

on which the diffusion term for $(S^{(k)})$, i.e. (1.1):$(S^{(d)})$ with d replaced by k, is uniformly elliptic. In practice, we need a bit more space for the Harnack inequality to hold, so that we consider the exit time:

$$T_H := \inf\{t \geq 0 \ ; \ \pi_k(X(t)) \notin \mathcal{Y}(y/2)\} < \tau_\partial. \tag{7.38}$$

The probability of such an escape is required to be very small, uniformly in $x \in E$, according to the following lemma.

Lemma 7.7.2. *With the above definitions, and $\mathbb{P}^{(k)}$ the law of the system given by $(S^{(k)})$:*

$$\sup_{x \in E} \mathbb{P}^{(k)}_{\pi_k(x)}(T_H \leq t_H) \to 0 \text{ as } t_H \to 0.$$

Since the system $(S^{(k)})$ is uniformly elliptic on $\mathcal{Y}(y/3)$, and recalling Proposition 7.4.3, Lemma 7.7.2 is easily deduced from classical results as for instance Proposition V.2.3 in [4].

Thanks to Proposition 7.4.3, we can thus choose $t_H \leq t_{\veebar}/2$ sufficiently small such that:

$$\sup_{x \in E} \mathbb{P}^{(k,d)}_x(T_H \leq t_H) \leq \epsilon. \tag{7.39}$$

7.7.2 Definition of U_A with a control of exceptional events

In this context, with the splitting starting at time t_H, $\tau^{[R]}_\partial := \inf\{t \geq t_H \ ; \ X_{[R]}(t) = 0\}$.

In view of Theorem 7.3, we define $U_A := t_{\veebar}$ on the event $\{t_H < T_H\} \cap \{t_{\veebar} < \tau_3^{m'} \wedge \tau_\partial\} \cap \{\tau^{[R]}_\partial \leq t_{\veebar}\}$, and otherwise $U_A := \infty$.

Exploiting Proposition 7.4.1 and recalling the definition of E, m' and t_H, we deduce:

$$\mathbb{P}_x(U_A = \infty, t_{\veebar} < \tau_\partial) \leq 2\,\mathbb{P}^{(k,d)}_x(t_H \leq T_H) + 2\,\mathbb{P}^{(k,d)}_x(t_{\veebar} < \tau^{[R]}_\partial) + \mathbb{P}^{(d)}_x\left(\tau_3^{m'} \leq t_{\veebar}\right) \leq 5\,\epsilon.$$

For Theorem 7.3, it means that the threshold is obtained with ϵ' such that $\epsilon = \epsilon' \times \exp[-\rho\, t_{\veebar}]/5$. Since ϵ is freely chosen, so is ϵ'.

It is technical but elementary that this definition of U_A gives rise to a stopping time U_A^∞ extending it in the sense described in Theorem 7.3. Rigorously, in the notations of this Subsection 7.7, note that E^0, see (7.35), takes the place of the set E in Theorem 7.3, which makes little difference in the proof.

7.7.3 Comparison of densities

Since the problem is reduced to a finite dimensional one, the Harnack inequality states, as in Subsection 5.4, that there exists $C_H > 0$ such that:

$$\begin{aligned}
&\inf_{x\,;\ \pi_k(x)\in\mathcal{Y}(y)} \mathbb{P}_x^{(k,d)}(\pi_k(X_{t_H}) \in dx'\ ,\ t_H < T_H \wedge \tau_{m'}^{(k),3}) \\
&\qquad \geq C_H \sup_{x'\,;\ \pi_k(x')\in\mathcal{Y}(y)} \mathbb{P}_{x'}^{(k,d)}(\pi_k(X'_{t_H}) \in dx'\ ,\ t_H < \tau_\partial \wedge \tau_{m'}^{(k),3}), \qquad (7.40)\\
&\text{where } \tau_{m'}^{(k),3} := \inf\left\{t \geq 0\ ;\ \textstyle\sum_{\{i\geq 0\}} (i\wedge k)^3 X_i(t) \geq m'\right\}.
\end{aligned}$$

Let $x, x_\zeta \in E$. Thanks to Proposition 7.4.1:

$$\begin{aligned}
\mathbb{P}_x^{(d)}&(X_{U_A} \in dx'\ ,\ U_A < \infty) \leq 2\,\mathbb{P}_x^{(k,d)}(X_{t_\veebar} \in dx'\ ,\ U_A < \infty)\\
&= 2\mathbb{E}_x^{(k,d)}[\mathbb{P}_{X_{t_H}}^{(k,d)}(X^{[G]}(t_\veebar - t_H) \in dx'\ ;\ \tau_\partial^{[R]} \leq t_\veebar - t_H < \tau_\partial \wedge \tau_{m'}^3)\ ;\ t_H < T_H \wedge \tau_{m'}^3]\\
&\leq 2\mathbb{E}_x^{(k,d)}[\mathbb{P}_{X(t_H)}^{(k,d)}(X^{[G]}(t_\veebar - t_H) \in dx'\ ;\ \tau_\partial^{[R]} \leq t_\veebar - t_H < \tau_\partial \wedge \tau_{m'}^{[G],3})\ ;\ t_H < T_H \wedge \tau_{m'}^{(k),3}]\\
&\leq 2C_H\,\mathbb{E}_{x_\zeta}^{(k,d)}[\mathbb{P}_{X(t_H)}^{(k,d)}(X^{[G]}(t_\veebar - t_H) \in dx'\ ;\ \tau_\partial^{[R]} \leq t_\veebar - t_H < \tau_\partial \wedge \tau_{m'}^{[G],3})\ ;\ t_H < T_H \wedge \tau_{m'}^{(k),3}],
\end{aligned} \qquad (7.41)$$

because of (7.40) and Fact 7.5.1, noting that τ_∂ and $\tau_\partial^{[R]}$ are measurable with respect to $\sigma(X^{[G]})$.

To go back to $\mathbb{P}^{(d)}$, we shall exploit again Proposition 7.4.1. So we need to again ensure upper-bounds on the third moments for the last components for which we lost the information.

For the time-interval $[0, t_H]$, in order to exploit independence as much as possible, we shall exploit the representation given in Proposition 7.4.3. Since $x_\zeta \in E$, we have $M_3^{[F]}(0) \leq m_3/y$. From Proposition 7.4.5, we thus define m_H such that for any $x_\zeta \in E$:

$$\mathbb{P}_{x_\zeta}^{(k,d)}\left(\tau_{m_H}^{[F],3} \leq t_H \,\middle|\, \mathcal{F}^{(k)}\right) \leq 1/2. \qquad (7.42)$$

Note that $M_3^{[R]}$, as in Lemma 7.5.5, is initialized (at time t_H in this context) by $M_3^{[R]}(t_H) = M_3^{[F]}(X^{[R]}(t_H)) \leq m_H$. Depending on the context for the start of the splitting with $X^{[F]}$, the definition of $\tau_{m_R}^{[F],3}$ may be adapted accordingly.

From Lemma 7.5.5, with $t = t_\veebar - t_H$, we then define m_R such that for any $x \in \mathcal{X}_d$ such that $M_3^{[R]}(x) \leq m_H$:

$$\mathbb{P}_x^{(k,d)}\left(\tau_{m_R}^{[R],3} \leq t_\veebar - t_H \,\middle|\, \mathcal{F}^{[G]}\right) \leq 1/2. \qquad (7.43)$$

From these results, we can come back to (7.41) and deduce first from (7.43) that, on the event $\{t_H < T_H \wedge \tau_{m'}^{(k),3}\}$:

$$\begin{aligned}
&\mathbb{P}_{X(t_H)}^{(k,d)}(X^{[G]}(t_\veebar - t_H) \in dx'\ ;\ \tau_\partial^{[R]} \wedge \tau_{m_R}^{[F],3} \leq t_\veebar - t_H < \tau_\partial)\\
&\qquad \geq (1/2)\times\mathbb{P}_{X(t_H)}^{(k,d)}(X^{[G]}(t_\veebar - t_H) \in dx'\ ;\ \tau_\partial^{[R]} \leq t_\veebar - t_H < \tau_\partial)
\end{aligned}$$

Then more generally, since $X^{[G]}$ is independent of what happens to $X^{[F]}$ in the time-interval $[0, t_H]$:

$$
\begin{aligned}
&\mathbb{E}^{(k,d)}_{x_\zeta}[\mathbb{P}^{(k,d)}_{X(t_H)}(X^{[G]}(t_\veebar - t_H) \in dx' \ ;\ \tau^{[R]}_\partial \le t_\veebar - t_H < \tau_\partial \wedge \tau^{[G],3}_{m'} \wedge \tau^{[F],3}_{m_R}) \\
&\qquad\qquad ;\ t_H < T_H \wedge \tau^{(k),3}_{m'} \wedge \tau^{[F],3}_{m_H}] \\
&\quad \ge (1/2)\, \mathbb{E}^{(k,d)}_{x_\zeta}[\mathbb{P}^{(k,d)}_{X(t_H)}(X^{[G]}(t_\veebar - t_H) \in dx' \ ;\ \tau^{[R]}_\partial \le t_\veebar - t_H < \tau_\partial \wedge \tau^{[G],3}_{m'}) \\
&\qquad\qquad ;\ t_H < T_H \wedge \tau^{(k),3}_{m'} \wedge \tau^{[F],3}_{m_H}] \\
&\quad \ge (1/4)\, \mathbb{E}^{(k,d)}_{x_\zeta}[\mathbb{P}^{(k,d)}_{X(t_H)}(X^{[G]}(t_\veebar - t_H) \in dx' \ ;\ \tau^{[R]}_\partial \le t_\veebar - t_H < \tau_\partial \wedge \tau^{[G],3}_{m'}) \ ;\ t_H < T_H \wedge \tau^{(k),3}_{m'}]
\end{aligned} \tag{7.44}
$$

The upper-bound is then simplified, with the Markov property and the fact that $M_3 \le M^{[G]}_3 + M^{[R]}_3$. Then, we exploit again Proposition 7.4.1 to state that there exists $C_G > 0$, independent of x, such that:

$$
\begin{aligned}
&\mathbb{E}^{(k,d)}_{x_\zeta}[\mathbb{P}^{(k,d)}_{X(t_H)}(X^{[G]}(t_\veebar - t_H) \in dx' \ ;\ \tau^{[R]}_\partial \le t_\veebar - t_H < \tau_\partial \wedge \tau^{[G],3}_{m'} \wedge \tau^{[F],3}_{m_R}) \\
&\qquad\qquad ;\ t_H < T_H \wedge \tau^{(k),3}_{m'} \wedge \tau^{[F],3}_{m_H}] \\
&\quad \le \mathbb{P}^{(k,d)}_{x_\zeta}(X(t_\veebar) \in dx' \ ;\ t_\veebar < \tau_\partial \wedge \tau^{3}_{m'+m_R}) \\
&\quad \le C_G\, \mathbb{P}^{(d)}_{x_\zeta}(X(t_\veebar) \in dx' \ ;\ t_\veebar < \tau_\partial).
\end{aligned} \tag{7.45}
$$

Combining (7.41), (7.44) and (7.45) yields that, with $C := 8\, C_H\, C_G > 0$, for any $x, x_\zeta \in E$:

$$
\mathbb{P}^{(d)}_x(X_{U_A} \in dx' \, ,\ U_A < \infty) \le C_G\, \mathbb{P}^{(d)}_{x_\zeta}(X(t_\veebar) \in dx' \ ;\ t_\veebar < \tau_\partial).
$$

This concludes the proof of Theorem 7.3. □

The proof of Theorem 2.3 is then finally concluded.

References

[1] Audiffren, J., Pardoux, E.; Müller's ratchet clicks in finite time, Stoch. Proc. and their Appl., V.123, pp. 2370–2397 (2013)

[2] Bansaye, V., Cloez, B., Gabriel, P., Marguet, A.; A non-conservative Harris' ergodic theorem, ArXiv: 1903.03946 (2019)

[3] Bansaye, V., Méléard, S.: Stochastic Models for Structured Populations, Scaling Limits and Long Time Behavior; Springer Switzerland (2015)

[4] Bass, R. F.; Diffusions and Elliptic Operators, Probab. and Its Applications, Springer, New York (1998)

[5] Champagnat, N., Villemonais, D.; Exponential convergence to quasi-stationary distribution and Q-process, Probab. Theory Relat. Fields, V. 164, pp. 243–283 (2016)

[6] Champagnat, N., Villemonais, D.; Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes, Stoch. Proc. Appl., V.135, pp.51-74 (2021)

[7] Champagnat, N., Villemonais, D.; General criteria for the study of quasi-stationarity, preprint on ArXiv: arxiv.org/abs/1712.08092v1 (2017)

[8] Collet, P., Martínez, S., San Martín, J.; Quasi-Stationary Distributions, Probab. and Its Appl., Springer, Berlin Heidelberg (2013)

[9] Di Tella, P., On the Predictable Representation Property of Martingales Associated with Levy Processes, PhD thesis (2013)

[10] Etheridge, A., Pfaffelhuber, P., Wakolbinger, A.; How often does the ratchet click? Facts, heuristics, asymptotics, in: Trends in Stochastic Analysis, in: London Math. Soc. Lecture Notes Series, Cambridge Univ. Press (2009)

[11] Haigh, J.; The accumulation of deleterious genes in a population — Müller's ratchet, Theoret. Population Biol. 14, 251–267 (1978)

[12] Jilani Ben Naouara, N.J.B., Trabelsi, F.; Boundary classification and simulation of one-dimensional diffusion processes, Int. J. Mathematics in Operational Research, V.11, N.1 (2017)

[13] Keeling, P., Palmer, J.; Horizontal gene transfer in eukaryotic evolution, Nat. Rev. Genet. 9 605–618 (2008)

[14] Kingman, J.F.C.; The coalescent, Stochastic Process. Appl. 13, 235-248 (1982)

[15] Lunardi., A.; Analytic semigroups and optimal regularity in parabolic problems, v. 16 of Progress in Nonlinear Diff. Eq. and their Appl.; Birkhauser Verlag, Basel (1995)

[16] Maynard Smith, J.; The Evolution of Sex, Cambridge University Press (1978)

[17] Metzger, J.J., Eule, S.; Distribution of the Fittest Individuals and the Rate of Müller's Ratchet in a Model with Overlapping Generations. PLoS ComputBiol V9, N11 (2013)

[18] Moran, M., Jarvik, T.; Lateral transfer of genes from fungi underlies carotenoid production in aphids, Science 328, 624–627 (2010)

[19] Müller, H. J.; The relation of recombination to mutational advance, Mutat. Res., V.1, pp.2-9 (1996)

[20] Ochman, H., Lawrence, J., Groisman, E.; Lateral gene transfer and the nature of bacterial innovation, Nature 405, 299–304 (2000)

[21] Pardoux, E., Răşcanu, A.; Stochastic Differential Equations, Backward SDEs, Partial Differential Equations. Stochastic Modelling and Applied Probability, 69, Springer (2014)

[22] Pfaffelhuber, P., Staab, P.R., Wakolbinger, A.; Müller's ratchet with compensatory mutations, Ann. Appl. Probab., V.22, N.5, pp.2108–2132 (2012)

[23] Szamecz, B., Boross, G. and al.; The Genomic Landscape of Compensatory Evolution PLoS Biology, V.12, N.8:e1001935 (2014)

[24] Tettelin, H., Riley, D., Cattuto, C., Medini, D.; Comparative genomics: the bacterial pan-genome, Curr. Opin. Microbiol. V.11, pp.472–477 (2008)

[25] Velleret, A.; Unique Quasi-Stationary Distribution, with a possibly stabilizing extinction, preprint available on ArXiv at: `https://arxiv.org/abs/1802.02409`

[26] Velleret, A.; Exponential quasi-ergodicity for processes with discontinuous trajectories, preprint available on ArXiv at: `https://arxiv.org/abs/1902.01441`

[27] Velleret, A.; Adaptation of a population to a changing environment under the light of quasi-stationarity; preprint available on ArXiv at: `https://arxiv.org/abs/1903.10165`

[28] Velleret, A.; Two level natural selection with a quasi-stationarity approach; preprint on ArXiv: `https://arxiv.org/abs/1903.10161`