ESAIM: PS 27 (2023) 867–912
https://doi.org/10.1051/ps/2023016

ESAIM: Probability and Statistics
www.esaim-ps.org

EXPONENTIAL QUASI-ERGODICITY FOR PROCESSES WITH DISCONTINUOUS TRAJECTORIES

Aurélien Velleret*

Abstract. This paper tackles the issue of establishing an upper-bound on the asymptotic ratio of survival probabilities between two different initial conditions, asymptotically in time for a given Markov process with extinction. Such a comparison is a crucial step in recent techniques for proving exponential convergence to a quasi-stationary distribution. We introduce a weak form of the Harnack's inequality as the essential ingredient for such a comparison. This property is actually a consequence of the convergence property that we intend to prove. Its complexity appears as the price to pay for the level of flexibility required by our applications, notably for processes with jumps on a multidimensional state-space. We show in our illustrations how simply and efficiently it can be used nonetheless. As illustrations, we consider two continuous-time processes on $\mathbb{R}^d$ that do not satisfy the classical Harnack's inequality, even in a local version. The first one is a piecewise deterministic process while the second is a pure jump process with restrictions on the directions of its jumps.

Mathematics Subject Classification. 37A30, 28D10, 60J25 and 92D15.

Received September 13, 2022. Accepted September 11, 2023.

1. Introduction

1.1. General presentation

This work is concerned with the long time behavior of quite general strong Markov processes, conditionally upon the fact that this process has not been absorbed in some "cemetery state" (*i.e.* that it is not "extinct"). The eventual interest is on the analogous of stationary distributions when such a conditioning is taken into account, namely quasi-stationary distributions (QSD).

In the aftermath of recent works by Champagnat and Villemonais, notably in [19], we are interested in highlighting key properties that ensure convergence results at exponential rate towards QSD. While the approach was initiated in the framework of a convergence in total variation that is uniform over the initial condition, how to deal with heterogeneity in the initial condition has already been the concern of further studies ([6, 20, 21, 24, 63]).

These works are inspired by the Harris recurrence techniques that is exploited for the proof of convergence towards a stationary distribution for conservative semi-groups (*cf.* [47], Sect. 2 or [51], Chap. 13, 15 for more details). The core of these techniques is a Doeblin minoration condition where the density of the marginal law is lower-bounded uniformly over some initial conditions. A Lyapunov criterion is then exploited to deal with the

Keywords and phrases: Continuous-time and continuous-space Markov process, jumps, quasi-stationary distribution, survival capacity, Q-process, Harris recurrence.

Université Paris-Saclay, INRAE, MaIAGE, Domaine de Vilvert, 78350 Jouy-en-Josas, France.

* Corresponding author: aurelien.velleret@nsup.org

heterogeneity in the initial conditions. Thanks to these two properties, we may exhibit a coupling procedure that make trajectories starting from different initial conditions coincide beyond a specific time-point [51], Section 13.2. While refinements of these two properties can be identified in [6, 24, 63], another key property is introduced for the generalization to non-conservative semi-groups that compare survival between different initial conditions.

This kind of property has a different nature as the others in that it is expressed asymptotically in large time. In this paper, as a representative of those properties, we focus on the following property stated in Equation (1.1). It is more simply expressed in that it is stated uniformly over the initial condition and the denominator is taken as a single reference measure $\zeta \in \mathcal{M}_1(\mathbb{X})$.

$$\limsup_{t\to\infty} \sup_{x\in\mathbb{X}} \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_\zeta(t < \tau_\partial)} < \infty. \tag{1.1}$$

In the expression, τ_∂ denotes the extinction time and the initial condition of the process is written in subscript. Although ζ is meant to relate to small sets, no property on ζ should be needed for the proof of (1.1). (1.1) is the archetype of the comparison properties of asymptotic survival probability, that are introduced in the abstract.

The derivation of a property like (1.1) is well understood in the cases of discrete-space processes (*cf.* [19], Proof of Thm. 4.1, p. 14), and processes that satisfy either two-sided estimates (*cf.* [17], Sect. 2), gradient estimates (*cf.* [17], Sect. 3) or the Harnack's inequality (given in [23], Sect. 4, step 4 of the proof, or more generally stated in [63], Sect. 4.2.3). In the case of many processes involving jumps, and notably in multi-dimensional settings, none of these three estimates are generally applicable: singularities due to the absence of jump (or their accumulation) in deterministic time-intervals interfer with direct comparison of densities between marginal laws starting from different initial conditions (see the paragraph next to our key property in Sect. 2.2.2). The other techniques that we know of do not appear easy to interpret for the level generality we aim at or require a coupling estimate that is much more constrained than what we expect to have (notably Lem. 5 in [37]).

This is why we worked at a weakened form of the Harnack's inequality with more flexibility, (that we call "almost perfect harvest", see Sect. 1.3.3), that would still suffice to obtain (1.1), in combination with the other key properties requested for exponential convergence towards QSD. The aim is to capture a much larger range of processes. In practice, the statement of our main result (see Thm. 2.3) relies on the other key properties that we proposed in [63].

The fact that the approach in [63] is specifically trajectorial gives more insight into relevant estimates, which contributed to the choice of this framework. The fact that the proofs in [63] were designed with the current article already in mind also played a role in this choice. Potential generalizations are nevertheless discussed in Sect. 1.4.2.

As a direct consequence, we present in Sect. 2.3 the new set of key conditions that is proved sufficient for the general results of [63] to be deduced. It implies not only the existence and uniqueness of the QSD, but also several results of exponential convergence.

The paper is organized as follows. In the next Sect. 1.2, we present two illustrative models that contribute to motivate the new conditions introduced in this article. With these two practical issues in mind, a concise presentation of the results can more readily be addressed in Sect. 1.3. There, we first specify the convergence objective, then how the contribution of [63] leads to proving such convergence results (see (1.2)) and finally how the present contribution eases the verification of one of the main conditions of [63], Section 2. In Sect. 1.4, we then place our method in the context of existing results on quasi-stationarity.

In complement to this concise overview of our results, a more detailed description is provided in Section 2. The focus is then on the main contribution of the article, namely the proposed key properties implying (1.1). The relation to the set of criteria given in [63] and the applicability of our combined results in practice is thus postponed to Sect. 2.3. Yet, it is also convenient to introduce or recall concise and precise notations at use, which we do first in Sect. 2.1. Sect. 2.2 is then devoted to the main contribution of this article, namely the derivation of Property (1.1) thanks to our new criterion. Sect. 2.3 is thus more generally concerned on the implications for the proof of exponential convergence to a unique QSD. These implications are complemented in Sect. 2.3.3 by reciprocal statements motivating the generality of our criteria and a property of uniform convergence in the parameter of a specific localization procedure (considering a restriction of the state space).

We detail our proofs in Section 3. Sections 4 and 5 are finally dedicated to each of our applications, which are to be introduced in their simpler form in the next paragraph.

1.2. Focal illustrations

The two illustrations of the current paper are meant to help the reader get insight on the adaptability of our new criterion of "almost perfect harvest". They fall into the class of piecewise deterministic Markov processes, the second one being actually a pure jump process. For the broader view of the applications we have in mind, we refer to the Discussion Section, and more precisely Sect. 6.2.

We first consider the following process on $\mathbb{R}^d$ for $d \geq 1$:

$$X_t = x - v\, t\mathbf{e}_1 + \sum_{\{i \leq N_t\}} W_i,$$

where $x \in \mathbb{R}^d$ is the deterministic initial condition, $v > 0$, $\mathbf{e}_1$ the first unit vector, N_t a Poisson process with intensity 1 and the family $(W_i)_{i\in\mathbb{Z}_+}$ is made up of i.i.d. normal variables with mean 0 and covariance matrix $\sigma^2 I_d$. This process gets extinct at a state-dependant rate. The extinction event, whose time is denoted τ_∂, is occurring at rate $t \mapsto \rho(X_t)$, where $\rho(x) := \|x\|^2$ (with the euclidian norm).

Then, as stated in Theorem 4.1, there exists a unique quasi-stationary distribution associated to this dynamics and it is the only attractor of the conditioned marginals as t goes to infinity.

While the one-dimensional case is treated in [25], with a stronger result than in [32], the convergence result is new for the multidimensional setting. For more details on the interpretation of this model and some generalization of its parameters, we refer to Section 4.

The second illustration, presented in Section 5, concerns a pure jump process on $\mathbb{R}^d$, for $d \geq 2$:

$$X_t := x + \sum_{\{i \leq N_t\}} \sigma W_i \cdot \mathbf{e}_{D_i}.$$

In this formula, x is the initial condition, N_t a standard Poisson process on $\mathbb{Z}_+$, while, for any $i \geq 1$, W_i is a standard 1d. normal random variable and D_i is uniform over $[\![1, d]\!]$. Moreover, all these random variables are independent from each others. Similarly as the previous example, this process gets extinct at a state-dependent rate given by $\rho_e : x \in \mathbb{R}^d \mapsto \|x\|_\infty^2$, where $\|x\|_\infty := \sup_{\{i\leq d\}} |x_i|$.

Then, as stated in Theorem 5.1, $\sigma \leq 1/8$ is a sufficient condition (uniform in the dimension d) for the existence of a unique quasi-stationary distribution associated to this dynamics.

Note that jumps are restricted to happen along the vectors of an orthonormal basis $(\mathbf{e}_1, ..., \mathbf{e}_d)$. One jump is thus unsufficient to erase the singularities with respect to the Lebesgue measure on $\mathbb{R}^d$. To our knowledge, there is no other result of quasi-stationarity for processes with such restrictions on the jumps. The result presented in [2] for the pure-jump case appears the closest to our. In this multidimensional setting however, the jump effect is assumed to have a density with respect to the Lebesgue measure on $\mathbb{R}^d$.

More details on the interpretation of this model and some generalization of its parameters are presented in Section 5.

1.3. Main implications of our result

1.3.1. Current statement of the convergence we aim at

For any Markov process $X = (X_t)_{t\in\mathbb{R}_+}$ with extinction time τ_∂, the property of quasi-stationary convergence that we aim to prove is stated in the current article as follows. The statement depends on a bounded function

h on the state space $\mathbb{X}$ and on a probability measure α on $\mathbb{X}$ that are uniquely defined by imposing that $\int_{\mathbb{X}} h(x)\alpha(\mathrm{d}x) = 1$ in addition to the property.

There exists $\lambda, C, \gamma > 0$ such that the following inequality holds for any $t \geq 0$ and any probability measure μ on $\mathbb{X}$:

$$\| \mathrm{e}^{\lambda t} \cdot \mathbb{P}_\mu \left(X_t \in \mathrm{d}x;\ t < \tau_\partial \right) - \alpha(\mathrm{d}x) \int_{\mathbb{X}} h(y)\mu(\mathrm{d}y) \|_{TV} \leq C \|\mu - \alpha\|_{TV} \cdot \mathrm{e}^{-\gamma\, t}. \tag{1.2}$$

The reader is refered to Definition 2.6 in Section 2.3.2 for a more precise statement and to Sect. 1.4.1 for the comparison with the result given in [63].

1.3.2. Main properties leading to the result following [63]

Up to minor adjustments justified in Sect. 3.2, it is the purpose of [63] that the proof of this property (1.2) can be deduced by the proofs of the following four properties. They depend on a probability measure ζ on $\mathbb{X}$ and on a subset E of $\mathbb{X}$, that have to be the same. The names are altered to ease the comparison with the key properties proposed in the current article. We refer to Sect. 2.2.1 for the intuition behind these properties.

($\overline{A0}$) "**Exhaustion of** $\mathbb{X}$": There exists a sequence $(\mathcal{D}_\ell)_{\ell \geq 1}$ of closed subsets of $\mathbb{X}$ such that for any $\ell \geq 1$, $\mathcal{D}_\ell$ is included in the interior of $\mathcal{D}_{\ell+1}$ and such that their union makes up the whole state space $\mathbb{X}$.

($A1$) "**Mixing property**": For any $\ell \geq 1$, there exists $L > \ell$ and $c, t > 0$ such that:

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

where $T_{\mathcal{D}_L}$ is the first exit time of $\mathcal{D}_L$.

($A2$) "**Escape from the Transitory domain**": *There exists $\rho > \rho_S$ such that the following boundedness property holds, where τ_E is the exit time of E and E is required to be included in $\mathcal{D}_\ell$ for ℓ sufficiently large:*

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

while the value ρ_S is defined as a survival estimate with the following definition:

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

($A3$) "**Asymptotic Comparison of Survival**":

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

1.3.3. Current proposal of an alternative to the last property

Proving the three first properties is quite straightforward for the two presented applications, as one can check in Sects. 4.3 and 5.4. The proof of Property ($A3$) is on the other hand much less direct as it involves an asymptotic in large time for the proposed ratio. It is only slighly easier than Equation (1.1) in that the inequality can be stated on a convenient subset E of $\mathbb{X}$ provided the exponential moment given in ($A2$).

As mentioned just after Equation (1.1) in the introduction, neither the two-sided estimates nor the gradient estimate or the Harnack's inequality are manifestly applicable. This is due to the fact that any singularity of the initial condition is maintained with just a decay in time (and possibly a translation in space), due to the jump event taking time.

This is why we make the emphasis on the following property that is quite easily proved in our examples. It still depends on E and ζ but also on a value ρ such that ($A2$) holds true and on a value $\epsilon > 0$ that is deduced from $(A1)$ and $(A2)$ as an intricate quantity. In practice, we thus expect the property to be deduced whatever this value of ϵ, by adjusting the other estimates accordingly. Notably, the property involves the design of a specific stopping time U_H, that we require to be infinite after a given threshold t_F in time. The design of the other stopping time V is to be adjusted according to U_H, without specific restrictions.

$(A3_F)$ "**Almost Perfect Harvest**": *There exist $t_F, c > 0$ such that for any $x \in E$ there exist two stopping times U_H and V with the following properties:*

$$\mathbb{P}_x\big(X(U_H) \in \mathrm{d}x' \,;\, U_H < \tau_\partial\big) \le c\,\mathbb{P}_\zeta\big(X(V) \in \mathrm{d}x' \,;\, V < \tau_\partial\big),$$

including the next conditions on U_H:

$$\mathbb{P}_x(U_H = \infty,\, t_F < \tau_\partial) \le \epsilon \exp(-\rho\, t_F), \qquad \text{where } \{\tau_\partial \wedge t_F < U_H\} = \{U_H = \infty\}\,.$$

Additional regularity condition of U_H are also required with respect to the Markov property, that we prove to be automatically satisfied under the mild assumption that the state space Ω is a path space (for instance prescribed as a canonical space, or according to [56], Def. (23.10), more details in the appendix Sects. A.2 and A.3).

Thanks to Theorem 2.3 in Sect. 2.2.3, assuming $(\overline{A0})$, we deduce from $(A1)$, $(A2)$ and $(A3_F)$ that $(A3)$ holds also.

1.3.4. Implications of the convergence results

Thanks to Theorem 2.8, which mostly relies on Theorems 2.1 and 2.2 of [63] together with previous Theorem 2.3, the property of quasi-stationary convergence is then deduced from $(A1)$, $(A2)$ and $(A3_F)$ (as well as from $(A1)$, $(A2)$ and $(A3)$). We then infer in Corollary 2.9 the following convergence result for α, that justifies its identification as a quasi-stationary distribution:

"**Convergence to** α": *For any $t \ge 0$ and $\mu \in \mathcal{M}_1(\mathbb{X})$ such that $\langle \mu \,|\, h\rangle > 0$:*

$$\|\, \mathbb{P}_\mu\,[\,X_t \in \mathrm{d}x \mid t < \tau_\partial] - \alpha(\mathrm{d}x)\,\|_{TV} \le C\,\frac{\|\mu - \alpha\|_{TV}}{\langle \mu \,|\, h\rangle}\; \mathrm{e}^{-\gamma\, t}. \tag{1.3}$$

It is actually justified in Theorem 2.3 and 2.8 that the family of sets $(\mathcal{D}_\ell)$ may not cover the whole set $\mathbb{X}$ for the new formulation given in Equation (1.2) to hold. Unexpectedly at first, this flexibility has been a great help for the proofs exploited in [66] and [49], even if h is still strictly positive in the latter case. In addition to the convergence depending on $\langle \mu \,|\, h\rangle$, it raises the question of identifying lower-bounds of h, which we tackle in Proposition 2.10 by proving the following property:

"**Lower-bounds of** h": *h is uniformly bounded away from zero on any set $H \subset \mathbb{X}$ for which there exists $t > 0$ and $\ell \ge 1$ such that* $\inf_{\{x \in H\}} \mathbb{P}_x(\tau_{\mathcal{D}_\ell} < t \wedge \tau_\partial) > 0$.

This justifies the identification of the domain of h as follows:

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

As in [63], Theorem 2.3, we additionally deduce the existence of the Q-process, that is of a Markov process living on this space $\mathcal{H}$ whose generator $(\mathbb{Q}_x)_{x \in \mathcal{H}}$ satisfies the following asymptotic property:

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

for any $x \in \mathcal{H}$, $s > 0$ and Λ_s any $\mathcal{F}_s$-measurable set.

As an extension of property (1.2), we also infer in Corollary 2.12 the following convergence property towards a unique stationary distribution β of this Q-process.

"**Convergence to** β": *for any probability measure* μ *on* $\mathcal{H}$ *satisfying* $\int_{\mathcal{H}} \frac{\mu(\mathrm{d}x)}{h(x)} < \infty$ *and* $t \geq 0$*:*

$$\|\mathbb{Q}_\mu\left[X_t \in \mathrm{d}x\right] - \beta(\mathrm{d}x)\|_{TV} \leq C\ \mathrm{e}^{-\gamma\, t} \cdot \left\|\mu(\mathrm{d}x) - \left(\int_{\mathcal{H}} \frac{\mu(\mathrm{d}y)}{h(y)}\right) \cdot \beta(\mathrm{d}x)\right\|_{1/h},$$

where

$$\mathbb{Q}_\mu(\mathrm{d}w) := \textstyle\int_{\mathcal{H}} \mu(\mathrm{d}x)\, \mathbb{Q}_x(\mathrm{d}w), \quad \|\mu\|_{1/h} := \left\|\dfrac{\mu(\mathrm{d}x)}{h(x)}\right\|_{TV}.$$

1.3.5. Additional robustness properties of these results

Approximation over restricted state space. The constants involved in the convergences are explicitly related to the parameters involved in the presented assumptions. Although the specific relation is very intricate, it implies that one can fairly approximate the quasi-stationary regime by restricting the state space $\mathbb{X}$. Indeed, for any $L \geq 1$ (thought to be large), let us consider the following approximation τ_∂^L of the extinction event that restricts the process X to remain in $\mathcal{D}_L$ in the time-interval $[0, \tau_\partial^L]$.

Then, as stated in Theorem 2.16, by proving any of the two sets of assumptions (for the extinction time τ_∂) we deduce that all the above results hold uniformly in L with the extinction time replaced by τ_∂^L. We mean that we can choose the constants C independent of L for the convergences to α_L, h_L, (Q_t^L) and β_L. Moreover, as L goes to infinity, λ_L converges to λ and α^L, h^L converge to α, h in total variation and pointwise respectively. Also, we deduce $\rho_S = \lambda$.

Reciprocal results. It is always satisfying to check that the sufficient conditions that one is attempting to verify are not too restrictive on the process. This is why we have also been concerned with the proof of properties as analogous as possible to our key assumptions, starting from the core result (1.2). In particular, given this property of quasi-stationary convergence, Proposition 2.13 namely ensures that Property $(A2)$ holds for a certain parameter $\rho > \lambda$, and a certain $\ell \geq 1$ regardless of the choice of the sequence $\mathcal{D}_\ell$ satisfying Assumption $(\overline{A0})$. Proposition 2.14 then asserts, under the same conditions and for the same value of ρ, that Property $(A3_F)$ is effectively met, where ζ can be chosen as any probability measure on $\mathbb{X}$.

1.4. Comparison with the literature

1.4.1. Similar recent contributions

The interest of our current result is mostly on generalizations of the Harris recurrence principle, among which [6, 19, 24, 25, 63] present the most general statements for homogeneous-in-time processes. The upper-bounds that are derived can be compared with the following inequality, that takes a more general form than Equation (1.3)

$$\|\,\mathbb{P}_\mu\left[X_t \in \mathrm{d}x \mid t < \tau_\partial\right] - \alpha(\mathrm{d}x)\,\|_{TV} \leq C(\mu)\, \mathrm{e}^{-\gamma\, t}, \tag{1.4}$$

where the main differences in the conclusions come from different expressions and interpretations of the constant $C(\mu)$. [19] first highlighted two necessary and sufficient conditions for this convergence result to be uniform, that is with $C(\mu)$ taken as a constant. The argument of a contraction in total variation norm is then simpler, yet inspired the other extensions considering a non-trivial dependency $C(\mu)$ related to the forms of the key properties required.

When considering extinction, the property of linearity over the initial condition is lost, so that linear expressions of $C(\mu)$ in terms of μ are not as natural as in the conservative case. This explains also why we came back to linear convergence statements in Equation (1.2).

Some similarity with our proposal of $(A3_F)$ can notably be found in Assumption $(H1')$ of [25], Hypothesis 2. The alternative techniques proposed in [6], [24] or [25] involve contraction estimates of the operator in specific norms that are weighted by specific functions. These functions satisfy some properties with respects to the semi-group (P_t) that justify their association with the principles of Lyapunov contraction in Harris recurrence techniques (as in [51], Chap. 16). With the help of such Lyapunov function W from $\mathbb{X}$ to $\mathbb{R}_+$, it is possible to generalize exponential moments as our property $(A2)$ and interpret the sequence $\mathcal{D}_\ell$ as level sets of W (*i.e.* $\mathcal{D}_\ell = \{x \in \mathbb{X}; W(x) \le \ell\}$). With intricate boundary conditions, providing an efficient definition of such functions in practice remains however challenging.

1.4.2. Generalization of our approach

Given the close interplay between our Assumptions $(A2)$ and $(A3_F)$, adapting the reasoning around $(A3_F)$ is not obvious. Notably, the contraction estimates exploited in [6] and [25] do not appear to relate as easily to a crucial property for our argument (namely (2.2) about the decay estimate of survival probability).

The Doob's "h-transform" is a typical technique to deduce asymptotic results of a generally non-conservative semi-group $(P_t)_t$ from the study of a related sub-Markovian semi-group. Provided that there exists a positive measurable function ψ on $\mathbb{X}$ and $\rho_\vee > 0$ such that for any $t > 0$, $P_t\psi \le \mathrm{e}^{\rho_\vee t}\psi$, the following definition indeed provides a sub-Markovian semi-group:

$$P_t^{\psi}(x, \mathrm{d}y) := \frac{\psi(y)}{\psi(x)} \mathrm{e}^{\rho_\vee t} P_t(x, \mathrm{d}y).$$

Implicitly, it means that ψ is exploited to weight the state space $\mathbb{X}$, as in the norms weighted by Lyapunov functions in [6, 24, 25]. Note that the uniqueness property of the QSD given as $\psi(y)\alpha(\mathrm{d}y)$ for P^ψ corresponds exactly to the fact that α is the unique QSD of P in the space $\mathcal{M}(\psi) := \{\mu \in \mathcal{M}(\mathbb{X})\,;\, \langle \mu \,|\, \psi \rangle < \infty\}$. This allows for other probability distributions ν to be QSD, in which case $\langle \nu \,|\, \psi \rangle = \infty$ holds necessarily.

On the other hand, our proofs would be easy to adapt to processes in discrete-time. Our techniques should generalize naturally to time-inhomogeneous processes, given the recent adaptations presented in [5, 22, 25, 37]. It can probably be extended to semi-Markov processes, *i.e.* pure jump processes for which the waiting time between jumps is not necessarily exponential (as in [1]).

1.4.3. Other frameworks

General surveys such as [26], [62] or more specifically those for population dynamics [50] give an overview on the huge literature dedicated to QSD, for which Pollett has collected quite an impressive bibliography, cf. [54].

When jumps in continuous space are involved, the reversibility property is generally expected not to hold true. They are even more exceptional when the state space is multidimensional (*cf.* [18], Appendix A).

Comparison of survival is also an essential part of perturbation techniques as in [35], Chapter 12 or in [36], yet it is mostly exploited for finite time and compared to an intrinsic convergence rate. In [40], results on the non-conservative semi-group are deduced from the study of the Q-process. Our approach may provide guidance in dealing with estimates of the poorly known survival capacity.

The other methods appear to bring less quantitative insight in terms of uniqueness (except [3] and possibly [1]) or rate of convergence. Besides the classical use of the Krein-Rutman theorem (we recall [13, 32]), extensions from fixed point argument [27] and the "renewal theory" [39], we refer in addition to the R-theory (cf [1] for semi-Markov processes, including continuous-time pure jump processes, see [59] or [60] for the original discrete-time setting). An approach specifically designed for pure jump process can also be found in [3]: auxilliary operators are then introduced that relate to the skeleton Markov chain with adjusted growth parameter.

The compactness of the semi-group is actually not required for our approach. We recall that many classical approaches rely on this property to deduce the existence of a QSD (*cf. e.g.* the reviews [26], [50]), often thanks to the Ascoli-Arzela theorem. Since the process is allowed in the illustrations given in Section 1.2 either not to jump or to have a large number of jumps in any time-interval $[0, t]$, we could not rely on this technique directly.

2. Detailed description of our results

Before we present our results in more details, it is convenient to use efficient notations, that we introduce or recall in the next subsection. We can then focus in Subsection 2.2 on the main contribution of the current article before clarifying in Susbection 2.3 the implications of this result combined with the ones of [63], Theorems 2.1-3 on the quasi-stationary convergence.

2.1. Notations

In Subsection 2.1.1 we describe our general notations, in Subsection 2.1.2 our specific setup of a càd-làg strong Markov process with extinction and clarify in Subsecton 2.1.3 some notations made to express various event restrictions.

2.1.1. Elementary notations

The most classical sets of integers are denoted as follows: $\mathbb{Z}_+ := \{0, 1, 2...\}$, $\mathbb{N} := \{1, 2, 3...\}$, $[\![m, n]\!] := \{m, m+1, ..., n-1, n\}$ (for $m \leq n$). By the notation $:=$, we simply emphasize that the equality is meant to explicit some notation. For maxima and minima, we use the following abbreviations: $s \vee t := \max\{s, t\}$, $s \wedge t := \min\{s, t\}$. In the paper, we may write $k \geq 1$ instead of $k \in \mathbb{N}$ and $t \geq 0$ (resp. $c > 0$) instead of $t \in \mathbb{R}_+ := [0, \infty)$ (resp. $c \in \mathbb{R}_+^* := (0, \infty)$), when there is no real ambiguity.

2.1.2. The stochastic process with absorption

In its generic form, we consider a strong Markov processes absorbed at some value ∂: the cemetery. More precisely, we assume that $X_s = \partial$ implies $X_t = \partial$ for all $t \geq s$. This implies that the extinction time: $\tau_\partial := \inf\{t \geq 0\,;\; X_t = \partial\}$ is a stopping time. In practice, the state space is denoted $\mathbb{X} \cup \partial$, where the cemetery ∂ is assumed to be isolated from the topology $\mathcal{B}$ on the Polish space $\mathbb{X}$.

Exploiting the same notations as in [55], Definition III.1.1, the law of the process starting from initial condition $x \in \mathbb{X} \cup \partial$ will be given by the probability $\mathbb{P}_x$. P_t would then be the semi-group of the process and the latter shall satisfy the usual measurability assumptions and the Chapman-Kolmogorov equation. Yet, we wish to avoid confusion for instance in the examples given in Subsection 1.2, where extinction is prescribed through some state-dependent rate defined *a posteriori* from an internal dynamics of X on $\mathbb{X}$. This is why we rather consider the family $(P_t)_{t \geq 0}$ as a non-conservative semigroup of operators on the set $\mathcal{B}_+(\mathbb{X})$ (respectively $\mathcal{B}_b(\mathbb{X})$) of positive (respectively bounded) $(\mathbb{X}, \mathcal{B})$-measurable real-valued functions.

In this way, restricting the state-space of observation for the process living on $\mathbb{X} \cup \partial$ amounts to compute the marginal law at time t in restriction to the event $\{t < \tau_\partial\}$. Since the latter restriction is expressed in the same way with a state-dependent extinction rate, we generally make the restriction on the event $\{t < \tau_\partial\}$ explicit (even when it would not be required for the statements of the absorbed process with state-space $\mathbb{X} \cup \{\partial\}$). Note also that the semi-group (P_t) defined through a state-dependent extinction rate, with extinction time τ_∂ given a process X with state-space $\mathbb{X}$, can also be interpreted as the above-introduced restriction for an auxilliary absorbed process $\hat{X}$ with state-space $\mathbb{X} \cup \{\partial\}$. $\hat{X}$ is simply defined as $\hat{X}_t = X_t$ for any $t < \tau_\partial$ and and as $\hat{X}_t = \partial$ for $t \geq \partial$, so that the the usual measurability assumptions are satisfied and τ_∂ indeed coincides with the hitting time of ∂ for $\hat{X}$ (*cf. e.g.* [56], Sect. 11).

In the study of the process, we will need to apply the Markov property at first entry times of either closed or open subsets. This is why we assume that the time homogeneous process X is strong Markov for the filtration $(\mathcal{F}_t)_{t \geq 0}$ that is right-continuous and complete and that it has càd-làg paths (left limited and right continuous). The hitting time (respectively the exit time out) of $\mathcal{D}$, for some domain $\mathcal{D} \subset \mathbb{X}$, will generally be denoted by

$\tau_{\mathcal{D}}$ (respectively by $T_{\mathcal{D}}$). Due to the filtration being right-continuous and complete, these are stopping times for any $\mathcal{D}$ that is either closed or open (*cf. e.g.* [56], Sect. 10 or [38, Theorem 52]).

For any probability measure μ on $\mathbb{X}$, and $f \in \mathcal{B}_+(\mathbb{X})$ (or $f \in \mathcal{B}_b(\mathbb{X})$), we use the notations:

$$\mathbb{P}_\mu(.) := \int_{\mathbb{X}} P_x(.)\ \mu(\mathrm{d}x), \quad \langle \mu \,|\, f \rangle := \int_{\mathbb{X}} f(x)\ \mu(\mathrm{d}x).$$

We also denote by $\mathbb{E}_x$ (respectively $\mathbb{E}_\mu$) the expectation according to $\mathbb{P}_x$ (respectively $\mathbb{P}_\mu$).

The set of respectively probability measures on $\mathbb{X}$, of positive and of signed measures are denoted respectively $\mathcal{M}_1(\mathbb{X})$, $\mathcal{M}_+(\mathbb{X})$ and $\mathcal{M}(\mathbb{X})$. For any $B \in \mathcal{B}$ and $\mu \in \mathcal{M}_1(\mathbb{X})$, $\mu P_t(B)$ is clearly defined as $\mathbb{P}_\mu(X_t \in B, t < \tau_\partial)$. As above-mentioned, the restriction on $\{t < \tau_\partial\}$ is made explicit but automatically satisfied provided X is indeed absorbed in ∂. We generally denote the action of P_t on $\mu \in \mathcal{M}_1(\mathbb{X})$ as follows:

$$\mu P_t(\mathrm{d}y) := \mathbb{P}_\mu(X_t \in \mathrm{d}y\,;\ t < \tau_\partial), \quad \text{ or } \langle \mu P_t \,|\, f \rangle = \langle \mu \,|\, P_t f \rangle = \mathbb{E}_\mu[f(X_t)\,;\ t < \tau_\partial],$$

where $f \in \mathcal{B}_+(\mathbb{X})$ or $f \in \mathcal{B}_b(\mathbb{X})$. Let us then define the family of conditioned operators $(A_t)_{t\geq 0}$ acting as follows on any probability measure $\mu \in \mathcal{M}_1(\mathbb{X})$:

$$\mu A_t(\mathrm{d}y) := \mathbb{P}_\mu(X_t \in \mathrm{d}y \,|\, t < \tau_\partial), \qquad \langle \mu A_t \,|\, f \rangle = \mathbb{E}_\mu[f(X_t) \,|\, t < \tau_\partial] = \frac{\mathbb{E}_\mu[f(X_t)\,;\ t < \tau_\partial]}{\mathbb{P}_\mu[t < \tau_\partial]}.$$

μA_t is what we call the MCNE at time t, with initial distribution μ, as it is "the Marginal distribution (at time t) Conditioned upon the fact that No Extinction has yet occurred" (also at time t).

In this setting, the family $(P_t)_{t\geq 0}$ (respectively $(A_t)_{t\geq 0}$) defines a linear but non-conservative semigroup (respectively a conservative but non-linear semigroup) of operators on $\mathcal{M}_1(\mathbb{X})$ endowed with the total variation norm. We consider the following definition of this norm, generally for any signed measure $\mu \in \mathcal{M}(\mathbb{X})$: $\|\mu\|_{TV} := \sup\{|\mu(A)|\,;\ A \in \mathcal{B}\}$. While the semi-group P is directly generalized by linearity for any signed measure, note that it is not as clear for the semi-group A because $\mathbb{P}_\mu[t < \tau_\partial]$ could then be equal to zero for some $t > 0$.

A probability measure α is said to be the *quasi-limiting distribution* of an initial condition $\mu \in \mathcal{M}_1(\mathbb{X})$ if: $\forall\, B \in \mathcal{B}, \quad \lim_{t\to\infty} \mathbb{P}_\mu(X_t \in B \,|\, t < \tau_\partial) := \lim_{t\to\infty} \mu A_t(B) = \alpha(B)$.
It is now classical (*cf. e.g.* Proposition 1 in [50]) that α is then a quasi-stationary distribution or QSD, in the sense that: $\alpha A_t(\mathrm{d}y) = \alpha(\mathrm{d}y)$ holds for any positive t. Also, for any such QSD, there exists a unique extinction rate $\lambda > 0$ such that: $\forall\, t \geq 0,\ \mathbb{P}_\alpha(t < \tau_\partial) = \exp[-\lambda\, t]$.

2.1.3. Conditions, stopping times and random events

While dealing with the Markov property between different stopping times, we wish to clearly indicate with our notation that we introduce a copy of X (ie a process with the same semigroup P_t) independent of X given its initial condition. This copy (and the associated stopping times) is then denoted with a tilde ($\widetilde{X}$, $\widetilde{\tau}_\partial$, $\widetilde{T}_{\mathcal{D}}$ *etc.*). For instance in the notation $\mathbb{P}_{X(\tau_E)}(t - \tau_E < \widetilde{\tau}_\partial)$, τ_E and $X(\tau_E)$ refer to the initial process X while $\widetilde{\tau}_\partial$ refers to the copy $\widetilde{X}$.

Besides, some notations of semi-colons and commas have meanings that are specific to our probabilistic notations and currently very efficient given that we often consider restrictions on various events. For expectations, the terms after the semi-colon indicate the sets of conditions under which the former term is counted (it is replaced by 0 if the conditions are not met). For sets, and notably random ones, the terms after the semi-colon indicate the sets of conditions under which the former term is included. Different conditions may be separated by commas, notably when combinations of those are introduced through indices. For instance, given any random variables F, $(X_i)_{i\leq d}$, for $d\geq 1$, T, T' and $t>0$, the following notation can be translated as follows:

$$\mathbb{E}(F\,;\ \forall\, i\leq d,\ X_i\geq 0\,,\ t<T) = \mathbb{E}\big(F\cdot \textstyle\prod_{\{i\leq d\}} \mathbf{1}_{\{X_i\geq 0\}}\cdot \mathbf{1}_{\{t<T\}}\big).$$

To give another example, the following notation:

$$\{s\geq T'\,;\ \forall\, i\leq d,\ X_i\geq 0\,,\ s<T\},$$

is to be understood as the random set $\mathcal{S}$ (that is thus dependent on $\omega\in\Omega$) defined as follows. If there exists i such that $X_i(\omega)$ is negative or if $T'(\omega)\geq T(\omega)$, then $\mathcal{S}=\emptyset$. Otherwise, $\mathcal{S}$ consists of all values $s\in\mathbb{R}_+$ such that $s\in[T'(\omega), T(\omega))$. The infimum of an empty set is generally to be taken as ∞. Without such semi-colon, the set has to be considered as the part of Ω for which the conditions are satisfied, for instance:

$$\{\forall\, i\leq d,\ X_i\geq 0\,,\ t<T\} = \cap_{\{i\leq d\}}\{X_i\geq 0\}\cap\{t<T\},$$

consists of all $\omega\in\Omega$ such that for any $i\leq d$, $X_i(\omega)\geq 0$ and $t<T(\omega)$.

2.2. Coupling approach including failures

The main contribution of this article is presented in this Sect. 2.2, namely the derivation of Property (1.1) thanks to our new criterion. We start in Sect. 2.2.1 by explaining the key properties derived from [63], Sect. 2 and the role of the associated parameters involved in our new criterion. After presenting this new criterion of "Almost Perfect Harvest" in Sect. 2.2.2, we state our main Theorem 2.3 in the next Sect. 2.2.3.

2.2.1. The associated assumptions

Our core property of "almost perfect harvest" exploits several parameters ($\rho>0$, E and ζ) whose choices are strongly tied to several key properties highlighted in [63]. We shall thus start by explaining the key properties of [63] on which we rely and thus the role of these parameters.

The approach is trajectory-based and designed to handle specific dependencies on the initial condition, so it appears efficient to consider a customizable covering by an increasing sequence of sets, as specified in the next property.

$(\underline{A0})$: **"Specification of space"** There exists a sequence $(\mathcal{D}_\ell)_{\ell\geq 1}$ of closed subsets of $\mathbb{X}$ such that for any $\ell\geq 1$, $\quad\mathcal{D}_\ell\subset int(\mathcal{D}_{\ell+1})\quad$ (with $int(\mathcal{D})$ the interior of $\mathcal{D}$).

This sequence serves as a reference for the other key properties, notably through the following notation, for subsets that are "regular" with respect to this specification:

$$\mathbf{D} := \{\mathcal{D}\,;\ \mathcal{D}\text{ is closed and }\exists\,\ell\geq 1,\ \mathcal{D}\subset\mathcal{D}_\ell\}\,. \tag{2.1}$$

As it is stated in [63], Sect. 2 (and by extension in the introduction), this property $(\underline{A0})$ is often strengthened as follows. This additional condition of complete covering enables to obtain uniqueness of the QSD as such and not only among the QSD with minimal extinction rate.

$(\overline{A0})$: the sequence $(\mathcal{D}_\ell)_{\ell\geq 1}$ satisfying $(\underline{A0})$ is such that $\quad\bigcup_{\ell\geq 1}\mathcal{D}_\ell=\mathbb{X}$.

The next assumption consists in a minoration of the density of the process after a given time. This minoration extends the classical Doeblin's inequality where the marginal law is lower-bounded uniformly over the initial conditions. This is the step that produces the mixing of the past dependencies. As in Harris recurrence technique, this lower-bound needs only to be uniform locally in the initial condition, at the expense of an associated contraction estimate. In our case, we further restrict the probability to trajectories that remain locally confined and exploit for practical convenience the minoration with the restriction on the event of survival $\{t < \tau_\partial\}$ rather than with the conditioning on the event.

We recall the following definitions for the exit and first entry times of any set $\mathcal{D}$:

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

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

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

The last assumption corresponds to the contraction estimate in Harris recurrence techniques, in the form of an exponential moment of return to some reference subset E of $\mathbb{X}$. To combine this property with the mixing stated in $(A1)$, we require this set E to be included in one of the $\mathcal{D}_\ell$. For practical convenience, we consider the moment estimate on the stopping time $\tau_\partial \wedge \tau_E$ for the unconditioned expectation. For the contraction to be of relevance for the conditioned process, an exponential moment estimate with a parameter larger than the extinction rate of the process appears crucial. Note that the extinction rate is often not directly accessible and has to be evaluated.

$(A2)[\rho, E]$: **"Escape from the Transitory domain"** *The value $\rho > 0$ and the subset $E \in \mathbf{D}$ are such that:*

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

ρ as stated in $(A2)$ and the next property $(A3_F)$ is required in the following theorems to be strictly larger than the following **"survival estimate"**:

$$\rho_S := \sup\Big\{\gamma \in \mathbb{R}\,\Big|\, \sup_{L\geq 1}\inf_{t>0}\ \mathrm{e}^{\gamma t}\,\mathbb{P}_\zeta(t < \tau_\partial \wedge T_{\mathcal{D}_L}) = 0\Big\}.$$

Moreover, the set E shall be common for $(A2)$ and $(A3_F)$.

From the proof of Corollary 5.2.1 in [63], we deduce the following property of survival as a direct consequence the above properties.

Proposition 2.1. *Assume that $(\underline{A0})$, $(A1)[\zeta]$ and $(A2)[\rho, E]$ hold for some $\zeta \in \mathcal{M}_1(\mathbb{X})$, $\rho > \rho_S$ and $E \in \mathbf{D}$. Then, there exists $t_S, c_S > 0$ and $\hat{\rho}_S \in (\rho_S, \rho)$ such that:*

$$\forall\, u \geq 0,\ \forall\, t \geq u + t_S, \qquad P_\zeta(t - u < \tau_\partial) \leq c_S\, \mathrm{e}^{\hat{\rho}_S\, u}\, \mathbb{P}_\zeta\left(t < \tau_\partial\right). \tag{2.2}$$

It means that for sufficiently large times, the exponential decay in time with parameter $\hat{\rho}_S$ provides a relevant estimation of the decay of the probability of survival starting from the initial conditon ζ.

2.2.2. The new key property

We are now in position to introduce the following weak form of Harnack's inequality, whose purpose is to imply the property (1.1):

$(A3_F)[\rho, E, \zeta]$: **"Almost perfect harvest"** *The value $\rho > 0$, the subset $E \in \mathbf{D}$ and the probability measure $\zeta \in \mathcal{M}_1(\mathbb{X})$ are such that the following condition holds:*
For any $\epsilon \in (0, 1)$, there exist $t_F, c > 0$ such that for any $x \in E$ there exist two stopping times U_H and V with

the following properties:

$$\mathbb{P}_x\big(X(U_H) \in \mathrm{d}x'\,;\ U_H < \tau_\partial\big) \le c\,\mathbb{P}_\zeta\big(X(V) \in \mathrm{d}x'\,;\ V < \tau_\partial\big), \tag{2.3}$$

including the next conditions on U_H:

$$\{\tau_\partial \wedge t_F < U_H\} = \{U_H = \infty\} \quad \text{and} \quad \mathbb{P}_x(U_H = \infty,\, t_F < \tau_\partial) \le \epsilon\, \exp(-\rho\, t_F). \tag{2.4}$$

Then, provided that the state space Ω is a "*path space*", we may extend the definition of U_H into a harvesting time that allows for failures in the harvest, as stated in the next proposition whose proof is deferred to Appendix A.2. Having Ω as a path space is a mild yet very convenient assumption for dealing with a trajectorial approach with extinction and iterated stopping times. It implies notably the existence of a family $(\theta[t])_{t>0}$ of shift operators on Ω such that for any $\omega \in \Omega$ and any adapted process $(X_s)_{s\ge 0}$, we have $X_s \circ \theta[t](\omega) = X_{t+s}(\omega)$.

Proposition 2.2. *Assume that Ω is a path space and that U_H is globally defined as a stopping time for any initial condition on E with the above requirement that $\{\tau_\partial \wedge t_F < U_H\} = \{U_H = \infty\}$ for some $t_F > 0$. Let $\tau_E^F := \inf\{s \ge t_F\,;\ X_s \in E\}$. Then there exists another stopping time U_H^∞ such that:*

(*i*) *$U_H^\infty = U_H$ on the event $\{\tau_\partial \wedge U_H \le \tau_E^F\}$,*
(*ii*) *For any ω in the event $\{\tau_E^F < \tau_\partial \wedge U_H\}$, it holds*

$$U_H^\infty(\omega) = \tau_E^F(\omega) + U_H^\infty \circ \theta\big[\tau_E^F\big](\omega).$$

Interpretation of the core assumption

This property is to be compared to the more classical Harnack's inequality, which should take the following form (as can be seen in [63]):
The subset $E \in \mathbf{D}$ and the probability measure $\zeta \in \mathcal{M}_1(\mathbb{X})$ are such that there exist $t_F, v, c > 0$ for which the following inequality holds for any $x \in E$:

$$\mathbb{P}_x(X_{t_F} \in \mathrm{d}x'\,;\ t_F < \tau_\partial) \le c\,\mathbb{P}_\zeta(X_v \in \mathrm{d}x'\,;\ v < \tau_\partial). \tag{2.5}$$

One main difficulty in obtaining such property is that the constant times t_F and v cannot be adjusted much for convenience to the specific process, while the upper-bound shall capture the realizations of the whole set of trajectories sarting from x and that have still not reached extinction. In many jump processes on the other hand, and notably those proposed in introduction, the distribution $\delta_x P_t$ keeps a singular component (say a Dirac mass) with respect to any distribution such as ζP_v that is to be typical of the asymptotic profile (so typically with a density with respect to the Lebesgue measure).

Thanks to $(A3_F)$, we will similarly be able to couple the trajectories of the processes starting with the initial conditions respectively $x \in E$ and ζ so that they coincide with a time lag for sufficiently large times. In other words, we want to embed the "tail trajectories" of the process starting from x into the trajectories starting from ζ, a time lag being allowed. When (2.5) holds true, the trajectories with initial condition x can namely be coupled after time t_F, with a time-lag of $t_F - v$.

This embedding procedure is what we describe as the "harvest" of the tail trajectories. With $(A3_F)$, this embedding is here to occur after time U_H^∞, the "harvesting time", the time-lag being more flexible and notably allowed to be random.

The Markov property being granted, Inequality (2.3) is what makes this embedding possible. This embedding is directly obtained on the event $\{U_H \le t_F \wedge \tau_\partial\}$, in which case U_H and U_H^∞ coincide. The Markov property being exploited for the coupling, we request both U_H and V to be stopping time.

It is however rarely expected that U_H can be obtained upper-bounded by a uniform constant t_F such that Inequality (2.3) still holds. Possibly several attempts of coupling may thus be required under some specific conditions, that we describe as failures. A "failure" to do the harvest is then characterized by the event $\{\tau_\partial \wedge t_F < U_H\}$ (which we require to coincide with $\{U_H = \infty\}$). The objective of establishing the embedding of tail trajectories except in case of rare events in probability is what motivates our denomination of "almost perfect harvest".

In case of failures, U_H^∞ must be larger than t_F, and a new attempt can directly be implemented only after τ_E^F. We need to ensure that the events of recurrent failures play a negligible role in the probability of long-term survival. This is why, in complement to the factor ϵ being sufficiently small, we require the penalisation by $\exp(-\rho\, t_F)$ to compensate for the extinction rate (during the time-interval $[0, t_F]$ for the process with initial condition ζ). For such a compensation to be exploitable, the value of ρ will be required to be greater than the known lower bound ρ_S for the decay of the survival probability.

The condition of failure where $\{U_H = \infty\}$ has to be adjusted for convenience to the model. As the allowed error ϵ goes to 0, we expect to see the failure conditions stated through certain thresholds being less and less stringent. This certainly leads to larger and larger constants c in (2.3), and often to larger and larger constants t_F.

Because the attempts of coupling are meant to be iterated after each failure, a condition related to the Markov property on U_H^∞ is requested. Assuming Ω to be a path space appears to be the most relevant way to deal with it. It takes into account the waiting time before the process comes back to E and a new attempt can be initiated. Provided that the definition of U_H do not depend in a singular way on the initial condition $x \in E$, this condition on U_H^∞ should be easily satisfied.

2.2.3. The central result of the paper

Theorem 2.3. *Assume that Ω is a path space and that there exist $\zeta \in \mathcal{M}_1(\mathbb{X})$, $t_S, c_S > 0$, $\rho \geq \hat{\rho}_S > \rho_S > 0$, and $E \in \mathbf{D}$ such that Inequality (2.2), Assumptions ($\underline{A0}$), ($A1$)[ζ], ($A2$)[ρ, E] and ($A3_F$)[ρ, E, ζ] hold. Then:*

$$\limsup_{t\to\infty} \sup_{x\in E} \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_\zeta(t < \tau_\partial)} < \infty. \tag{2.6}$$

For the implication that (1.1) is a consequence (($\underline{A0}$), ($A1$) and ($A2$) being granted) of this more local property (which is exactly ($A3$)), we refer to [63], Thm. 5.2. A careful check of the proof ensures that the additional requirement $\bigcup_{\ell\geq 1}\mathcal{D}_\ell = \mathbb{X}$ is actually not exploited therein.

Remark 2.4. In ($A1$), a confinement in some subspace $\mathcal{D}_L$ is required. This confinement part of the assumption is actually not involved in the proof of Theorem 2.3, but is in fact exploited to deduce Theorem 2.8. We kept it to avoid too frequent variations of our assumptions.

2.3. Exponential convergence to a unique *QSD*

Thanks to Theorem 2.3, we see that the general conclusions of [63], Theorems 2.1-3 can be derived from a new set of assumptions, as we present it in this Subsection.

2.3.1. Two general sets of assumptions

We say that Assumption ($\mathbf{A_F}$) holds, whenever:

"There exists $\zeta \in \mathcal{M}_1(\mathbb{X})$ such that ($A1$)[ζ] holds for a specific sequence ($\mathcal{D}_\ell$) satisfying ($\underline{A0}$). Moreover, there exists $\rho > \rho_S$ and $E \in \mathbf{D}$ such that Assumptions ($A2$)[ρ, E] and ($A3_F$)[ρ, E, ζ] hold. Finally, Ω is a path space".

Slightly adapting [63] (regarding ($\underline{A0}$)), we say that Assumption ($\mathbf{A}$) holds, whenever:

"There exists ζ such that $(A1)[\zeta]$ holds for a specific sequence $(\mathcal{D}_\ell)$ satisfying $(\underline{A0})$. Moreover, there exists $\rho > \rho_S$ and $E \in \mathbf{D}$ such that Assumptions $(A2)[\rho, E]$ and $(A3)[E, \zeta]$ hold".

Remark 2.5. ⋆ The purpose of Theorem 2.3 is to prove that Assumption $(\mathbf{A_F})$ actually implies Assumption $(\mathbf{A})$, for which the conclusions of [63] can be adapted (see Sect. 3.2 for the adjustments).

⋆ Almost sure extinction is not at all needed for our proof (which would in fact include the case where there is no extinction, or only in some "transitory domain").

⋆ In our set of assumptions, the case where $\mathcal{D}_\ell := \mathbb{X}$ for any ℓ is actually allowed and fits in the setting of uniform convergence covered in [19].

2.3.2. Main consequences of our main result

The main property that we deduce in Theorem 2.8 and in our subsequent applications can be stated as in the following definition:

Definition 2.6. For any linear positive and bounded semi-group $(P_t)_{t\geq 0}$ acting on a state space $\mathbb{X}$, we say that P displays a uniform exponential quasi-stationary convergence with characteristics $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathbb{X}) \times B(\mathbb{X}) \times \mathbb{R}$ if $\langle \alpha \,|\, h \rangle = 1$ and there exists $C, \gamma > 0$ such that for any $t > 0$ and for any measure $\mu \in \mathcal{M}(\mathbb{X})$ with $\|\mu\|_{TV} \leq 1$:

$$\left\| \mathrm{e}^{\lambda t} \mu P_t - \langle \mu \,|\, h \rangle \cdot \alpha \right\|_{TV} \leq C \mathrm{e}^{-\gamma t}. \tag{2.7}$$

This definition slightly extends the one given in introduction, cf (1.2), by considering a general class of semi-groups and allows the value λ to be negative. Since $\langle \alpha \,|\, h \rangle = 1$ and since P_t is linear, elementary computations show that the above property implies the apparently stronger one given in (1.2) with $\|\mu - \alpha\|_{TV}$ as additional factor.

The property given in Definition 2.6 implies that α is a QSD with extinction rate λ, as stated in the following fact, whose proof is deferred to the appendix.

Fact 2.7. If a semi-group P displays a uniform exponential quasi-stationary convergence with characteristics (α, h, λ), then for any $t > 0$, $\alpha P_t(\mathrm{d}s) = \mathrm{e}^{-\lambda t} \alpha(\mathrm{d}s)$.

Remark 2.7. ⋆ It is elementary that $h_t : x \mapsto \mathrm{e}^{\lambda t} \langle \delta_x P_t \,|\, \mathbf{1} \rangle$ converges in the uniform norm to h, implying that h is non-negative. Since $h_{t+t'} = \mathrm{e}^{\lambda t} P_t h_{t'}$, one can also easily deduce that $\mathrm{e}^{\lambda t} P_t h = h$.

⋆ By "characteristics", we express that they are uniquely defined.

Theorem 2.8. *Assume that either $(\mathbf{A_F})$ or $(\mathbf{A})$ holds. Then, P displays a uniform exponential quasi-stationary convergence with some characteristics (α, h, λ) in $\mathcal{M}_1(\mathbb{X}) \times B(\mathbb{X}) \times \mathbb{R}_+$. It also holds true that h is bounded away from zero on any $\mathcal{D}_\ell$ and that $\alpha(\mathcal{D}_\ell) > 0$ for ℓ sufficiently large.*

The convergence to α is made more precise by the following corollary:

Corollary 2.9. *Assume (2.7). Then for any $t \geq 0$ and $\mu \in \mathcal{M}_1(\mathbb{X})$ such that $\langle \mu \,|\, h \rangle > 0$:*

$$\| \, \mathbb{P}_\mu(\, X_t \in \mathrm{d}x \mid t < \tau_\partial) - \alpha(\mathrm{d}x) \, \|_{TV} \leq C \frac{\|\mu - \alpha\|_{TV}}{\langle \mu \,|\, h \rangle} \, \mathrm{e}^{-\gamma \, t}. \tag{2.8}$$

Thanks to Theorem 2.8, it is elementary that h is positive under Assumption $(\overline{A0})$. It might be useful otherwise to exploit the following proposition to identify *a posteriori* the domain on which h is positive.

Proposition 2.10. *Assume that $(\mathbf{A_F})$ or $(\mathbf{A})$ holds. Then, the survival capacity h is uniformly bounded away from zero on any set $H \subset \mathbb{X}$ that satisfies the following condition:*
(H_0) : *there exists $t > 0$, $\ell \geq 1$ such that* $\inf_{\{x \in H\}} \mathbb{P}_x(\tau_{\mathcal{D}_\ell} < t \wedge \tau_\partial) > 0.$

It implies the following identification :

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

Remark 2.11. Corolllary 2.9 implies that if ν were a QSD different from α, then $\nu A_t = \nu \neq \alpha$ thus $\langle \nu \,|\, h\rangle = 0$. This implies $\nu(\mathcal{H}) = 0$. So by contraposition, the previous proposition provides a practical way to ensure the uniqueness of the QSD *a posteriori*.

Once the set $\mathcal{H}$ is clarified, we can study Doob's h-transform of the semi-group P with weight given by the survival capacity. This is the so-called Q-process that is actually a conservative Markov process as stated in the next corollary:

Corollary 2.12. *Under again* $(\mathbf{A_F})$ *or* $(\mathbf{A})$, *with* (α, h, λ) *the characteristics of exponential convergence of* P, *the following properties hold:*
(i) **Existence of the Q-process:**
There exists a family $(\mathbb{Q}_x)_{x\in\mathcal{H}}$ *of probability measures on* Ω *defined by:*

$$\lim_{t\to\infty} \mathbb{P}_x(\Lambda_s \,|\, t < \tau_\partial) = \mathbb{Q}_x(\Lambda_s), \tag{2.9}$$

for any $x \in \mathcal{H}$, $s > 0$ *and* Λ_s *any* $\mathcal{F}_s$*-measurable set. The process* $(\Omega; (\mathcal{F}_t)_{t\geq 0}; (X_t)_{t\geq 0}; (\mathbb{Q}_x)_{x\in\mathcal{H}})$ *is an* $\mathcal{H}$*-valued homogeneous strong Markov process.*

(ii) Weighted exponential ergodicity of the Q-process:
The measure $\beta(\mathrm{d}x) := h(x)\,\alpha(\mathrm{d}x)$ *is the unique invariant probability measure under* $\mathbb{Q}$.
Moreover, for any $\mu \in \mathcal{M}_1(\mathcal{H})$ *satisfying* $\langle \mu \,|\, 1/h\rangle < \infty$ *and* $t \geq 0$*:*

$$\left\| \mathbb{Q}_\mu\left[\, X_t \in \mathrm{d}x\right] - \beta(\mathrm{d}x)\right\|_{TV} \leq C\, \|\mu - \langle \mu \,|\, 1/h\rangle\, \beta\|_{1/h}\, \mathrm{e}^{-\gamma\, t}, \tag{2.10}$$

where $\mathbb{Q}_\mu(\mathrm{d}w) := \int_{\mathcal{H}} \mu(\mathrm{d}x)\, \mathbb{Q}_x(\mathrm{d}w)$, $\quad \|\mu\|_{1/h} := \left\| \dfrac{\mu(\mathrm{d}x)}{h(x)} \right\|_{TV}$.

The constant $\langle \mu \,|\, 1/h\rangle$ before β in (2.10) is optimal up to a factor 2, due to the following fact, that is proved in the appendix. This is also the case for the implicit constant 1 before α in (2.8).

Fact 2.13. The next two minorations hold for any probability measures μ and any $u > 0$:

$$\|\mu - u\,\alpha\|_{1/h} \geq (1/2)\cdot \|\mu - \langle \mu \,|\, 1/h\rangle \beta\|_{1/h}, \qquad \|\mu - u\,\alpha\|_{TV} \geq (1/2)\cdot \|\mu - \alpha\|_{TV}.$$

2.3.3. Additional robustness properties of the results

Reciprocal results The following propositions are stated to justify that our assumptions are not particularly restrictive. We shall see in the following Theorem 2.16 that ρ_S actually equals λ if $(\mathbf{A_F})$ is satisfied. We thus aim at stating reciprocal statements in terms of a parameter ρ satisfying $\rho > \lambda$. The first proposition concerns $(A2)$, that is derived under the assumption of exponential convergence together with property $(\overline{A0})$.

Proposition 2.13. *Assume that* $\mathbb{X} = \bigcup_{\ell\geq 1} \mathcal{D}_\ell$ *where* $(\mathcal{D}_\ell)_{\ell\geq 1}$ *is an increasing subsequence. Assume that* (2.7) *is satisfied. Then, for* $\rho := \lambda + \gamma/2$, *there exists* $\ell \geq 1$ *such that:*

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

Since $(A3_F)$ is the most intricate property, one may more likely suspect that it leads to restrictions. The following proposition shows to what extend we may generally reply that this is not the case, in that the convergence property (1.2) actually implies a general form of property $(A3_F)$. Note that μ in Proposition 2.14 is meant to represent the reference measure ζ in the original statement of $(A3_F)$.

Proposition 2.14. *Assume that (2.7) is satisfied. Then, for $\rho := \lambda + \gamma/2$, there exist $c_F > 0$ such that for any $\mu \in \mathcal{M}_1(\mathbb{X})$, and any $t_F > 0$ sufficiently large, there exists for any $x \in \mathbb{X}$ two stopping times U_H and V satisfying the three following properties:*

$$\mathbb{P}_x(t_F < \tau_\partial \wedge U_H) \le \frac{c_F}{\langle \mu \,|\, h\rangle} \mathrm{e}^{-\rho t_F}, \qquad \textit{where } \{\tau_\partial \wedge t_F < U_H\} = \{U_H = \infty\}\,;$$
$$\textit{and} \quad \mathbb{E}_x(X_{U_H} \in \mathrm{d}y\,;\, U_H < \tau_\partial) \le \frac{\|h\|_\infty}{\langle \mu \,|\, h\rangle} \mathbb{E}_\mu(X_V \in \mathrm{d}y\,;\, V < \tau_\partial).$$

To deduce that $(A3_F)$ holds for any measure μ satisfying $\langle \mu \,|\, h\rangle > 0$, one has simply to remark that the following inequality holds with $\hat{\rho} := \lambda + \gamma/4$ for any $\epsilon > 0$ provided t_F is sufficiently large:

$$(c_F/\langle \mu \,|\, h\rangle) \cdot \mathrm{e}^{-\rho t_F} \le \epsilon \cdot \mathrm{e}^{-\hat{\rho} t_F}.$$

The proof of Proposition 2.14 is also meant to provide a relevant intuition on the role of the parameters involved in the expression of $(A3_F)$. It follows in Sect. 3.4 the one of Proposition 2.13.

Remark 2.15. The derivation of the mixing property $(A1)$ from a convergence property in total variation such as 2.7 appears much more difficult to obtain. It indeed requires a lower-bound with respect to some measure ζ that is uniform over the initial condition in some $\mathcal{D}_\ell$, while the total variation informs about the discrepancy between two measures. Given our additional restriction involving $T_{\mathcal{D}_L}$, a property of uniform convergence would also at least be required, as stated in the next Theorem 2.16.

Uniformity in the localization procedure The constants involved in the convergences are explicitly related to the ones in Assumptions $(\mathbf{A_F})$ or $(\mathbf{A})$. Although the specific relation is very intricate, it implies the following corollary of approximation:

Theorem 2.16. *Assume that $(\mathbf{A_F})$ or $(\mathbf{A})$ holds. Then, there exists $L_\vee \ge 1$ such that all the preceding results hold true with the same constants involved in the convergences when τ_∂ is replaced by $\tau_\partial^L := \tau_\partial \wedge T_{\mathcal{D}_L}$ for any $L \ge L_\vee$, or more generally replaced by $\tau_\partial \wedge T_{\mathcal{D}'}$ for any subset $\mathcal{D}'$ of $\mathbb{X}$ that contains $\mathcal{D}_{L_\vee}$.*

Let λ_L, α^L, h^L be the corresponding extinction rates, QSD and survival capacities. Then as L goes to infinity, λ_L converges to λ and α^L, h^L converge to α, h in total variation and pointwise respectively. Also, we deduce $\rho_S = \lambda$.

Remark 2.17. In Theorem 2.16, when we replace τ_∂ by $\tau_\partial^L := \tau_\partial \wedge T_{\mathcal{D}_L}$ in the convergence statements, it means that we consider instead of the original semi-group the one generated by the altered Markov process $X^{(L)}$ evolving on $\mathcal{D}_L \cup \{\partial\} \subset \mathbb{X} \cup \{\partial\}$, still with cemetery ∂, such that $X_t^{(L)} = X_t$ for any $t < \tau_\partial^L$ (thus $t < \tau_\partial$) and such that $X_t^{(L)} = \partial$ for any $t \ge \tau_\partial^L$.

3. Proofs of the general results

3.1. Proof of Theorem 2.3

Recall the expression of c_S in Inequality (2.2). Thanks to Assumption $(A2)$ that controls the escape time from the transitory domain $\mathcal{T}$, we may define the constant $e_{\mathcal{T}}$ as follows:

$$e_{\mathcal{T}} := \sup_{\{x \in \mathbb{X}\}} \mathbb{E}_x \left(\exp\left[\rho\left(\tau_\partial \wedge \tau_E\right)\right]\right). \tag{3.1}$$

In the following, every time we will apply Assumption $(A3_F)$, we will exploit the following parameter $\epsilon := (2\, c_S\, e_{\mathcal{T}})^{-1}$. For t_F the associated deterministic upper-bound on U_H, we define $t_A := t_F + t_S$.

The proof of Theorem 2.3 is completed in four steps, and the conclusions of the first three steps are summarized in respectively Lemma 3.1, 3.2 and 3.3.

3.1.1. Step 1: the case without failures.

Step 1 consists in proving the following lemma,

Lemma 3.1. *Assume that Inequality* (2.2) *and Assumption* $(A3_F)$ *hold with the above parameters. Then, there exists* $C_0 > 0$ *such that the following upper-bound holds for any* $x \in E$ *and any* $t \geq t_A$*:*

$$\mathbb{P}_x(U_H < t < \tau_\partial) \leq C_0\, \mathbb{P}_\zeta(t < \tau_\partial).$$

Its proof presented next implicitly relies on the coupling of trajectories starting from x on the event $\{U_H < t < \tau_\partial\}$ with trajectories starting from ζ.

Proof: The Markov property is exploited as follows in relation to $(A3_F)$:

$$\begin{aligned}
\mathbb{P}_x(U_H < t < \tau_\partial) &= \mathbb{E}_x\left[\mathbb{P}_{X[U_H]}(t - U_H < \widetilde{\tau}_\partial)\,;\, U_H < \tau_\partial\right] \\
&\leq C\, \mathbb{E}_\zeta\left[\mathbb{P}_{X[V]}(t - t_F < \widetilde{\tau}_\partial)\,;\, V < \tau_\partial\right] \\
&\leq C\, \mathbb{P}_\zeta\left[t - t_F < \tau_\partial\right],
\end{aligned}$$

where we exploited Assumption $(A3_F)$ for the first inequality with the fact that $U_H \leq t_F$ holds on the event $\{U_H < \tau_\partial\}$, then noted that $t - t_F + V \geq t - t_F$ for the second inequality. Thanks to Inequality (2.2), with $t \geq t_F + t_S$, the following upper-bound is derived:

$$\mathbb{P}_x(U_H < t < \tau_\partial) \leq C \cdot c_S \cdot \mathrm{e}^{\rho\, t_F}\, \mathbb{P}_\zeta\left[t < \tau_\partial\right].$$

Lemma 3.1 is thus satisfied with: $C_0 := C \cdot c_S \cdot \mathrm{e}^{\rho\, t_F}$. □

3.1.2. Step 2: the case where the process avoids the set E for long after a failure

Step 2 consists in proving the following lemma:

Lemma 3.2. *Assume that Assumption* $(A2)$, *Inequalities* (2.2) *and Assumption* $(A3_F)$ *hold with* $\rho > \hat{\rho}_S$. *Then, there exists* $C_F > 0$ *such that the following upper-bound holds for any* $x \in E$ *and* $t \geq t_A$*:*

$$\mathbb{P}_x(U_H = \infty\,,\; t - t_A \leq \tau_E^F\,,\; t < \tau_\partial) \leq C_F\, \mathbb{P}_\zeta(t < \tau_\partial).$$

Proof: Thanks to Equation (3.1), the following upper-bound holds a.s. on $\{U_H = \infty, t_F < \tau_\partial\}$:

$$\mathbb{E}_{X_{t_F}}\left[\exp\left(\rho[\widetilde{\tau}_E \wedge \widetilde{\tau}_\partial]\right)\right] \leq e_{\mathcal{T}}.$$

Thanks to the Markov's inequality, this implies the next upper-bound a.s. on the same event:

$$\mathbb{P}_{X_{t_F}}(t - t_A - t_F \leq \widetilde{\tau}_E\,;\, t - t_F < \widetilde{\tau}_\partial) \leq e_{\mathcal{T}} \cdot \mathrm{e}^{-\rho[t - t_A - t_F]}.$$

In combination with the Markov property, we deduce the next upper-bound:

$$\mathbb{P}_x(U_H = \infty\,,\; t - t_A - \tau_E \leq \tau_E^F\,,\; t < \tau_\partial) \leq e_{\mathcal{T}}\, \mathrm{e}^{-\rho[t - t_A - t_F]} \mathbb{P}_x\left(U_H = \infty\,,\; t_F < \tau_\partial\right).$$

Thanks to Inequality (2.2), we can relate the decay $\mathrm{e}^{-\rho t}$ to $\mathbb{P}_\zeta(t < \tau_\partial)$. Thanks to $(A3_F)$ and recalling the definition of ϵ as $(2\, c_S\, e_\mathcal{T})^{-1}$, we then conclude the proof of Lemma 3.2 where:

$$C_F := \frac{e_\mathcal{T}\, \mathrm{e}^{\rho t_A}\, c_S\, \epsilon}{\mathrm{e}^{\rho\, t_S}\, \mathbb{P}_\zeta(t_S < \tau_\partial)} = \frac{\mathrm{e}^{\rho t_A}}{2\mathrm{e}^{\rho\, t_S}\, \mathbb{P}_\zeta(t_S < \tau_\partial)} > 0. \qquad \square$$

3.1.3. Step 3: we look at propagating some induction property

Let us consider the following random variable:

$$J(t) := \sup\left\{ j \geq 0\,;\ \tau_E^j < (t - t_A) \wedge \tau_\partial \wedge U_H^\infty \right\}, \tag{3.2}$$

where we exploit the sequence (τ_E^j) that is defined inductively as follows:

$$\tau_E^{i+1} := \inf\{s \geq \tau_E^i + t_F\,;\ X_s \in E\} \wedge \tau_\partial, \quad \text{and} \quad \tau_E^0 = 0. \tag{3.3}$$

Step 3 consists in proving the following lemma, at the core of the induction over the value of $J(t)$ presented in Step 4:

Lemma 3.3. *Assume that Ω is a path space and that Assumption (A2), Inequalities (2.2) and Assumption $(A3_F)$ hold with $\rho > \hat{\rho}_S$. If there exists $j \geq 0$ and $C_j > 0$ such that the following upper-bound holds for any $x \in E$ and any $t \geq t_A$:*

$$\mathbb{P}_x(t < \tau_\partial\,,\ J(t) = j) \leq C_j\, \mathbb{P}_\zeta(t < \tau_\partial), \tag{3.4}$$

then the inequality holds also in the next step $j+1$ as follows for any $x \in E$ and any $t \geq t_A$:

$$\mathbb{P}_x(t < \tau_\partial\,,\ J(t) = j+1) \leq \tfrac{C_j}{2}\, \mathbb{P}_\zeta(t < \tau_\partial). \tag{3.5}$$

Lemma 3.3 states that the extension of survival during the failed coupling procedure, for a time-length t_F, then outside E (before τ_E^F) is not sufficient to compensate much compared to the cost of the complementary event of failure, *i.e.* $\{\tau_\partial \wedge t_F \leq U_H\} = \{U_H = \infty\}$.

Proof: Recalling Definition (3.2), we note that τ_E^1 coïncides with τ_E^F on the event $\{J(t) \geq 1\}$. Thanks to Proposition 2.2 and to the Markov property, we deduce:

$$\begin{aligned}
&\mathbb{P}_x(t < \tau_\partial\,,\ J(t) = j+1) \\
&\quad = \mathbb{E}_x\left[\mathbb{P}_{X[\tau_E^1]}(t - \tau_E^1 < \widetilde{\tau}_\partial\,,\ \widetilde{J}(t - \tau_E^1) = j)\,;\ J(t) \geq 1\right] \\
&\quad \leq C_j\, \mathbb{E}_x\left[\mathbb{P}_\zeta(t - \tau_E^1 < \widetilde{\tau}_\partial)\,;\ J(t) \geq 1\right],
\end{aligned}$$

where we exploited (3.4) and the fact that $t - \tau_E^1 \geq t_A$ on the event $\{J(t) \geq 1\}$. Then, thanks to Inequality (2.2):

$$\mathbb{P}_x(t < \tau_\partial\,,\ J(t) = j+1) \leq C_j \cdot c_S \cdot \mathbb{P}_\zeta(t < \tau_\partial) \cdot \mathbb{E}_x\left[\exp(\rho\, \tau_E^1);\ J(t) \geq 1\right]. \tag{3.6}$$

We then decompose τ_E^1 into the sum of t_F and $\widetilde{\tau}_E$ to arrive at the next upper-bound on the expectation in the right-hand side:

$$\mathbb{E}_x\left[\exp(\rho\,\tau_E^1);\ J(t)\geq 1\right] \leq \mathbb{E}_x\Big[\mathbb{E}_{X_{t_F}}\Big(\exp(\rho\,\widetilde{\tau}_E)\,;\ \widetilde{\tau}_E < (t-t_F-t_A)\wedge\widetilde{\tau}_\partial\Big)\cdot\exp(\rho t_F)\,;\ U_H=\infty\,,\ t_F<\tau_\partial\Big]. \quad (3.7)$$

Now, thanks to Assumption ($A2$), the following upper-bound holds $a.s.$ on $\{U_H=\infty\,,\ t_F<\tau_\partial\}$:

$$\mathbb{E}_{X_{t_F}}\left(\exp(\rho\,\widetilde{\tau}_E)\,;\ \widetilde{\tau}_E<(t-t_F-t_A)\wedge\widetilde{\tau}_\partial\right)\leq e_{\mathcal{T}}. \quad (3.8)$$

The next upper-bound is derived for any $x\in E$ thanks to Assumption ($A3_F$):

$$\exp(\rho t_F)\cdot\mathbb{P}_x\left(U_H=\infty\,,\ t_F<\tau_\partial\right)\leq\epsilon=\frac{1}{2\,c_S\,e_{\mathcal{T}}}. \quad (3.9)$$

Combining the four Inequalities (3.6), (3.7), (3.8) and (3.9) yields the following upper-bound for any $x\in E$ and any $t\geq t_A$:

$$\mathbb{P}_x(t<\tau_\partial\,,\ J(t)=j+1)\leq\tfrac{C_j}{2}\,\mathbb{P}_\zeta(t<\tau_\partial),$$

which concludes the proof of Lemma 3.3. □

3.1.4. Concluding the proof of Theorem 2.3

With $C:=2\,(C_0+C_F)$, thanks to Lemmas 3.1, 3.2, the following upper-bound holds for any $x\in E$ and any $t\geq t_A$:

$$\mathbb{P}_x(t<\tau_\partial\,,\ J(t)=0)\leq(C/2)\cdot\mathbb{P}_\zeta(t<\tau_\partial).$$

By immediate induction thanks to Lemma 3.3, we deduce that the following upper-bound holds for any $x\in E$, any $t\geq t_A$ and any $j\in\mathbb{Z}_+$:

$$\mathbb{P}_x(t<\tau_\partial\,,\ J(t)=j)\leq 2^{-j-1}\,C\cdot\mathbb{P}_\zeta(t<\tau_\partial).$$

With this decomposition, we simply conclude the proof of Theorem 2.3 as follows:

$$\mathbb{P}_x(t<\tau_\partial)=\textstyle\sum_{\{j\geq0\}}\mathbb{P}_x(t<\tau_\partial\,,\ J(t)=j)\leq C\cdot\mathbb{P}_\zeta(t<\tau_\partial).$$ □

3.2. A more refined convergence result

Thanks to [63], Corollary 5.2.1 and to Theorem 2.3, Assumption (**A**) of [63] is nearly implied by our Assumption ($\mathbf{A_F}$), except that $\bigcup_{\ell\geq1}\mathcal{D}_\ell=\mathbb{X}$ is no longer assumed. We let the reader check that the proofs given in [63], Section 5 apply mutatis mutandis, except for the following facts:

- ⋆ As presented in several examples in [66], α might not be the unique QSD for $\mathbb{X}$. By adapting the argument given in the following proof of Proposition 2.10 we can nonetheless deduce that any other QSD ν must satisfy that: $\nu(\bigcup_\ell\mathcal{D}_\ell)=0$.
- ⋆ The lower-bound on $h(x)$ are only obtained for $x\in\mathcal{D}_\ell$, so that $h(x)$ might be equal to 0 for $x\in\mathbb{X}\setminus\bigcup_\ell\mathcal{D}_\ell$.
- ⋆ The reasoning on the Q-process can only be applied for initial conditions on $\mathcal{H}:=\{x\in\mathbb{X}\,;\ h(x)>0\}\supset\bigcup_\ell\mathcal{D}_\ell$.

- ⋆ The uniqueness of β as a stationary distribution for the Q-process holds among all distributions with support on $\mathcal{H}$ and not necessarily on $\mathcal{M}_1(\mathbb{X})$.

By adapting [63], Theorem 2.2, we deduce that there exists h, λ $\gamma, C_h > 0$ such that:

$$\forall\, t > 0,\ \|\mathrm{e}^{\lambda t}\mathbb{P}_{.}(t < \tau_\partial) - h\|_\infty \leq C_h\, \mathrm{e}^{-\gamma\, t}. \tag{3.10}$$

In addition, by adapting [63], Theorem 2.1, there exists also $\alpha \in \mathcal{M}_1(\mathbb{X})$ and a family of constants $C_\alpha(\ell, \xi) > 0$, defined for any $\ell \geq 1$ and $\xi \in (0, 1]$, such that the following property holds true for any $\mu \in \mathcal{M}_1\,(\mathbb{X})$ such that $\mu(\mathcal{D}_\ell) \geq \xi$ and any $t > 0$:

$$\|\, \mu A_t(\mathrm{d}x) - \alpha(\mathrm{d}x)\, \|_{TV} \leq C_\alpha(\ell, \xi)\, \mathrm{e}^{-\gamma\, t}. \tag{3.11}$$

Moreover, $\langle \alpha \,|\, h\rangle = 1$. The first condition on the Q-process is also directly deduced. It only remains to prove the convergence results as they are stated.

3.2.1. Proof of Theorem 2.8

We recall that μ is chosen in Definition 1 such that $\|\mu\|_{TV} \leq 1$. Denote by μ_+ (respectively μ_-) the positive (respectively negative) component of μ so that: $\mu = \mu_+ - \mu_-$ and $\|\mu\|_{TV} = 1 = \mu_+(\mathbb{X}) + \mu_-(\mathbb{X})$. Let $y \in \mathcal{D}_1$ and define the positive measures $\hat{\mu}_+$ and $\hat{\mu}_-$ as follows:

$$\hat{\mu}_+(\mathrm{d}x) := \frac{1}{1 + \mu_+(\mathbb{X})}\,[\delta_y + \mu_+(\mathrm{d}x)] \geq 0, \qquad \hat{\mu}_-(\mathrm{d}x) := \frac{1}{1 + \mu_-(\mathbb{X})}\,[\delta_y + \mu_-(\mathrm{d}x)] \geq 0.$$

Note that $\hat{\mu}_+(\mathbb{X}) = \hat{\mu}_-(\mathbb{X}) = 1$, so that both $\hat{\mu}_+$ and $\hat{\mu}_-$ are probability measures. In particular, this implies the following identification:

$$\hat{\mu}_+.(\mathrm{e}^{-t\,\lambda} P_t)(\mathrm{d}x) = \mathrm{e}^{-t\,\lambda} \cdot \hat{\mu}_+ P_t(\mathrm{d}x) = \langle \hat{\mu}_+ \,|\, h_t\rangle \cdot \hat{\mu}_+ A_t(\mathrm{d}x), \tag{3.12}$$

and similarly for $\hat{\mu}_-$. These measures are also constructed so as to satisfy the two following properties:

$$\mu = [1 + \mu_+(\mathbb{X})] \cdot \hat{\mu}_+ - [1 + \mu_-(\mathbb{X})] \cdot \hat{\mu}_-,$$
$$\hat{\mu}_+(\mathcal{D}_1) \wedge \hat{\mu}_-(\mathcal{D}_1) \geq \frac{1}{1 + \|\mu\|_{TV}} \geq 1/2.$$

Thus, first thanks to (3.11), then to (3.10), there exists $\gamma, C > 0$ independent from μ such that the four following upper-bounds hold on μ_+ and μ_- for any $t \geq 0$:

$$\|\, \hat{\mu}_+ A_t(\mathrm{d}x) - \alpha(\mathrm{d}x)\, \|_{TV} \leq C\, \mathrm{e}^{-\gamma\, t}, \quad |\, \hat{\mu}_- A_t(\mathrm{d}x) - \alpha(\mathrm{d}x)\, \|_{TV} \leq C\, \mathrm{e}^{-\gamma\, t},$$
$$|\langle \hat{\mu}_+ \,|\, h_t - h\rangle| \leq C\, \mathrm{e}^{-\gamma\, t}, \qquad |\langle \hat{\mu}_- \,|\, h_t - h\rangle| \leq C\, \mathrm{e}^{-\gamma\, t}.$$

The linear decomposition of μ between μ_+ and μ_- thus implies the next estimations, thanks also to Equation (3.12) and to the fact that h and the family h_t are uniformly bounded:

$$\begin{aligned}
\mu.(\mathrm{e}^{-t\,\lambda}\, P_t)(\mathrm{d}x) &= [1 + \mu_+(\mathbb{X})] \cdot \langle \hat{\mu}_+ \,|\, h_t\rangle \cdot \hat{\mu}_+.A_t(\mathrm{d}x) - [1 + \mu_-(\mathbb{X})] \cdot \langle \hat{\mu}_- \,|\, h_t\rangle \cdot \hat{\mu}_-.A_t(\mathrm{d}x) \\
&= \Big([1 + \mu_+(\mathbb{X})] \cdot \langle \hat{\mu}_+ \,|\, h\rangle - [1 + \mu_-(\mathbb{X})] \cdot \langle \hat{\mu}_- \,|\, h\rangle\Big) \cdot \alpha(\mathrm{d}x) + O_{TV}(\mathrm{e}^{-t\,\gamma}) \\
&= \langle \mu \,|\, h\rangle\, \alpha(\mathrm{d}x) + O_{TV}(\mathrm{e}^{-t\,\gamma}).
\end{aligned}$$

Thus, we conclude the proof of Theorem 2.8 in that there exists some $C' > 0$ such that the following upper-bound holds for any $t \geq 0$ and for any signed measure such that $\|\mu\|_{TV} \leq 1$:

$$\left\|\mu.(\mathrm{e}^{-t\,\lambda}\,P_t) - \langle \mu\,|\,h\rangle\,\alpha\right\|_{TV} \leq C'\,\exp[-t\,\gamma]. \qquad \square$$

3.2.2. Proof of Corollary 2.9

Let $\bar{\mu} = \mu - \alpha$. We recall that by definition, $\langle \alpha\,|\,h_t\rangle = \langle \alpha\,|\,h\rangle = 1$. Then we can express the difference between μA_t and α as follows in terms of $\bar{\mu}$ (recall also the analogous of Eq. (3.12) for μ)

$$\begin{aligned}\mu A_t - \alpha &= \frac{\exp[t\lambda]\,\mu P_t - \langle \mu\,|\,h\rangle\,\alpha}{\langle \mu\,|\,h\rangle} + \frac{\langle \mu\,|\,h - h_t\rangle}{\langle \mu\,|\,h\rangle\cdot\langle \mu\,|\,h_t\rangle}\,\exp[t\lambda]\,\mu P_t\\ &= \frac{\exp[t\lambda]\,\bar{\mu} P_t - \langle \bar{\mu}\,|\,h\rangle\,\alpha}{\langle \mu\,|\,h\rangle} + \frac{\langle \bar{\mu}\,|\,h - h_t\rangle}{\langle \mu\,|\,h\rangle}\,\mu A_t.\end{aligned}$$

Thanks to Theorem 2.8, this immediately implies the estimate on the convergence to α. $\square$

3.2.3. Proof of Proposition 2.10

From (H_0), we consider $H \subset \mathbb{X}$, $t, c > 0$ and $\ell \geq 1$ such that the following lower-bound holds for any $x \in H$:

$$\mathbb{P}_x(\tau_{\mathcal{D}_\ell} < t \wedge \tau_\partial) \geq c. \tag{3.13}$$

Thanks to Theorem 2.8, h is bounded away from zero by a positive constant on any $\mathcal{D}_\ell$, so let h_ℓ be such lower-bound. The property of h being an eigenfunction of the semi-group (P_t) can be rephrased by saying that $(h(X_t)\mathrm{e}^{\rho_0 t}\mathbf{1}_{\{t<\tau_\partial\}})$ is a martingale. We also recall that h is non-negative. The following upper-bound for any $x \in H$ is derived thanks to the martingale property, then the definition of h_ℓ and finally Inequality (3.13):

$$\begin{aligned}h(x) &= \mathbb{E}_x\left[h(X(\tau_{\mathcal{D}_\ell} \wedge t))\exp[\rho_0(\tau_{\mathcal{D}_\ell} \wedge t)]\,;\ \tau_{\mathcal{D}_\ell} \wedge t < \tau_\partial\right]\\ &\geq h_\ell \mathbb{P}_x(\tau_{\mathcal{D}_\ell} < t \wedge \tau_\partial) \geq c\cdot h_\ell > 0.\end{aligned}$$

This proves the uniform lower-bound of h on H.

For the second point, recalling that the sequence of stopping times $\tau_{\mathcal{D}_\ell}$ is necessarily decreasing, we deduce the following inclusion, that leads to the first desired inclusion:

$$\begin{aligned}\{x \in \mathbb{X}\,;\ \exists\,\ell \geq 1,\ \mathbb{P}_x(\tau_{\mathcal{D}_\ell} < \tau_\partial) > 0\} &= \textstyle\bigcup_{\ell \geq 1}\{x \in \mathbb{X}\,;\ \mathbb{P}_x(\tau_{\mathcal{D}_\ell} < \ell \wedge \tau_\partial) \geq 1/\ell\}\\ &\subset \{x \in \mathbb{X}; h(x) > 0\}.\end{aligned}$$

For the reciprocal inclusion, let any x be such that $h(x) > 0$. Thanks to Corollary 2.9, $\delta_x A_t$ converges to α. Choosing $\ell \geq 1$ by Theorem 2.8 such that $\alpha(\mathcal{D}_\ell) > 0$, it implies that $\delta_x A_t(\mathcal{D}_\ell) > 0$ for t sufficiently large, thus $\mathbb{P}_x(\tau_{\mathcal{D}_\ell} < t \wedge \tau_\partial) > 0$. This ends the proof of Proposition 2.10. $\square$

3.2.4. Proof of (ii) in Corollary 2.12

Assume that $\mu \in \mathcal{M}_1(\mathbb{X})$ satisfies $\langle \mu\,|\,1/h\rangle < \infty$. We may define $\nu(\mathrm{d}x) := \dfrac{\mu(\mathrm{d}x)}{h(x)\,\langle \mu\,|\,1/h\rangle}$, which trivially satisfies that $\nu \in \mathcal{M}_1(\mathbb{X})$ and that $\mu(\mathrm{d}x) = \langle \nu\,|\,h\rangle^{-1}\cdot h(x)\cdot \nu(\mathrm{d}x)$. Let $\bar{\nu} = \nu - \alpha$. The difference between μQ_t

and β is expressed as follows in terms of $\bar{\nu}$ and of a real-valued bounded measurable function f:

$$\begin{aligned}\langle \mu Q_t \,|\, f\rangle - \langle \beta \,|\, f\rangle &= \frac{\mathrm{e}^{t\lambda}\cdot\langle \nu P_t \,|\, h\cdot f\rangle - \langle \nu \,|\, h\rangle\cdot\langle \alpha \,|\, h\cdot f\rangle}{\langle \nu \,|\, h\rangle}\\ &= \frac{\mathrm{e}^{t\lambda}\cdot\langle \bar{\nu} P_t \,|\, h\cdot f\rangle - \langle \bar{\nu} \,|\, h\rangle\cdot\langle \alpha \,|\, h\cdot f\rangle}{\langle \nu \,|\, h\rangle}\\ &\leq \frac{\|\bar{\nu}\|_{TV}}{\langle \nu \,|\, h\rangle}\cdot\|h\|_\infty\cdot\|f\|_\infty\cdot\mathrm{e}^{-\gamma t},\end{aligned}$$

thanks to Theorem 2.8. Moreover, by the definition of β and ν:

$$\frac{\|\nu-\alpha\|_{TV}}{\langle \nu \,|\, h\rangle} = \left\|\frac{\mu}{\langle \mu \,|\, 1/h\rangle} - \beta\right\|_{1/h}\cdot\frac{\langle \mu \,|\, 1/h\rangle}{\langle \mu \,|\, \mathbf{1}\rangle} = \|\mu - \langle \mu \,|\, 1/h\rangle\,\beta\|_{1/h}.$$

Injecting this equality into the preceding upper-bound yields (ii) and concludes the proof of Corollary 2.12. □

3.3. Proof of Propositions 2.13

For $x\in\mathbb{X}$ and $t\geq 0$, define $\nu_{x,t}(\mathrm{d}y) := (\mathrm{e}^{\lambda t}\delta_x P_t - h(x)\alpha)_+(\mathrm{d}y)\geq 0$. Thanks to (2.7), we know that there exists $C,\gamma>0$ such that $\nu_{x,t}(\mathbb{X})\leq C\mathrm{e}^{-\gamma t}$. Let $\rho:=\lambda+\gamma/2$ and $t_E := 2\cdot[\gamma\cdot\log(4C)]^{-1}$. We thus ensure that $C\cdot\mathrm{e}^{-(\lambda+\gamma)t_E}\leq (1/4)\cdot\mathrm{e}^{-\rho t_E}$.

Since $\mathbb{X}=\bigcup_{\ell\geq 1}\mathcal{D}_\ell$, and denoting as $\mathcal{D}_\ell^c$ the complementary of $\mathcal{D}_\ell$, we choose ℓ sufficiently large to ensure $\alpha(\mathcal{D}_\ell^c)\leq (4\|h\|_\infty)^{-1}\cdot\exp(-\gamma\, t_E/2)$. Recalling the definition of ν_{x,t_E}, the above results imply the following upper-bound for any $x\in\mathbb{X}$:

$$\begin{aligned}\delta_x P_{t_E}(\mathcal{D}_\ell^c) &\leq \mathrm{e}^{-\lambda t_E} h(x)\alpha(\mathcal{D}_\ell^c) + \nu_{x,t_E}(\mathbb{X})\\ &\leq \mathrm{e}^{-\rho t_E}/2.\end{aligned}\tag{3.14}$$

We split $\mathbb{R}_+$ into time-intervals of length t_E, so as to decompose the following expectation for any $K\geq 2$:

$$\begin{aligned}&\mathbb{E}_x\left(\exp\left[\rho\left(\tau_\partial\wedge\tau_{\mathcal{D}_\ell}\wedge(Kt_E)\right)\right]\right)\\ &\qquad\leq \mathrm{e}^{\rho t_E} + \sum_{k=1}^{K-1}\exp([k+1]t_E)\cdot\mathbb{P}_x(\forall\, k'\in[\![1,k]\!],\ X_{k' t_E}\notin\mathcal{D}_\ell\,;\ kt_E<\tau_\partial).\end{aligned}\tag{3.15}$$

Thanks to the Markov property and to Inequality (3.14), an elementary induction proves the following upper-bound for any $k\geq 1$:

$$\mathbb{P}_x(\forall\, k'\in[\![1,k]\!],\ X_{k' t_E}\notin\mathcal{D}_\ell\,;\ kt_E<\tau_\partial)\leq\frac{\mathrm{e}^{-\rho k t_E}}{2^k}.\tag{3.16}$$

Thanks to Inequalities (3.15) and (3.16), we obtain the following upper-bound:

$$\mathbb{E}_x\left(\exp\left[\rho\left(\tau_\partial\wedge\tau_{\mathcal{D}_\ell}\wedge(Kt_E)\right)\right]\right)\leq \mathrm{e}^{\rho t_E}(1+\textstyle\sum_{k\geq 1}2^{-k})\leq 2\mathrm{e}^{\rho t_E}<\infty.$$

Letting K tend to infinity concludes the proof of Propositions 2.13. □

3.4. Proof of Propositions 2.14

Let $x\in\mathbb{X}$, $\mu\in\mathcal{M}_1(\mathbb{X})$, and a value $t_F>1$ that is to be sufficiently large.

Step 1: Definition of U_H

We define for any $t \geq 0$:

$$\nu^t_{x,\mu} := \left(\delta_x P_t - \frac{\|h\|_\infty}{\langle \mu \,|\, h\rangle} \mu P_t\right)_+ .$$

We impose that U_H takes values t_F or ∞ in such a way that $\nu^{t_F}_{x,\mu}$ exactly correspond to the harvested measure:

$$\mathbb{E}_x(X_{t_F} \in \mathrm{d}y\,;\, U_H = t_F < \tau_\partial) = \delta_x P_{t_F}(\mathrm{d}y) - \nu^{t_F}_{x,\mu}(\mathrm{d}y) \geq 0. \tag{3.17}$$

A natural choice of U_H is defined through U being a uniform random variable on $(0, 1)$, independent of the process X. Thanks to the Radon-Nikodym Theorem, $\nu^{t_F}_{x,\mu}$ has a density with respect to the measure $\delta_x P_{t_F}$, density that may be chosen to take values in $[0, 1]$. So we simply impose $U_H = \infty$ if $U \geq \frac{\partial \nu^{t_F}_{x,\mu}}{\partial \delta_x P_{t_F}}(X_{t_F})$, where $\frac{\partial \nu^{t_F}_{x,\mu}}{\partial \delta_x P_{t_F}}(x)$ denotes the above-mentioned density at location x, and $U_H = t_F$ otherwise. With this choice, (3.17) is satisfied.

To ensure that such a uniform random variable U can be expressed within the framework of a path space representation Ω, we may for instance couple the original process with an independent Brownian Motion (B_t) and consider Ω as the canonical description of the coupled process (X, B). Since $t_F \geq 1$, B_1 is a $\mathcal{F}_{t_F}$-measurable random variable, with a gaussian centered distribution. With F_G the cumulative function of this latter distribution, the variable $U := F_G^{-1}(B_1)$ satisfies our conditions: it is uniformly distributed on $[0, 1]$, $\mathcal{F}_{t_F}$-measurable and independent of X.

Step 2: Control of the densities

Thanks to the definitions of U_H and $\nu^{t_F}_{x,\mu}$, the following inequality is straightforward:

$$\mathbb{E}_x(X_{U_H} \in \mathrm{d}y\,;\, U_H < \tau_\partial) \leq \frac{\|h\|_\infty}{\langle \mu \,|\, h\rangle} \mathbb{E}_\mu(X_{t_F} \in \mathrm{d}y\,;\, t_F < \tau_\partial).$$

Step 3: Control of failures

With this definition of U_H, recalling (3.17):

$$\mathbb{P}_x(U_H = \infty, t_F < \tau_\partial) = 1 - (\delta_x P_{t_F}(\mathbb{X}) - \nu^{t_F}_{x,\mu}(\mathbb{X})) = \nu^{t_F}_{x,\mu}(\mathbb{X}), \tag{3.18}$$

which is thus the quantity for which we want an upper-bound.

Consider the following positive measure:

$$\hat{\nu}^{t_F}_{x,\mu} := \mathrm{e}^{-\lambda t_F} \cdot \left[\left(\mathrm{e}^{\lambda t_F} \delta_x P_{t_F} - h(x)\alpha\right)_+ + \frac{\|h\|_\infty}{\langle \mu \,|\, h\rangle} \cdot \left(\langle \mu \,|\, h\rangle \alpha - \mathrm{e}^{\lambda t_F} \mu P_{t_F}\right)_+\right].$$

Thanks to Inequality (2.7) with $\rho := \lambda + \gamma/2$, the mass of this measure can be efficiently upper-bounded:

$$\hat{\nu}^{t_F}_{x,\mu}(\mathbb{X}) \leq C\mathrm{e}^{-(\lambda+\gamma)t_F} \left(1 + \frac{\|h\|_\infty}{\langle \mu \,|\, h\rangle}\right) \leq \frac{c_F}{\langle \mu \,|\, h\rangle} \mathrm{e}^{-\rho t_F}, \tag{3.19}$$

where $c_F := 2C \cdot \|h\|_\infty$ is independent of x, μ and t_F.

On the other hand, $\hat{\nu}^{t_F}_{x,\mu}$ is such that the following property holds:

$$\delta_x P_{t_F}(\mathrm{d}y) \le \frac{\|h\|_\infty}{\langle \mu \,|\, h\rangle} \mu P_{t_F}(\mathrm{d}y) + \hat{\nu}^{t_F}_{x,\mu}(\mathrm{d}y),$$

which implies that $\nu^{t_F}_{x,\mu} \le \hat{\nu}^{t_F}_{x,\mu}$. Combining it with (3.18) and (3.19), the intended inequality is obtained:

$$\mathbb{P}_x(U_H = \infty, t_F < \tau_\partial) \le \frac{c_F}{\langle \mu \,|\, h\rangle} \mathrm{e}^{-\rho t_F}.$$

Since t_F is indeed allowed to take any sufficiently large values, this concludes the proof of Proposition 2.14. □

3.5. Proof of Theorem 2.16

Recall that we wish to describe the approximations of the previous dynamics when extinction happens at $\tau^L_\partial := \tau_\partial \wedge T_{\mathcal{D}_L}$ instead of τ_∂.

There is an explicit relation between all the constants introduced in the proofs of Theorems 2.8-4 (requiring also the proofs in [63]). Moreover, the proof actually relies on a single value of $\rho > \rho_S$ and a specific set E. Note that for any L such that $E \subset \mathcal{D}_L$, we have:

$$\sup_{\{x\in\mathbb{X}\}} \mathbb{E}_x\left(\exp\left[\rho\left(\tau^L_\partial \wedge \tau_E\right)\right]\right) \le \sup_{\{x\in\mathbb{X}\}} \mathbb{E}_x\left(\exp\left[\rho\left(\tau_\partial \wedge \tau_E\right)\right]\right) := e_{\mathcal{T}}.$$

Likewise, Assumption $(A3_F)$ extends naturally for τ^L_∂. $(A1)$ is stated with extinction already occurring at the exit of some set $\mathcal{D}_{L(\ell)}$ prescribed by the value of ℓ. Considering the proof of Theorem 2.8 in Sect. 3.2.1, we see that (3.3) and (3.4) are only required for $\ell = 1$ and $\xi = 1/2$. Once these two values are fixed, the proof given in [63], Section 5 treats uniformly initial conditions μ such that $\mu(\mathcal{D}_1) \ge 1/2$ by exploiting $(A1)$ a finite number of times. One can thus identify an upper-bound $L_\veebar \ge 1$ of the values $L(\ell)$ involved in the successive applications of $(A1)$. So it suffices to take L such that $\mathcal{D}_{L_\veebar} \subset \mathcal{D}_L$ to ensure that all the results extend for the extinction time τ^L_∂ instead of τ_∂. Under this condition, our proof ensures that uniform exponential quasi-stationary convergence also holds for the process with extinction at time τ^L_∂ and that the constants involved in the convergences can be taken uniformly over these values L.

Provided the subset $\mathcal{D}'$ of $\mathbb{X}$ is such that $\mathcal{D}_{L_\veebar} \subset \mathcal{D}'$, the arguments above holds with τ^L_∂ replaced by $\tau_\partial \wedge T_{\mathcal{D}'}$.

To compare λ to λ_L, we can observe that for any $t > 0$:

$$\frac{-1}{t} \log \mathbb{P}_\zeta(t < \tau_\partial) \le \frac{-1}{t} \log \mathbb{P}_\zeta(t < \tau^L_\partial),$$

so that $\lambda \le \lambda_L$ is deduced by taking the limit $t \to \infty$ and exploiting the convergence to the survival capacities. The same argument ensures that λ_L is a decreasing sequence in L.

Assume then by contradiction that there exists $\eta > 0$ such that $\lim_L \lambda_L \ge \lambda + \eta$. Recall that we have an explicit upper-bound $\|h_*\|_\infty$ that is valid uniformly for the functions h^L and h. Thanks to Theorem 2.8 and the analogous result with τ^L_∂, for t sufficiently large, the next two estimates on the survival probability holds:

$$\mathrm{e}^{\lambda t}\mathbb{P}_\zeta(t < \tau_\partial) \ge \frac{1}{2}\langle \zeta \,|\, h\rangle, \qquad \mathrm{e}^{\lambda_L t}\mathbb{P}_\zeta(t < \tau^L_\partial) \le 2\|h_*\|_\infty.$$

By exploiting the property of η, we deduce:

$$\begin{aligned} 0 < \langle \zeta \,|\, h\rangle &\le 2\mathrm{e}^{(\lambda-\lambda_L)t} \cdot \mathrm{e}^{\lambda_L t} \cdot \mathbb{P}_\zeta(t < \tau^L_\partial) + 2\mathrm{e}^{\lambda t}\,\left|\mathbb{P}_\zeta(t < \tau_\partial) - \mathbb{P}_\zeta(t < \tau^L_\partial)\right| \\ &\le 4\,\|h_*\|_\infty \mathrm{e}^{-\eta t} + 2\mathrm{e}^{\lambda t}\,\left|\mathbb{P}_\zeta(t < \tau_\partial) - \mathbb{P}_\zeta(t < \tau^L_\partial)\right|. \end{aligned} \tag{3.20}$$

The first term in the upper-bound becomes negligible uniformly over L by taking t sufficiently large. In order to obtain a contradiction, we merely have to prove that $\mathbb{P}_\zeta(t < \tau_\partial^L)$ converges to $\mathbb{P}_\zeta(t < \tau_\partial)$ at any fixed (large) time t. The difference is $\mathbb{P}_\zeta(T_{\mathcal{D}_L} < t < \tau_\partial)$ thus upper-bounded by $\mathbb{P}_\zeta(T_{\mathcal{D}_L} < t)$. So it suffices to prove that a.s. $\lim_L T_{\mathcal{D}_L} = \infty$.

Assume by contradiction that the limit T_∞ of this increasing sequence is at a finite value. Yet, thanks to ($\underline{A0}$) and to the fact that X is càd-làg, $X_{T_\infty -} \in \mathbb{X}$ belongs to $\mathcal{D}_M$ for some M. Thus, there exists a vicinity to the left of T_∞ on which $T_{\mathcal{D}_L}$ for $L > M$ cannot happen. But this precisely contradicts the definition of T_∞. Consequently, $\lim_L T_{\mathcal{D}_L} = \infty$ holds a.s. and $\mathbb{P}_\zeta(t < \tau_\partial^L)$ converges to $\mathbb{P}_\zeta(t < \tau_\partial)$ as $L \to \infty$. The contradiction with (3.20) makes us conclude that λ_L tends to λ as $L \to \infty$.

The next step is to look at the survival capacities, thanks again to Theorem 2.8 (with the measures evaluated on $\mathbb{X}$). The following upper-bound holds for any $x \in \mathbb{X}$:

$$|h(x) - h_L(x)| \leq \mathrm{e}^{\lambda t} |\mathbb{P}_x(t < \tau_\partial) - \mathbb{P}_x(t < \tau_\partial^L)| + |\mathrm{e}^{\lambda t} - \mathrm{e}^{\lambda_L t}| + C\,\mathrm{e}^{-\gamma t}.$$

Again, we can choose t sufficiently large to make $C\,\mathrm{e}^{-\gamma t}$ negligible. We already know that λ_L tends to λ and as previously, we prove that $\mathbb{P}_x(t < \tau_\partial^L)$ tends to $\mathbb{P}_x(t < \tau_\partial)$, as $L \to \infty$. This concludes the punctual convergence of h_L to h. The conclusion would be the same if one replaces x by any probability measure μ, for instance α.

Concerning the QSD:

$$\begin{aligned}\|\alpha - \alpha_L\|_{TV} \leq \left\|\mathrm{e}^{\lambda_L t} \delta_\alpha P_t^L - \langle \alpha \,|\, h_L \rangle \alpha_L \right\|_{TV} + |\mathrm{e}^{\lambda_L t} - \mathrm{e}^{\lambda t}| \\ + |\langle \alpha \,|\, h_L - h \rangle| + \mathrm{e}^{\lambda t} \left\| \delta_\alpha P_t^L - \delta_\alpha P_t \right\|_{TV},\end{aligned}$$

where as $L \to \infty$, for t fixed, we have just shown that the following quantity tends to 0 by proving that $\lim_L T_{\mathcal{D}_L} = \infty$ holds a.s.:

$$\left\| \delta_\alpha P_t^L - \delta_\alpha P_t \right\|_{TV} = \mathbb{P}_\alpha (T_{\mathcal{D}_L} < t < \tau_\partial) \to 0.$$

Thanks again to Theorem 2.8 and the previous convergence results to α_L, h and λ, the right-hand side can be made negligible by taking t then L sufficiently large, concluding the convergence of α_L to α in total variation. This concludes the proof of Theorem 2.16. □

Remark 3.4. More generally, thanks to the arguments above, there exists a sequence $(\epsilon_L)_{L \geq 1}$ converging to 0 such that the following properties hold for any subset $\mathcal{D}'$ of $\mathbb{X}$ that contains $\mathcal{D}_L$, with $L \geq L_\vee$ and λ', h' and α' the corresponding extinction rates, survival capacities and QSD restricted on $\mathcal{D}'$: $\lambda' \in [\lambda, \lambda_L]$, $\|h' - h\|_\infty \leq \epsilon_L$ and $\|\alpha' - \alpha\|_{TV} \leq \epsilon_L$. To deduce convergence, the sequence of subsets of $\mathbb{X}$ on which the localization is done is thus much more flexible than simply $(\mathcal{D}_L)$.

Applications

4. Mutations compensating a drift leading to extinction

4.1. A first simple process

We recall that we wish to prove uniform exponential quasi-stationary convergence for the following process :

$$X_t = x - v\,t\,\mathbf{e}_1 + \textstyle\sum_{\{i \leq N_t\}} W_i, \tag{4.1}$$

with a state-dependent extinction rate given by $\rho_e : x \mapsto \|x\|^2$. The number N_t of mutations at time t is given as a classical Poisson process on $\mathbb{R}_+$. Each mutation effect W_i is distributed as a normal variable with covariance matrix $\sigma^2 I_d$, and drawn independently of each others and of N_t. Between jumps, the process is translated at constant speed $v > 0$ along the first coordinate (*i.e.* along $\mathbf{e}_1$).

Theorem 4.1. *Consider P the semi-group associated to the process X as above (including the extinction). Then, for any $v, \sigma > 0$, P displays a uniform exponential quasi-stationary convergence with some characteristics $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathbb{R}^d) \times B(\mathbb{R}^d) \times \mathbb{R}_+$ (cf. Def. 2.6). Moreover, h is locally bounded away from 0.*

4.2. The main required properties

This application is related to non-local reaction-diffusion equations with a drift term. The one dimensional case has been studied recently by [32], with existence results obtained with compactness argument, and in [25], Section 2, with the use of Lyapunov functions.

To highlight the generality of our approach, we specify next the main properties of X that we exploit. We consider generally a càd-làg process X on $\mathbb{R}^d$, confronted to an extinction at a state-dependent rate given by $\rho_e : \mathbb{R}^d \mapsto \mathbb{R}_+$, and of the following form:

$$X_t = x - v\, t\, \mathbf{e}_1 + \int_{[0,t]\times\mathbb{R}^d\times\mathbb{R}_+} w\, \mathbf{1}_{\left\{u \le g(X_{s^-}, w)\right\}}\, M(\mathrm{d}s, \mathrm{d}w, \mathrm{d}u). \tag{4.2}$$

where M is a Poisson Random Measure (PRaMe) over $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+$, with intensity $\pi(\mathrm{d}s, \mathrm{d}w, \mathrm{d}u) = \mathrm{d}s\, \mathrm{d}w\, \mathrm{d}u$, while $g(x, w)$ describes the jump rate from x to $x + w$. In our focal example,

$$g(x, w) := \Big(\frac{1}{\sqrt{2\pi}\sigma}\Big)^d \cdot \exp\Big(-\frac{\|w\|^2}{2\sigma^2}\Big).$$

The infinitesimal generator $\mathcal{L}$ of such generic process is defined on all C^1 and bounded function f on $\mathbb{R}^d$ (thanks *e.g.* to [53], Cor. A.6), as follows:

$$\mathcal{L}f(x) := -v\partial_{x_1} f(x) + \int_{\mathbb{R}^d} (f(x+w) - f(x)) \cdot g(x, w)\mathrm{d}w - \rho_e(x) f(x).$$

The dynamics prescribed by the dual $\mathcal{L}^\star$ of $\mathcal{L}$ by the equation $\partial_t u = \mathcal{L}^\star u$, which is the starting point of [32], then corresponds to the dynamics in time of the density of the measure-valued process μP_t (see notably [25], Sect. 2).

The next properties are stated for these measurable functions g and ρ_e, with $B(x, r)$ the open ball around x of radius r for the Euclidian norm. The fact that an upper-bound holds locally in $\mathbb{X}$ means that for any compact subset K of $\mathbb{X}$, the upper-bound holds uniformly for $x \in K$.

Assumption (P): (for Piecewise-Deterministic)

- **(P1)** ρ_e is locally upper-bounded and $\lim_{\|x\|\to\infty} \rho_e(x) = +\infty$.
  Also, explosion implies extinction: $\tau_\partial \le \sup_{\{\ell\ge 1\}} T_{\mathcal{D}_\ell}$.
- **(P2)** The jump-rate $\rho_J(x) := \int_{\mathbb{R}^d} g(x, w)\, \mathrm{d}w$ is locally upper-bounded.
- **(P3)** Locally in $\mathbb{X}$, there exists $0 < \eta < a$ such that the restriction of g to $\mathbb{X} \times B(a \cdot \mathbf{e}_1, \eta)$ is lower-bounded.
- **(P4)** The jump size is tight locally in $\mathbb{X}$.
- **(P5)** The density for each jump vector w is upper-bounded locally in $\mathbb{X}$.

Theorem 4.2. *Provided the above conditions (**P**) are satisfied, P displays a uniform exponential quasi-stationary convergence with some characteristics $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathbb{R}^d) \times B(\mathbb{R}^d) \times \mathbb{R}_+$. Moreover, h is locally bounded away from 0.*

Besides, the Q-process exists and is exponentially ergodic with weight $1/h$ as stated in Corollary 2.12. Also, the uniformity in the localization procedure holds as stated in Theorem 2.16 for the restriction to any closed subset $\mathcal{D}'$ that contains $B(0, L)$, $L > 0$ being sufficiently large.

The proof of Theorem 4.2 is given in Sect. 4.3. It entails the proof of Theorem 4.1 since it is elementary that the process described in (4.1) satisfies (**P**). Let us nonetheless first clarify the meaning of these different properties.

To fix ideas on the role of condition (**P**), let us consider any compact subset K of $\mathbb{X}$. By virtue of Property (**P**) there exists $\rho_\vee$ such that $\rho_e(x) \leq \rho_\vee$ for any $x \in K$. This makes it possible to have simple lower-bounds on the survival of a given trajectory of $(X_t)_{t\geq 0}$ provided it remains confined in the compact set K. On the other hand, the fact that $\rho_e(x)$ tends to infnity as x tends to infinity makes it possible to justify the complementary of $B(0, \ell)$ as transitory for ℓ sufficiently large (according to property $(A2)$).

By virtue of Property (**P2**), there exists $\rho_J^\vee > 0$ such that for any $x \in K$, $\rho_J(x) \leq \rho_J^\vee$. We can thus consider events of positive probabilities such that the corresponding trajectories of X have no other jumps than the one we carefully describe.

By virtue of Property (**P3**), there exists also $g_\wedge > 0$ and $0 < \eta < a$ such that the following inequality holds for any $x \in K$ and any $w \in B(a\,\mathbf{e}_1, \eta)$, $g(x, w) \geq g_\wedge$. This will make it possible to consider trajectories in which each jump compensate the drift of the previous time-interval, with a small variation.

In addition, by virtue of Property (**P4**), for any $\epsilon > 0$, there exists $w_\vee$ such that the following upper-bound holds uniformly in $x \in K$: $\int_{\mathbb{R}^d} g(x, w)\, \mathbf{1}_{\{\|w\| \geq w_\vee\}}\, \mathrm{d}w \leq \epsilon \cdot \rho_J(x)$. This makes it possible to restrict the size of the jumps with a probability close to 1.

Finally, by virtue of Property (**P5**), there exists $g_\vee$ such that the following upper-bound holds uniformly in $x \in K$ and $w \in \mathbb{R}^d$: $g(x, w) \leq g_\vee \cdot \rho_J(x)$. Thanks to this property, we will deduce upper-bounds of the marginal density of X_t with respect to the Lebesgue measure after some jumps.

4.3. Proof of Theorem 4.2

We aim at proving Assumption $(\mathbf{A_F})$ for the sequence $\mathcal{D}_\ell := B(0, \ell \cdot a)$.

$(A0)$ is clearly satisfied. The proof of $(A1)$ is deduced from the following proposition, whose proof is deferred to the end of this subsection:

Proposition 4.3. *Under* $(\mathbf{P1, 2, 3})$*, for any $\ell \geq 1$, with $L := \ell + 2$, there exists $c, t > 0$ such that:*

$$\forall\, x \in \mathcal{D}_\ell, \qquad \mathbb{P}_x\left[X_t \in \mathrm{d}x\, ;\; t < \tau_\partial \wedge T_{\mathcal{D}_L}\right] \geq c\, \mathbf{1}_{\mathcal{D}_\ell}(\mathrm{d}x).$$

In particular, it implies that Assumption $(A1)$ holds with ζ uniform over $\mathcal{D}_1$. From Lemma 3.0.2 in [63], noting that $\zeta(\mathcal{D}_1) > 0$ in particular, we can (explicitly) deduce a strict upper-bound ρ of ρ_S. Since the extinction rate outside of $\mathcal{D}_\ell$ tends to infinity while $\ell \to \infty$, for any $\rho > 0$, we can find some $L \geq 1$ such that Assumption $(A2)$ holds true for $E := \mathcal{D}_\ell$ (*cf.* Sect. 4.1.2 in [63]). The proof of Assumption $(A3_F)$ for these choices is a clear consequence of the next proposition, whose proof is given in the next Sect. 4.3.1:

Proposition 4.4. *Suppose that Assumption* (**P**) *holds true. Consider any $\rho_E > \hat{\rho}_S$ and $\ell_E \geq 1$ such that the set $E = \bar{B}(0, \ell_E)$ satisfies $\forall\, y \notin E$, $\rho_e(y) \geq \rho_E$. Set also any $\epsilon > 0$. Then, there exists $t_F, t_V, c > 0$, such that for any $x \in E$, there exists a stopping time U_H with the following properties:*

$$\begin{aligned} \mathbb{P}_x(U_H = \infty,\; t_F < \tau_\partial) \leq \epsilon \cdot \exp[\hat{\rho}_S\, t_F], \qquad &\textit{where } \{\tau_\partial \wedge t_F < U_H\} = \{U_H = \infty\}, \\ \mathbb{P}_x(X(U_H) \in \mathrm{d}x'\, ;\; U_H < \tau_\partial) \leq c\, \mathbb{P}_\zeta\big(&X(t_V) \in \mathrm{d}x'\, ;\; t_V < \tau_\partial\big). \end{aligned} \tag{4.3}$$

Given these two propositions 4.3 and 4.4, we can conclude that Assumption $(\mathbf{A_F})$ holds true. Thanks to Theorem 2.8, Corollary 2.12 and Theorem 2.16, this directly implies Theorem 4.2.

4.3.1. Proof of Proposition 4.4

With the notations of the proposition, we first define t_F by the relation:

$$\exp[\hat{\rho}_S \cdot (2\ \ell_E/v) - (\rho_E - \hat{\rho}_S) \cdot (t_F - 2\ \ell_E/v)] = \epsilon/2.$$

The left-hand side is decreasing and converges to 0 when $t_F \to \infty$, so that t_F is well-defined. Let T_J be the first jump time of X. On the event $\{t_F < T_J\}$, we set $U_H = \infty$. The choice of t_F is done to ensure that the probability associated to the failure is indeed exceptional enough (with threshold $\epsilon/2$ and penalty-rate $\hat{\rho}_S$). Any jump occurring before t_F occurs from a position $X(T_J-) \in \bar{B}(0, \ell_E + v\, t_F) := K$. By virtue of Assumption (**P4**), we can then define $w_\vee$ such that:

$$\forall\, x \in K, \quad \int_{\mathbb{R}^d} g(x, w)\, \mathbf{1}_{\{\|w\| \ge w_\vee\}}\, \mathrm{d}w \le \epsilon/2 \cdot \exp[\hat{\rho}_S\, t_F].$$

A jump size larger than $w_\vee$ is then the other criterion of failure.

On the event $\{T_J \le t_F\} \cap \{T_J < \tau_\partial\} \cap \{\|W\| \le w_\vee\}$, where W is the size of the first jump (at time T_J), we thus set $U_H := T_J \le t_F$. Otherwise $U_H := \infty$.

In particular, $\{\tau_\partial \wedge t_F < U_H\} = \{U_H = \infty\}$ is clearly satisfied.

We prove next that the failures are indeed exceptional enough:

$$\mathbb{P}_x(U_H = \infty,\ t_F < \tau_\partial) \le \mathbb{P}_x(t_F \le T_J \wedge \tau_\partial) + \mathbb{P}_x(T_J < \tau_\partial \wedge t_F\,,\ \|W\| > w_\vee).$$

By the definition of $w_\vee$, we deal with the second term:

$$\begin{aligned}\mathbb{P}_x(T_J < \tau_\partial \wedge t_F\,,\ \|W\| > w_\vee) &\le \mathbb{P}_x(\|W\| > w_\vee \,|\, T_J < \tau_\partial \wedge t_F)\\ &\le \epsilon/2 \cdot \exp[\hat{\rho}_S\, t_F].\end{aligned}$$

On the event $\{t_F \le T_J \wedge \tau_\partial\}$ it holds a.s. for any $t \le t_F$ that $X_t = x - v\, t\, \mathbf{e}_1$. Thus, X is outside of E in the time-interval $[(2\ell_E/v), t_F]$, with an extinction rate at least ρ_E. By the definition of t_F, it implies:

$$\mathbb{P}_x(t_F \le T_J \wedge \tau_\partial) \le \exp[-\rho_E\ (t_F - 2\ \ell_E/v)] \le \epsilon/2 \cdot \exp[\hat{\rho}_S\, t_F].$$

This concludes (4.3).

On the other hand, recall that a.s. on the event $\{U_H < \infty\}$, $X_{U_H} = W + X(T_J-)$ where $X(T_J-) \in K$. Thus, by virtue of Assumption (**P5**), there exists $g_\vee > 0$ such that the following upper-bound on the density of $X(U_H)$ holds uniformly in $x \in E$:

$$\mathbb{P}_x(X(U_H) \in \mathrm{d}x'\,;\ T_J < t_F \wedge \tau_\partial\,,\ \|W\| \le w_\vee) \le g_\vee\, \mathbf{1}_{\left\{x' \in \bar{B}(0, \ell_E + v\, t_F + w_\vee)\right\}} \mathrm{d}x'.$$

We know also thanks to Proposition 4.3 that there exists $t_M, c_M > 0$ such that:

$$\mathbb{P}_\zeta\big(X(t_M) \in \mathrm{d}x\,;\ t_M < \tau_\partial\big) \ge c_M\, \mathbf{1}_{\left\{x' \in \bar{B}(0, \ell_E + v\, t_F + w_\vee)\right\}} \mathrm{d}x'.$$

With $t := t_M$, $c := g_\vee / c_M$ and thanks to Inequality (4.3), this concludes the proof of Proposition 4.4. □

4.3.2. Proof of Proposition 4.3

We consider a characteristic length of dispersion given by $r := \eta/4$. Given some $\ell \ge 1$, $x_I \in \mathcal{D}_\ell$, $L := \ell + 2$ and $t, c > 0$, we propose the following definition of the range of positions that can be reasonably reached by the process at time t starting from the vicinity of $x_I \in \mathcal{D}_\ell$, with density lower-bounded by c and restricted on specific domains (of the form $\mathcal{D}_L$):

$$\mathcal{R}^{(L)}(x_I, t, c) := \Big\{ x_F \in \mathbb{R}^d \,;\, \forall x_0 \in B(x_I, r),\ \mathbb{P}_{x_0}\left(X_t \in \mathrm{d}x \,;\ t < T_{\mathcal{D}_L} \wedge \tau_\partial\right) \geq c\, \mathbf{1}_{B(x_F, r)}(x)\, \mathrm{d}x \Big\}. \tag{4.4}$$

The proof then relies on the next two elementary lemmas.

Lemma 4.5. *Given any $L \geq 3$, there exists $c_0, t_0 > 0$, such that for any $x_I \in \mathcal{D}_\ell$:*

$$B(x_I, r) \subset \mathcal{R}^{(L)}(x_I, t_0, c_0).$$

Lemma 4.5 is proved in Sect. 4.3.3 by compensating the drift component with exactly one jump and adjusting the time to the allowed variations in jump vector. Thanks to Lemma 4.5, we can initiate small paths with time-frame t_0 and an efficiency c_0 chosen uniformly over any initial position in $\mathcal{D}_\ell$. As the second step formalized in the next Lemma 4.5, we show that we may expand the range in the feature dimension $\mathbb{X}$, at the expense of a specific increase of time and reduction of the reference density:

Lemma 4.6. *For any $L \geq 3$, with $\ell = L - 2$, there exists $t_a, c_a > 0$ such that the following implication holds uniformly for any $x_I, x \in \mathcal{D}_\ell$ and any $t, c > 0$:*

$$x \in \mathcal{R}^{(L)}(x_I, t, c) \quad \Rightarrow \quad B(x, r) \subset \mathcal{R}^{(L)}(x_I, t + t_a, c \cdot c_a).$$

Lemma 4.6 is proved in Sect. 4.3.4, as a corrolary of Lemma 4.5. The proof of Proposition 4.3 will then be achieved at the end of the third step, by expanding the range obtained in Lemma 4.5, inductively thanks to Lemma 4.6.

4.3.3. Step 1 : Initialisation - proof of Lemma 4.5

Thanks to Assumptions (**P2, 3**) (recalling that $r = \eta/4$ as stated in (**P3**)), there exists $g_\wedge > 0$ such that the following lower-bound hold for any $x \in \mathcal{D}_L$ and $w \in B(a \cdot \mathbf{e}_1,\ 4r)$:

$$\rho_J(x)^{-1} \cdot g(x, w) \geq g_\wedge. \tag{4.5}$$

Let T_1^J, T_2^J be respectively the first and second time of jump of X. Thanks to Assumptions (**P1, 2**), exploiting that $\mathcal{D}_L$ is convex, there exists $p_\wedge > 0$ such that the following lower-bound holds for any $x \in \mathcal{D}_L$ and $t \geq 0$ such that $x - v \cdot t \cdot \mathbf{e}_1 \in \mathcal{D}_L$:

$$\mathbb{P}_x(t < T_1^J \wedge \tau_\partial) = \exp[-\textstyle\int_0^t [\rho_J + \rho_e](x - v \cdot s \cdot \mathbf{e}_1)\mathrm{d}s] \geq p_\wedge. \tag{4.6}$$

On the other hand, thanks to Assumption (**P1, 3**), there exists also $q_\wedge > 0$ such that the following lower-bound holds for any $x \in \mathcal{D}_L$ and $t \geq a/v$ such that $x - v \cdot t \cdot \mathbf{e}_1 \in \mathcal{D}_L$:

$$\mathbb{P}_x(T_1^J < t \wedge \tau_\partial) \geq q_\wedge. \tag{4.7}$$

Let $t_0 := a/v,\ x_0 \in B(x_I, r)$. Concerning the constraint $t_0 < T_{\mathcal{D}_L}$, note that the following set is part of $B(x_I, 6r)$:

$$\mathcal{A} := \{x_0 - v \cdot s \cdot \mathbf{e}_1 \,;\ s \leq t_0\} \cup \{x_0 - v \cdot s \cdot \mathbf{e}_1 + w \,;\ s \leq t_0\,,\ w \in B(a \cdot \mathbf{e}_1,\ 4r)\}\,.$$

Since in addition $x_I \in \mathcal{D}_\ell = B(0, \ell \cdot a)$, $6r = 3\eta/2 \leq 2a$ and $L = \ell + 2$, this set is itself a subset of $\mathcal{D}_L$. Thus, by imposing at most one such jump, with a jump effect $w \in B(a \cdot \mathbf{e}_1,\ 4r)$, we keep the process inside of $\mathcal{D}_L$. Let us denote by $W = \Delta X_{T_1^J}$ the size of the first jump. We therefore restricts our analysis to the following event:

$$\mathcal{J} := \{T_1^J < t_0 < T_2^J \wedge \tau_\partial\} \cap \{W \in B(a \cdot \mathbf{e}_1,\ 4r)\}. \tag{4.8}$$

Both $t_0 < T_{\mathcal{D}_L}$ and $X(t_0) = x_0 - v \cdot t_0 \cdot \mathbf{e}_1 + W$ hold a.s. on $\mathcal{E}$. It implies the following lower-bound for any real-valued positive test function f on $\mathbb{X}$:

$$\mathbb{E}_{x_0}\left(f[X(t_0)]\,;\ t_0 < T_{\mathcal{D}_L} \wedge \tau_\partial\right) \geq \mathbb{E}_{x_0}\left(f[x_0 - v \cdot t_0 \cdot \mathbf{e}_1 + W]\,;\ \mathcal{J}\right). \tag{4.9}$$

Let $Y := X(T_1^J-)$ be the position of the process just before the first jump. Given the definition of $p_\wedge$ in (4.6) and since $\mathcal{A} \subset \mathcal{D}_L$, the following lower-bound holds a.s. on the event $\{T_1^J < t_0 \wedge \tau_\partial\} \cap \{W \in B(a \cdot \mathbf{e}_1,\ 4r)\}$:

$$\mathbb{P}_Y(t_0 - T_1^J < \widetilde{T}_1^J \wedge \widetilde{\tau}_\partial) \geq p_\wedge.$$

In combination with (4.9) and the Markov's inequality, this implies the next lower-bound:

$$\mathbb{E}_{x_0}\left(f[X(t_0)]\,;\ t_0 < T_{\mathcal{D}_L} \wedge \tau_\partial\right) \ \geq\ p_\wedge \mathbb{E}_{x_0}\left(f[x_0 - v \cdot t_0 \cdot \mathbf{e}_1 + W]\,;\ T_1^J < t_0 \wedge \tau_\partial, W \in B(a \cdot \mathbf{e}_1,\ 4r)\right). \tag{4.10}$$

Conditionally on Y (or equivalently on the time T_1^J), the law of W is given by the measure $\rho_J(Y)^{-1} \cdot g(Y, w)\mathrm{d}w$. Given the definition of $g_\wedge$ in (4.5), the law of W is lower-bounded by the measure $g_\wedge \cdot \mathbf{1}_{B(a \cdot \mathbf{e}_1,\ 4r)}(w)\mathrm{d}w$. Given also the definition of $q_\wedge$ in (4.7), we can derive from (4.10) the next lower-bound:

$$\mathbb{E}_{x_0}\left(f[X(t_0)]\,;\ t_0 < T_{\mathcal{D}_L} \wedge \tau_\partial\right) \geq p_\wedge \cdot g_\wedge \cdot q_\wedge \cdot \int_{B(a \cdot \mathbf{e}_1,\ 4r)} f[x_0 - v \cdot t_0 \cdot \mathbf{e}_1 + w]\mathrm{d}w. \tag{4.11}$$

Recalling that $x_0 \in B(x_I, r)$ and $t_0 = a/v$, we note that the following inclusions are valid for any $x_F \in B(x_I, r)$:

$$B(x_F, r) \subset B(x_I, 2r) \subset x_0 - v \cdot t_0 \cdot \mathbf{e}_1 + B(a \cdot \mathbf{e}_1,\ 4r).$$

With the fact that (4.11) holds for any positive test-function f, this concludes the proof of Lemma 4.5 where $c_0 := p_\wedge \cdot g_\wedge \cdot q_\wedge > 0$. □

4.3.4. Step 2 : Expansion - proof of Lemma 4.6

By definition, the fact that $\hat{x} \in \mathcal{R}^{(L)}(x_I, t, c)$ means that the following lower-bound holds uniformly in $x_0 \in B(x_I, r)$:

$$\mathbb{P}_{x_0}\left(X_t \in \mathrm{d}x_1\,;\ t < T_{\mathcal{D}_L} \wedge \tau_\partial\right) \geq c\, \mathbf{1}_{B(\hat{x}, r)}(x_1)\, \mathrm{d}x_1. \tag{4.12}$$

We see in the above proof of Lemma 4.5 that the definitions of t_0, c_0 can be stated in terms of L uniformly in x_I within $\mathcal{D}_\ell$. In particular, applying Lemma 4.5 to $\hat{x}$ instead of x_I, we deduce the following lower-bound for any $\hat{x} \in \mathcal{D}_\ell$ and any $x_1, x_F \in B(\hat{x}, r)$:

$$\mathbb{P}_{x_1}(X(t_0) \in \mathrm{d}x\,;\ t_0 < T_{\mathcal{D}_L} \wedge \tau_\partial) \geq c_0\, \mathbf{1}_{B(x_F, r)}(x)\, \mathrm{d}x. \tag{4.13}$$

In combination with the Markov property, (4.12) and (4.13) imply the following lower-bound uniformly in $x_0 \in B(x_I, r)$:

$$\mathbb{P}_{x_0}\left(X(t + t_0) \in \mathrm{d}x\,;\ t + t_0 < T_{\mathcal{D}_L} \wedge \tau_\partial\right) \geq c \cdot c_0 \cdot Leb(B(\hat{x}, r)) \cdot \mathbf{1}_{B(x_F, r)}(x)\, \mathrm{d}x.$$

Note that $c_a = c_0 \cdot Leb(B(\hat{x}, r)) = c_0 \cdot r^d \cdot Leb(B(0,1))$ is a positive constant independent of $\hat{x}$. With $t_a = t_0$, this implies that $B(\hat{x}, r) \subset \mathcal{R}^{(L)}(x_I, t + t_a, c \cdot c_a)$, concluding the proof of Lemma 4.6. □

4.3.5. Step 3: Conclude the proof of Proposition 4.3

Thanks to Lemma 4.5, there exists (t_0, c_0) such that, for any $x_I \in \mathcal{D}_\ell$, $x_I \in \mathcal{R}^{(L)}(x_I, t_0, c_0)$. By an immediate induction thanks to Lemma 4.6, we deduce that for any $k \geq 1$, there exist $t_k, c_k > 0$ such that $B(x_I, k \cdot r) \cap \mathcal{D}_\ell \subset \mathcal{R}^{(L)}(x_I, t_k, c_k)$. Note that both t_k and c_k do not depend on the particular choice of x_I. For k sufficiently large, the left-hand side of the inclusion is equal to $\mathcal{D}_\ell$ whatever $x_I \in \mathcal{D}_\ell$. This completes the proof of Proposition 4.3. □

Now that Propositions 4.4 and 4.3 are proved, as mentioned just after their statements, the proof of Theorem 4.2 is achieved. As noted after the statement of Theorem 4.2, this also ends the proof of Theorem 4.1.

5. The case of jumps occurring as in a Gibbs sampler

5.1. The core typical example

X is a pure jump process on $\mathbb{R}^d$, for $d \geq 2$, confronted to an extinction rate at state x given by $\rho_e(x) := \|x\|_\infty^2$, where $\|x\|_\infty := \sup_{\{i \leq d\}} |x_i|$. Jumps are restricted to happen along the vectors of an orthonormal basis $(\mathbf{e}_1, ..., \mathbf{e}_d)$. Independently of these directions and of previous jumps, each jump occurs at rate 1 and its size follows an exponential distribution with mean σ. This entails the following representation:

$$X_t := x + \sum\nolimits_{\{i \leq N_t\}} \sigma W_i \, \mathbf{e}_{D_i}. \tag{5.1}$$

In this formula, x is the initial condition, N_t a standard Poisson process on $\mathbb{Z}_+$, while, for any $i \geq 1$, W_i is a standard normal random variable on $\mathbb{R}$, D_i is uniform over $[\![1, d]\!]$. Moreover, all these random variables are independent from each others.

Theorem 5.1. *Consider P the semi-group associated to the process X given by equation (5.1) and weighted by the extinction event at rate ρ_e. Assume that $\sigma \leq 1/8$. Then, P displays a uniform exponential quasi-stationary convergence with some characteristics $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathbb{R}^d) \times B(\mathbb{R}^d) \times \mathbb{R}_+$ (cf. Def. 2.6). Moreover, h is locally lower-bounded.*

5.2. The main required properties

The process under consideration in Section 5 is a specific instance of pure jump processes. We refer to [48] for a detailed presentation of existence results of a QSD for a pure jump process. Contrary to the former approaches given in [13, 28, 29, 31, 57] or [58] and relying on adaptations of the Krein-Rutman theorem, the ones in [2, 48] exploits some maximum principle, which makes it possible to obtain uniqueness. No estimation of spectral gap, that is quantitative result of convergence, appeared to be known before [2]. This result apply to specific non-local kernels, among symmetric convolution kernels with positive density with respect to the Lebesgue measure on $\mathbb{R}^d$.

To our knowledge, the restriction of having jumps only along specific directions seems not to have been analyzed until the current article. As a motivation, the process X could for instance characterize an ecosystem where each coordinate corresponds to a single species.

Remark 5.2. In order to prevent concentration effects, several assumptions are additionally provided by the authors, see also [7]. Proposition 2.14 suggests to look in more detail into their connection to our assumption of "almost perfect harvest", which is also meant for this purpose.

To highlight the generality of our approach, we specify also in this case the main properties that we exploit. Let $(X_t)_{t\geq 0}$ be the pure jump process on $\mathbb{X} := \mathbb{R}^d$ defined by:

$$X_t := x + \sum_{i\leq d} \int_{[0,t]\times\mathbb{R}\times\mathbb{R}_+} w\, \mathbf{e}_i\, \mathbf{1}_{\left\{u\leq g_i(X_{s^-},w)\right\}} M_i(\mathrm{d}s,\, \mathrm{d}w,\, \mathrm{d}u), \tag{5.2}$$

where M_i are mutually independent PRaMes on $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$ with intensities $\mathrm{d}s\,\mathrm{d}w\,\mathrm{d}u$, and the $(g_i)_{i\in[\![1,d]\!]}$ are real-valued measurable function on $\mathbb{R}^d \times \mathbb{R}$. The process is also associated to a state-dependent extinction rate given by $\rho_e : \mathbb{R}^d \mapsto \mathbb{R}_+$.

In our focal example, $\rho_e(x) := \|x\|_\infty^2$ and g_i is defined as follows for any $i \leq d$:

$$g_i(x, w) := \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\tfrac{|w|^2}{2\sigma^2}\Big).$$

Remark that the infinitesimal generator $\mathscr{M}$ of such generic process is defined on all bounded-measurable function f on $\mathbb{R}^d$ (thanks *e.g.* to [53], Cor. A.6), as follows:

$$\mathscr{M} f(x) := \textstyle\sum_{i\leq d} \int_{\mathbb{R}^d} (f(x + w\mathbf{e}_i) - f(x)) \cdot g_i(x, w)\mathrm{d}w - \rho_e(x) f(x).$$

$(X_t)_{t\geq 0}$ is a Markov Process with piecewise constant trajectories. Conditionally upon $X_t = x$, the waiting-time and size of the next jump are independent, the law of the waiting-time is exponential of rate $\rho_J(x) := \sum_{i\leq d} \rho_J^i(x)$, where the jump rate can be decomposed along each direction $i \in [\![1, d]\!]$ according to $\rho_J^i(x) := \int_{\mathbb{R}^d} g_i(x, w)\ \mathrm{d}w < \infty$. The jump occurs on the i-th coordinate with probability $\rho_J^i(x)/\rho_J(x)$, then with size given by $g_i(x, w)/\rho_J^i(x)\,\mathrm{d}w$.

- **Assumption (J)**: (for Jumps)

(**J1**) The global jump rate ρ_J is upper-bounded locally in $\mathbb{X}$.
(**J2**) Locally in $\mathbb{X}$, there exists $\eta > 0$ such that the restriction of g to $\mathbb{X} \times B(0, \eta)$ is lower-bounded.
(**J3**) The jump size has a tight law locally in $\mathbb{X}$.
(**J4**) The density for each jump vector is upper-bounded locally in $\mathbb{X}$.
(**J5**) The probability that each direction gets involved in the jump is uniformly lower-bounded.
(**J6**) ρ_e is bounded away from zero by $\rho > \rho_S$ outside some compact set. Moreover, ρ_e is locally bounded and explosion implies extinction: $\tau_\partial \leq \sup_{\{\ell\geq 1\}} T_{\mathcal{D}_\ell}$.
(**J7**) $\rho_S < \rho_F$ where $\rho_F := \inf_{\{x\in\mathbb{R}^d,\, i\leq d\}} \left\{\rho_J^i(x) + \rho_e(x)\right\}$.

Theorem 5.3. *Provided the above conditions* (**J**) *are satisfied, P displays a uniform exponential quasi-stationary convergence with some characteristics $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathbb{R}^d) \times B(\mathbb{R}^d) \times \mathbb{R}_+$. Moreover, h is locally bounded away from* 0.

Besides, the Q-process exists and is exponentially ergodic with weight $1/h$ as stated in Corollary 2.12. Also, the uniformity in the localization procedure holds as stated in Theorem 2.16 for the restriction to any closed subset $\mathcal{D}'$ that contains $B(0, L)$, $L > 0$ being sufficiently large.

We also refer to Sects. 6.3.2 for the connection with reaction-diffusion equations, which holds in the same way, this time for a non-local dispersion operator of the following form:

$$\mathscr{M}^\star u(x) := \sum_{i\leq d} \left[\int_{\mathbb{R}} g_i(x - w_i\,\mathbf{e}_i,\, w_i)\, u(x - w_i\,\mathbf{e}_i)\, \mathrm{d}w_i - \left(\int_{\mathbb{R}} g_i(x, w_i)\, \mathrm{d}w_i\right) u(x)\right]$$

Theorem 5.1 is deduced from Theorem 5.3 once we prove that the process indeed satisfies (**J**), for which only (**J**7) is not elementary. Let us first clarify the meaning of these assumptions.

The main difference with Assumption (**P**) lies in the last three assumptions. Remark that we could not avoid the comparison with the quantity ρ_S in the last two. Note however that, as we can see in Sect. 5.3, Lemma 3.0.2 of [63] provides an efficient way to get upper-bounds of ρ_S. It is hopefully enough to ensure such properties as (**J6**, **7**).

While (**J6**) plays a similar role as (**P1**), (**J5**, **7**) are really specific to the fact that some jump directions are restricted. In this illustrative example, (**J5**) is required to make sure that the different directions can efficiently be explored by the process. Thanks to it, there exists $p_\wedge$ such that the following lower-bound holds uniformly for $x \in \mathbb{R}^d$ and $i \le d$: $\rho_J^i(x) \ge p_\wedge \cdot \rho_J(x)$. (**J7**) is required to prevent some directions in $\mathbb{R}^d$ from being avoided by the process, meaning that the probability to do so for a long time becomes negligible even compared to extinction.

For clarity of the other properties, let us consider a compact K subset of $\mathbb{X}$. (**J1**) is exactly as (**P2**): there exists thus $\rho_J^\vee > 0$ such that the following upper-bound holds uniformly for $x \in K$: $\rho_J(x) \le \rho_J^\vee$. By virtue of (**J2**), there exist $r, g_\wedge > 0$ such that the following upper-bound holds uniformly for $x \in K$, $i \in [\![1, d]\!]$ and $w \in B(0, r)$: $g_i(x, w) \ge g_\wedge$. This is analogous to (**P3**) except that there is no drift to compensate here. (**J3**) is analogous to (**P4**): by its virtue, for any $\epsilon > 0$, there exists $w_\vee$ such that the following upper-bound holds uniformly for $x \in K$, and $i \in [\![1, d]\!]$: $\int_{\mathbb{R}} g_i(x, w)\, \mathbf{1}_{\{\|w\|_\infty \ge w_\vee\}}\, \mathrm{d}w \le \epsilon \cdot \rho_J(x)$. Finally, (**J4**) is similar to (**P5**): there exists $g_\vee$ such that the following upper-bound holds uniformly for $x \in K$, $i \in [\![1, d]\!]$ and $w \in \mathbb{R}$: $g_i(x, w) \le g_\vee \cdot \rho_J(x)$.

5.3. Proof of (J7) for our typical example

In this example, $\rho_J^i \equiv 1$ while the minimal value of ρ_e is simply 0. Thus, we need to prove that provided $\sigma \le 1/8$, $\rho_S < \rho_F = 1$ holds. We rely on the criteria proposed in Lemma 3.0.2 of [63] and aim at finding some set $\mathcal{D}_S$, $L \ge 1$ and $t > 0$ such that:

$$\inf_{\{x \in \mathcal{D}_S\}} \mathbb{P}_x(X_t \in \mathcal{D}_S\,,\ t < \tau_\partial \wedge T_{\mathcal{D}_L}) > \mathrm{e}^{-t}.$$

We justify next our choice of $t := (4/3) \cdot d \cdot \log 4$, $\mathcal{D}_S$ and $\mathcal{D}_L$ being of the form respectively $B(0, a)$ and $\bar{B}(0, 2a)$ for $a := 1/4$. Since the jumps of X occur at a uniform rate $1/d$ along each direction and with a distribution independent of the position, the increase process $(X_t^i - x^i)_{t \ge 0, i \le d}$ on each coordinate can be expressed as a i.i.d. family of processes whose law is given by:

$$Y_t := \sigma \textstyle\sum_{\{j \le N'_t\}} W_j,$$

where $(N'_t)_{t \ge 0}$ a standard Poisson process on $\mathbb{Z}_+$ with intensity $(1/d)$ while for any $j \ge 1$ W_j is an normal random variable. N' and the family $(W_j)_j$ are independent. We remark that Y is a martingale with predictable quadratic variation $\langle Y \rangle_t := \sigma^2 t/d$, with the same law as $-Y$ as a symmetry.

Exploiting also the fact that ρ_e is upper-bounded by $4a^2 = 1/4$ on $\mathcal{D}_L$ and thanks to the symmetries of the process, we deduce:

$$\begin{aligned}
&\inf_{\{x \in \mathcal{D}_S\}} \mathbb{P}_x(X_t \in \mathcal{D}_S\,,\ t < \tau_\partial \wedge T_{\mathcal{D}_L}) \\
&\quad \ge \mathrm{e}^{-t/4} \inf_{\{x \in \mathcal{D}_S\}} \mathbb{P}_x(\sup_{\{s \le t\}} \|X_s - x\|_\infty \le a\,,\ \forall\, i \le d,\ x_i \cdot (X_t^i - x_i) \le 0) \\
&\quad \ge \mathrm{e}^{-t/4} \left[(1/2) \cdot \mathbb{P}(\sup_{\{s \le t\}} |Y_s| \le a)\right]^d.
\end{aligned}$$

Thanks to the Doob's inequality, and recalling our expressions for a and t:

$$\mathbb{P}(\sup_{\{s\leq t\}} |Y_s| \geq a) \leq \frac{\mathbb{E}[\langle Y\rangle_t]}{a^2} = \frac{16\sigma^2 t}{d} = \frac{64\sigma^2 \log 4}{3} < 1/2,$$

provided $\sigma \leq 1/8 < \sqrt{3/[128 \cdot \log(4)]}$. Since the definition of t is made such that $\mathrm{e}^{-t/4}/4^d \geq \mathrm{e}^{-t}$, this concludes the following uniform lower-bound:

$$\inf_{\{x\in\mathcal{D}_S\}} \mathbb{P}_x(X_t \in \mathcal{D}_S\,,\ t < \tau_\partial \wedge T_{\mathcal{D}_L}) > \mathrm{e}^{-t}.$$

Thanks to Lemma 3.0.2 of [63], ρ_S is thus necessarily smaller than 1, which concludes the proof of $(J7)$ for our example. □

Remark 5.4. i) The condition $\sigma \leq 1/8$ comes only from the way we prove $(J7)$ and is likely not to be optimal. For too large values of σ however, singular concentration effects around 0 may play a substantial role, as in [7]. The event consisting of forbidding any jump when starting at a Dirac Mass around 0 might lead to a lower rate of decay in probability than the one consisting of accumulating jumps, because these jumps mostly send the process to deadly regions.

ii) The purpose of Assumption $(\mathbf{J7})$ is to bound the time T_c at which either extinction occurs or all of the coordinates have changed. Assumption $(\mathbf{J7})$ indeed ensures an exponential moment of T_c with parameter $\hat{\rho}_S$ (*cf.* (5.3) below).

5.4. Proof of Theorem 5.3

For this example, we consider the family $(\mathcal{D}_\ell)_{\ell\geq 1}$ as the open balls $\mathcal{D}_\ell := \bar{B}(0,\ell)$, now for the supremum norm $\|.\|_\infty$ for commodity.

Remark 5.5. Because this norm is equivalent to the Euclidian norm, it is not difficult to see that the statements of Assumption $(\mathbf{A_F})$ are actually equivalent for these two choices.

Assumption $(\overline{A0})$ is clearly satisfied. The proof of $(A1)$ as stated in the following proposition is very similar to the one of Proposition 4.3. By these means, we deal with each coordinate one by one so as to get a uniform lower-bound of the density on a subspace of inductively increasing dimension. The reader will be spared further details.

Proposition 5.6. *Assumptions* $(\mathbf{J1}, \mathbf{2}, \mathbf{5})$ *imply Assumption* $(A1)$*, with* ζ *the uniform distribution over* $\mathcal{D}_1$*. More generally, for any* $\ell \geq 1$*, there exist* $L > \ell$ *and* $t, c > 0$ *such that the follwoing inequality holds for any* $x \in \mathcal{D}_\ell$*:*

$$\mathbb{P}_x\left[X_t \in \mathrm{d}y\,;\ t < \tau_\partial \wedge T_{\mathcal{D}_L}\right] \geq c\,\mathbf{1}_{\{y\in\mathcal{D}_\ell\}}\,\mathrm{d}y.$$

Thanks to Proposition 5.6, we know in particular that Assumption $(A1)$ holds true for the uniform distribution over $\mathcal{D}_1$, *i.e.*: $\zeta(\mathrm{d}y) := \mathbf{1}_{\{y\in\mathcal{D}_1\}}/Leb(\mathcal{D}_1)\,\mathrm{d}y$. Assumption $(A2)$ is implied by Assumption $(\mathbf{J6})$ for some $\rho > \rho_S$ and for $E := \mathcal{D}_L$ where $L \geq 1$ is chosen sufficiently large. The proof of Assumption $(A3_F)$ for these choices is a clear consequence of the next proposition, whose proof is given in the next subsection:

Proposition 5.7. *Assumption* $(\mathbf{J})$ *implies that for any* $E \in \mathbf{D}$ *and* $\rho > 0$*, Assumption* $(A3_F)$ *holds.*

With this result, we can conclude that Assumption $(\mathbf{A_F})$ holds true. Thanks to Theorem 2.8, Corollary 2.12 and Theorem 2.16 it directly implies Theorem 5.3.

5.4.1. Proof of Proposition 5.7

We consider here three types of "failed attempts". Either the process has not done all of its required jumps despite a very long time of observation, or there are too many of these jumps, or at least one of these jumps is too large.

Definition of the stopping times and time of observation

For $k \leq d$, let T_k^J be the first time at which (at least) k jumps have occurred in different coordinates. On the event $\left\{T_k^J < \tau_\partial\right\}$ (for $0 \leq k \leq d-1$), and conditionally on $\mathcal{F}_{T_k^J}$, we know thanks to Assumption (**J7**) that $(T_{k+1}^J \wedge \tau_\partial) - T_k^J$ is upper-bounded by an exponential variable with rate parameter $\rho_F > \hat{\rho}_S$. Thus, with $\hat{\rho}'_S := (\rho_F + \hat{\rho}_S)/2$, we may define the finite quantity $e_f := [2\,\rho_F/(\rho_F - \hat{\rho}_S)]^d$ and deduce the following upper-bound of the exponential moment uniformly in $x \in \mathbb{R}^d$: $\mathbb{E}_x \exp[\hat{\rho}'_S \cdot (T_d^J \wedge \tau_\partial)] \leq e_f$. Thanks to the Markov's inequality, this implies the following upper-bound on the probability that $T_d^J \wedge \tau_\partial$ takes large values:

$$\mathbb{P}_x[T_d^J \wedge \tau_\partial > t] \cdot \exp[\hat{\rho}_S \cdot t] \leq e_f \, \exp[-(\rho_F - \hat{\rho}_S) \cdot t/2], \tag{5.3}$$

which tends to 0 as t tends to infinity.

Let $\epsilon > 0$. Thanks to Inequality (5.3), we can choose $t_F > 0$ such that:

$$\exp[\hat{\rho}_S \, t_F]\, \mathbb{P}_x(t_F < T_d^J \wedge \tau_\partial) \leq \epsilon/3. \tag{5.4}$$

On the event $\left\{t_F < T_d^J\right\}$, we set $U_H := \infty$. This clearly implies that $T_d^J \leq t_F$ holds a.s. on the event $\{U_H < \tau_\partial\}$.

Upper-bound on the number of jumps

Thanks to Assumption (**J5**), at each new jump, conditionally on the past until the previous jump, there is a lower-bounded probability that a new coordinate gets altered. The number N_J of jumps before T_d^J (on the event $\{T_d^J < \tau_\partial\}$) is thus upper-bounded by a sum of d mutually independent geometric random variables. Therefore, we can define $n_J^\vee \geq 1$ such that the following upper-bound holds uniformly in $x \in \mathbb{R}^d$:

$$\mathbb{P}_x(N_J \geq n_J^\vee\,,\; T_d^J < \tau_\partial) \leq \epsilon \, \exp[-\hat{\rho}_S \, t_F]/3. \tag{5.5}$$

We thus declare a failure if the $n_J^\vee$-th jump occurs while T_d^J still is not reached.

Upper-bound on the size of the jumps

The crucial argument on the jump size is given by the following lemma.

Lemma 5.8. *Suppose that Assumption* (**J3**) *holds. Consider any $L > 0$, any $N \geq 1$ and any $\epsilon > 0$. Let $(W_i, i \geq 0)$ denote the time-ordered sequence of jump effects. Then, there exists $w_\vee > 0$ such that the following lower-bound holds uniformly in the initial conditions $x \in \bar{B}(0, L)$:*

$$\mathbb{P}_x(\sup\nolimits_{i \leq N} \|W_i\|_\infty \leq w_\vee) \geq 1 - \epsilon.$$

Proof of Lemma 5.8: The property is proved by induction over N, where one needs to adjust at each step both ϵ and $w_\vee$. The initialization is directly implied by Assumption (**J3**). For some N and $w_\vee^N > 0$, consider the event $\mathcal{W}_N(w_\vee^N)$ according to which the N first jumps have a size that is upper-bounded by $w_\vee^N > 0$, that is $\mathcal{W}_N(w_\vee^N) := \{\sup_{i \leq N} \|W_i\|_\infty \leq w_\vee^N\}$

Assume by the induction hypothesis that $w_\vee^N$ is chosen such that the following lower-bound is ensured uniformly for $x \in \bar{B}(0, L)$: $\mathbb{P}_x(\mathcal{W}_N(w_\vee^N)) \geq 1 - \epsilon/2$. On the event $\mathcal{W}_N(w_\vee^N)$, with $\|x\|_\infty \leq L$, we deduce that $\|X(U_N^J)\|_\infty \leq L + N \cdot w_\vee^N$, where U_N^J denotes the N-th time of jump of X. Recall that $\mathcal{F}_{U_N^J}$ describe the information of the process up to its N-th jump time. Thanks to Assumption (**J3**), there exists $w_\vee^{N+1} \geq w_\vee^N$

such that the event $\{\|W_{N+1}\|_\infty \le w_\vee^{N+1}\}$ occurs with probability greater than $1-\epsilon/2$ conditionally on $\mathcal{F}_{U_N^J}$ restricted to the event $\mathcal{W}_N(w_\vee^N)$ and uniformly on $x \in \bar{B}(0,L)$. Note also the following inclusion:

$$\mathcal{W}_N(w_\vee^N) \cap \{\|W_{N+1}\|_\infty \le w_\vee^{N+1}\} \subset \mathcal{W}_{N+1}(w_\vee^{N+1}).$$

Thanks to the Markov property, the following upper-bound is then derived for any $x \in \bar{B}(0,L)$:

$$\begin{aligned}\mathbb{P}_x(\mathcal{W}_{N+1}(w_\vee^{N+1})) &\ge \mathbb{E}_x\Big[\mathbb{P}_x(\|W_{N+1}\|_\infty \le w_\vee^{N+1} \,|\, \mathcal{F}_{U_N^J})\,;\; \mathcal{W}_N(w_\vee^N)\Big]\\ &\ge [1-\epsilon/2]^2 \ge 1-\epsilon.\end{aligned}$$

The induction over N then concludes the proof of Lemma 5.8. □

Thanks to Lemma 5.8, we can choose a value $w_\vee > 0$ such that $\epsilon \cdot \exp[-\hat{\rho}_S\, t_F]/3$ is for any $x \in E$ an upper-bound of the probability for the process starting from x that there is a jump before the $n_J^\vee$-th jump and $T_d^J \wedge \tau_\partial$ with size larger than $w_\vee$. We thus declare a failure if a jump larger than $w_\vee$ occurs. From this we deduce that, on the event $\{U_H < \tau_\partial\}$, the process has stayed in $\mathcal{D}_L$ for $L := \ell_E + n_J^\vee \cdot w_\vee$.

We set $U_H(\omega) = T_d^J(\omega)$ for any ω such that the three following conditions hold:

$$(i)\, T_d^J(\omega) < \tau_\partial(\omega),\; T_d^J(\omega) \le t_F,\; (ii) N_J(\omega) \le n_J^\vee, \;\text{ and } (iii) \text{ for any } i \le N_J(\omega), W_i(\omega) \le w_\vee.$$

Otherwise $U_H(\omega) := \infty$.

Given our construction (see (5.4), (5.5) and the above definition of $w_\vee$), it is clear that:

$$\{\tau_\partial \wedge t_F < U_H\} = \{U_H = \infty\} \quad \text{and} \quad \mathbb{P}_x(U_H = \infty,\, t_F < \tau_\partial) \le \epsilon\, \exp(-\rho\, t_F).$$

The proof of Proposition 5.7 is then completed thanks to Proposition 5.6 together with the next lemma, whose proof constitutes the last step:

Lemma 5.9. *Suppose that Assumptions* (**J3** − **7**) *hold, with the preceding notations. Then, there exists* $c > 0$ *such that:*

$$\mathbb{P}_x(X(U_H) \in \mathrm{d}x'\,;\; U_H < \tau_\partial) \le c\,\mathbf{1}_{\{x' \in \mathcal{D}_L\}} \mathrm{d}x'.$$

□

5.4.2. Proof of Lemma 5.9

The proof is based on an induction on the coordinates affected by jumps in the time-interval $[0, t_F]$. We recall that, thanks to our criterion of exceptionality, we can restrict ourselves to trajectories where any coordinate is affected by at least one jump in the time-interval $[0, t_F]$, while at most $n_J^\vee$ jumps have occurred in this time-interval. We consider the sequence of directions that the process follows at each successive jumps. There is clearly a finite number of such sequences. In order to deduce the upper-bound on the density of $X(U_H)$ presented in Lemma 5.9, we merely need to prove the restricted versions for any such possible sequence of directions.

Let $(i(k))$ for $k \le n_J \le n_J^\vee$ be a given sequence of directions in $[\![1, d]\!]$ such that, at $k = n_J$, all the d directions have been listed. Let also $I(k) \in [\![1, d]\!]$ for $k \le n_J^\vee$, be the sequence of random directions followed by the $n_J^\vee$ first successive jumps of X. We recall the notation U_N^J for the time of the k-th jump of X

Since in our model, all directions are defined in a similar way, we can simplify a bit our notations without loss of generality by relabeling some of the directions. Since we will go backwards to progressively forget about the conditioning, we order the coordinates by the time they appear for the last time in the sequence $(i(k))_{k\leq n_J}$.

It means that $i(n_J) = d$ and that, up to the relabeling, we exploit the unique non-decreasing function $j : [\![1, n_J]\!] \to [\![1, d]\!]$ such that for any $K \in [\![1, n_J]\!]$, $\{i(k)\,;\ K \leq k \leq n_J\} = [\![j(K), d]\!]$. Let then $K[j]$ be the largest integer $k \leq n_J$ such that $j(k) \leq j$. With this definition, it holds for any $j \in [\![1, d]\!]$ that $i(K[j]) = j$ and that for any $k \in [\![K[j] + 1, n_J]\!]$, $i(k) \in [\![j + 1, d]\!]$.

Remark 5.10. In our case, n_J is naturally chosen as the first integer for which all the directions have been listed. Yet, our induction argument is more clearly stated if we do not assume this condition on n_J.

We define the sequence $(\mathcal{A}(k))_{k\leq n_J}$ of events that encode the fact that U^J_k has not reached $\tau_\partial \wedge t_F$ and that, up to the k-th jump, the random sequence of directions coincide with the sequence i and the size of the jumps remain uniformly bounded by $w_\vee$. Namely, for $K \in [\![1, n_J]\!]$:

$$\mathcal{A}(K) := \left\{U^J_K < \tau_\partial \wedge t_F\right\} \cap \left\{\forall\, k \leq K - 1,\ I(k) = i(k), \|\Delta X(U^J_k)\|_\infty \leq w_\vee\right\}.$$

Then, we look for a lower-bound that is uniform in $x \in E$ on the following expectation that involves any given non-negative and measurable functions $(f_j)_{j\leq d}$:

$$E^d := \mathbb{E}_x\Big[\prod\nolimits_{j\leq d} f_j[X^j(U^J_{n_J})]\,;\ \mathcal{A}(n_J)\Big].$$

Define the information up to time $U^J_{n_J}$ deprived from the last jump size as follows:

$$\mathcal{F}^*_{U^J_{n_J}} := \sigma\big(\mathcal{F}_{U^J_{n_J-1}}, \{I(n_J) = d\} \cap \{U^J_{n_J} < \tau_\partial \wedge t_F\}\big).$$

To compute E^d, we then need to compute the expectation of the following quantity:

$$\prod_{j\leq d-1} f_j[X^j(U^J_{n_J-1})] \cdot \mathbb{E}_x\Big[f_d[X^d(U^J_{n_J})]\,;\ |\Delta X^d(U^J_{n_J})| \leq w_\vee \,|\, \mathcal{F}^*_{U^J_{n_J}}\Big], \tag{5.6}$$

restricted to the following event:

$$\mathcal{A}(n_J - 1) \cap \left\{U^J_{n_J} < \tau_\partial \wedge t_F\right\} \cap \{I(n_J) = d\}.$$

Note that $X(U^J_{n_J}-) = X(U^J_{n_J-1})$ is $\mathcal{F}_{U^J_{n_J-1}}$-measurable, since we consider a pure jump process. Thanks to the Markov property, the law of the next jump only depends on $x' := X(U^J_{n_J-1})$ through the functions $(w \mapsto g_j(x', w))_{j\leq d}$. With the σ-algebra $\mathcal{F}^*_{U^J_{n_J}}$, we include the knowledge of the direction of the jump at time $U^J_{n_J}$, so that only the size of this jump (possibly negative) remains random. With $L := \ell_E + n^\vee_J \cdot w_\vee$, which is clearly independent of n_J and of the particular choice of the sequence $(i(k))$, we note the following containment property on the event $\mathcal{A}(n_J)$:

$$\|X^d(U^J_{n_J}-)\|_\infty \vee \|X^d(U^J_{n_J})\|_\infty \leq L.$$

This implies thanks to Assumption (**J4**) that the following inequality holds a.s. on the event $\mathcal{A}(n_J - 1) \cap \left\{U^J_{n_J} < \tau_\partial \wedge t_F\right\} \cap \{I(n_J) = d\}$:

$$\mathbb{E}_x\Big[f_d[X^d(U^J_{n_J})]\,;\ |\Delta X^d(U^J_{n_J})| \leq w_\vee \,|\, \mathcal{F}^*_{U^J_{n_J}}\Big] \leq g_\vee \int_{[-L,L]} f_d(x^d)\, \mathrm{d}x^d. \tag{5.7}$$

In what follows, the probability of the event $\{U^J_{n_J} < \tau_\partial \wedge t_F\} \cap \{I(n_J) = d\}$ is simply upper-bounded by 1. Combining Inequalities (5.6), (5.7), and our ordering with the definition of $K[j]$, we deduce:

$$E^d \leq g_\vee \int_{[-L,L]} f_d(x^d)\, \mathrm{d}x^d \cdot \mathbb{E}_x\left[\prod_{j\leq d-1} f_j[X^j(U^J_{n_J-1})]\,;\ \mathcal{A}(n_J-1)\right]$$
$$\leq g_\vee \int_{[-L,L]} f_d(x^d)\, \mathrm{d}x^d \cdot \mathbb{E}_x\left[\prod_{j\leq d-1} f_j[X^j(U^J_{K[d-1]})]\,;\ \mathcal{A}(K[d-1])\right].$$

Recall in particular that $i(K[d-1]) = d-1$ and that $K[d-1] \leq n^\vee_J$. The procedure can be iterated as follows:

$$E^{(d-1)} = \mathbb{E}_x\left[\prod_{j\leq d-1} f_j[X^j(U^J_{K[d-1]})]\,;\ \mathcal{A}(K[d-1])\right]$$
$$\leq g_\vee \int_{[-L,L]} f_{d-1}(x^{d-1})\, \mathrm{d}x^{d-1} \cdot \mathbb{E}_x\left[\prod_{j\leq d-2} f_j[X^j(U^J_{K[d-2]})]\,;\ \mathcal{A}(K[d-2])\right],$$

and so on until finally:

$$E^d \leq (g_\vee)^d \cdot \prod_{i\leq d}\left(\int_{[-L,L]} f_i(x)\, \mathrm{d}x\right).$$

We then sum over all sequences $(i(k))$ possibly observed up to time T^J_d. With the definition of the range of sequence i up to step $n \geq 1$ as $\mathcal{R}^i_n := \{i(k)\,;\ k \leq n\}$, the set of these sequences can be rigorously defined as follows:

$$\left\{(i(k))_{\{k\leq n_J\}} \in [\![1,d]\!]^{n_J}\,;\ n_J \leq n^\vee_J, \quad n_J = \min\{n\geq 1; \mathcal{R}^i_n = [\![1,d]\!]\}\right\}.$$

There are clearly less than $d^{n^\vee_J}$ possibilities (there is a surjection from the set of all sequences of length $n^\vee_J$). Since for any positive and measurable functions $(f_j)_{j\leq d}$, the following upper-bound is deduced uniformly for any $x \in E$:

$$\mathbb{E}_x\left[\prod_{j=1}^d f_j[X(U_H)]\,;\ U_H < \tau_\partial\right] \leq d^{n^\vee_J} \cdot (g_\vee)^d \int_{\bar{B}(0,L)} \prod_{j=1}^d f_j(x_j) \mathrm{d}x_1 ... \mathrm{d}x_d.$$

It is classical that it implies the following lower-bound on the marginal density:

$$\forall\, x \in E, \quad \mathbb{P}_x\left[X(U_H) \in \mathrm{d}x'\,;\ U_H < \tau_\partial\right] \leq d^{n^\vee_J} \cdot (g_\vee)^d\, \mathbf{1}_{\{x' \in \bar{B}(0,L)\}}\, \mathrm{d}x'.$$

It concludes the proof of Lemma 5.9. □

Recall with the statement just before Lemma 5.9 that the proof of Proposition 5.7 is now completed. Note also with the statement just after Proposition 5.7 that the proof of Theorem 5.3 is then completed. With the statement just after Theorem 5.3, this also concludes the proof of Theorem 5.1.

6. Discussion

In this final section, we first discuss the choices and parameters of our newly proposed condition, in Sect. 6.1, then the range of applications for which we expect it to be mostly helpful, see Sect. 6.2. We conclude in Sect. 6.3 with the practical implications of our results for the considered eco-evolutionary applications.

6.1. Assumption ($A3_F$) of "Almost Perfect Harvest"

How the different parameters have to be adjusted? In fact, we will exploit this assumption only for a given single value of $\epsilon > 0$, which is explicitly related to the other parameters (*cf.* Sect. 3.1). But in generic proofs, this explicit value is not expected to be really tractable.

The random variable U_H and V are thus expected to depend both on $x \in E$ and on ϵ, and to be related to t_F and c, while these two constants must be uniform in x.

Is it really important to consider failures? The purpose of introducing failure is to handle singularities, *i.e.* events which are rare in probability but for which comparison estimates are poor or simply impossible.

Notably in pure jump models, waiting for a jump is *a priori* needed to loosen the dependency on the initial condition (especially when the latter is a Dirac mass). Yet, this implies that the event of a very late jump (being one condition of failure in the harvest) has to be considered carefully, to prove that its probability is negligible compared to the whole survival probability. In multidimensional model, we may also need to wait for a jump on a specific coordinate to happen, while there is often a positive probability for very singular behavior to happen meanwhile on the other coordinates. It is generally needed to adjust the singularity level (and implicitly the efficiency of the harvesting step), for the associated probability to be sufficiently small and for such events to be treated as a failure in the harvest.

If this issue is made easily manageable in the applications of Section 4 and 5, this is mainly because we allow for both random stopping times and failure events.

Considering failures can also be of interest in order to exploit a Girsanov transform to simplify the dynamics of the process. As can be seen in [65] and [49], this transform is very efficient to relate the original dynamics to one that is more easily described, notably by decoupling different components of the dynamics. Namely, the original and simplified models are related through a change of the densities by a multiplicative factor that is upper-bounded except for rare events in probability. The statement of our property ($A3_F$) is very adapted to deal with such imprecision: exceptional behavior is treated as a failure, so that a uniform bounds on the multiplicative factor can be ensured. From these bounds, we can then deduce the appropriate constant c in (2.7).

Is it crucial to assume that the state space Ω is a path space? This assumption is not essential yet very helpful. The assumption is only exploited in order to have a proper and direct expression of the final harvesting time U_H^∞ for any possible choice of U_H as a stopping time. Without this assumption on Ω, it should usually not be a big issue to construct such a time U_H^∞ ("by hand", depending on how U_H and X are defined). Yet one would have to deal with more technical details in checking that it indeed is a stopping time and in adapting how point (*ii*) in Proposition 2.2 is exploited for the proof of Lemma 3.3.

Could we improve the assumption with less restrictions on the parameters? Would it be worth it? The first condition in (2.4) means indeed that U_H is required to be less than t_F for a first success in the harvest to be achieved. This implies the following equalities: $\{U_H < \tau_\partial\} = \{U_H \leq t_F\} = \{U_H = \infty\}^c$. The requirement that U_H must be less than t_F is not as stringent as it might seem and it makes the statement of ($A3_F$) much more tractable.

We believe it is generally compatible with the upper-bound on the failure event that one restricts any candidate $\widetilde{U}_H$ to be less than t_F, provided t_F is large enough, possibly by reducing the considered value of ρ towards a value closer to ρ_S and by considering the event $\{\widetilde{U}_H \geq t_F\}$ as an additional criterion of failure.

Nonetheless, a refinement of the assumption with a looser upper-bound on U_H may still provide a better estimate of the constants involved. Note simply that it requires to specify the times at which failures are stated, since there is no more reason for each step to end before time t_F. Since the statement would be much more technical, it is not included in the current article. Still, one may find a version of the proofs adapted for this context in the second ArXiv version [64] of the current article.

6.2. Brief overview of the intended applications

Although greatly simplified, the two applications of the current article relate to eco-evolutionary models. The growth rate or the persistence of a population is related to the individual characteristics of its members, in other words their "features" or "traits". This effect shall be represented by the state-dependent extinction rate. The dynamics of these traits may depend on mutations, a changing environment or the ageing of the individuals, for which our applications provide archetypal models. These effects are expected to be represented in a "discontinuous" fashion, *i.e.* with brutal transitions, for which our approach is adapted.

Eco-evolutionary models thus form a large class of applications. Our assumption of a constant drift term in our first application is merely taken for simplicity given our multidimensional state space. Our approach could simply be adjusted for a drift term depending on the position as in [32] or [25], as long as it brings the process to infinity. The proof is then much more specific to the biological motivation.

Such a drift term could as well be interpreted as the ageing of individuals in age-structured population models (as in [61]), or as the growth rate of the units in growth-fragmentation models (as in [15, 33, 52]). Our hope is to see our technique fruitful for these applications extended to a multidimensional setting (where age or size is coupled to other individual features). The applications are not restricted to ecology, and may for instance come from chemistry (notably for polymer growth as in [43]), neuroscience (see the elapsed-time models *e.g.* in [61]) or epidemiology (notably when the infection rate depends on the elapsed time after the infection as presented in [14], Sect. 1.1.2).

More detailed ecological models have also been studied thanks to the theorems of the current paper. Notably in [65], we couple a diffusive process specifying the population size to a piecewise deterministic process specifying the adaptation of the population. The proof is more involved than in the current paper, notably because we use the Girsanov transform to decouple the diffusive and the piecewise deterministic components of the system. In [49], we study another related application, in which accumulation of deleterious mutations is slowed down by natural selection: the conditions of the present article are exploited to obtain the convergence to a unique QSD of a diffusion in an infinite dimensional state space.

More generally, our techniques provide conditions ensuring the existence and uniqueness of the positive eigenvector of general linear non-local reaction diffusion equations (see notably Sect. 6.3.2 for some partial results and Sect. 4.1 for the related conditions). The long-time behavior of structured branching processes can typically be captured by such results, thanks to the many-to-one formula (see [42]).

6.3. Practical implications of these results

6.3.1. Biological motivations

The processes presented in our applications can be seen as models for the adaptation of a population to its environment, known as replicator-mutator processes in evolutionary biology (see *e.g.* [2]). The density of the quasi-stationary distribution is then usually known as the ground state (of the population). In our first application, forcing by a regularly changing environment is considered (in the spirit of [9]), whereas in the second application, dependent but distinct subpopulations contribute to global adaptation to an otherwise fixed environment.

Since the pioneering work of Kimura [46], non-local kernels are considered for the generation of mutation effects. ρ_e is known as the confining potential, as it prevents in unbounded domain the population features of adaptation from spreading to any extent. The work of Bürger [3] has brought to light the interest, at least in the absence of drift and of diffusion term, to derive conditions that prevent concentration effects: concentration towards singular measures like Dirac Mass are indeed observed for confining potentials that are too concentrated with respect to the non-local kernel. Such conditions can be found notably in [3, 4, 28, 48].

The environmental change in the first application is represented by a translation of the fitness landscape at constant speed v. We can consider X as a summary of the individual characters of the population. Then, the jumps come from the fixation of new mutations in the population, whose rate depends on the adaptation of the mutant subpopulation (with trait $X_{t-} + w$) as compared to the resident individuals (with trait X_{t-}). A much more detailed description is proposed in [65]. There, we extend the proof to a coupled process involving additionally continuous fluctuations of the population size.

Considering distinct directions of jumps in the second application is motivated by the interpretation that each of these directions corresponds to the variation of a single species, where the various d species contribute to the survival of the community. Many communities are then subjects to death and reproduction events and we can describe the state of the meta-community in this formalism as in [66].

6.3.2. Connection with reaction-diffusion equations

The quasi-stationary regime of the process generally prescribed in (4.2) is expected be related to the behavior of the solution $(u(t,x))_{t\geq 0, x\in\mathbb{R}^d}$ at low densities ($u \approx 0$) to reaction-diffusion evolution equations of the form:

$$\partial_t u(t,x) := v\,\partial_{x_1} u(t,x) + \int_{\mathbb{R}^d} g(y, x-y)\, u(t,y)\,\mathrm{d}y - \left(\int_{\mathbb{R}^d} g(x,w)\,\mathrm{d}w\right) u(t,x) + r(x, u(t,x))\, u(t,x). \quad (6.1)$$

In this setting, $u(t,x)$ is usually meant to represent the relative abundance of individuals with trait x at time t in a population of very large size. At low density, the approximation of the growth rate $r(x,u)$ by $r_0(x) := r(x,0)$ is usually valid, so as to linearize (6.1). Looking at the linear problem provides a criterion for the possibility and rate of invasion, cf for instance [10] and [30].

Also, if we consider $r_0(x)$ as an upper-bound of $r(x,u)$ for any density u, the solution $\bar{u}$ to the linear problem with r_0 shall provide an upper-bound of u by maximum principle approaches. If the eigenvalue $\lambda^\star$ of the linear problem is negative, the solution u is expected to asymptotically decline at least quicker than at rate $-\lambda^\star$. Thus, results such as ours have implications as a criteria for non-persistence, as in [8–12, 16].

In view of these interpretations, several authors are looking at characterizing such eigenvalue problems when there is possibly no regular eigenvector (see *e.g.* [28, 32, 41, 44, 48]). For now, we simply conjecture that, provided Theorem 4.2 applies (with a translation of the growth rate by a constant to deduce an extinction rate), they all coincide to the value prescribed with λ.

6.3.3. Ecological relevance of the results

In practice, the dynamics for which results of quasi-stationarity can be derived usually come as an approximation. The relevance of the approximation is then of course at stake, yet the quasi-stationary regime may provide insight on the conditions of relevance.

Considering for instance our first application in Section 4.2, the population is certainly doomed to extinction for too strong environmental drift. When population size strongly declines, the estimation of individual features through the marginal of X is not relevant. On the other hand, our result of convergence for the process X is not directly affected: it holds for any value of v. As v tends to ∞, we shall simply have the asymptotic extinction rate λ going to $-\infty$. Considering the asymptotic extinction rate, notably in comparison to the convergence rate, can nonetheless inform about the validity of the marginal of X to capture the indivual features.

Appendix A.

A.1 Elementary facts in the absorbed setting

Proof of Fact 2.7: Let us demonstrate that the property given in Definition 2.6 implies that α is a QSD with extinction rate λ. For $u > 0$, let us define $\mu_u(\mathrm{d}s) := \mathrm{e}^{\lambda u}\alpha P_u(\mathrm{d}s) - \alpha(\mathrm{d}s)$. Then, for any $t > 0$:

$$\begin{aligned}\left\|\mathrm{e}^{\lambda t}\alpha P_t - \alpha\right\|_{TV} &\leq \left\|\mathrm{e}^{\lambda t}\alpha P_t - \mathrm{e}^{\lambda(t+u)}\alpha P_{t+u}\right\|_{TV} + \left\|\mathrm{e}^{\lambda(t+u)}\alpha P_{t+u} - \alpha\right\|_{TV}\\ &= \left\|\mathrm{e}^{\lambda t}\mu_u.P_t\right\|_{TV} + \left\|\mu_{u+t}\right\|_{TV}\\ &\leq C\,\mathrm{e}^{-\gamma u}(\mathrm{e}^{\lambda t}\left\|\!\left| P_t \right|\!\right\| + \mathrm{e}^{-\gamma t}).\end{aligned}$$

Letting u tends to ∞ concludes the equality $\alpha P_t(\mathrm{d}s) = \mathrm{e}^{-\lambda t}\alpha(\mathrm{d}s)$. □

Proof of Fact 2.13 For any $u \geq 0$, $\|\mu - \alpha\|_{TV} \leq \|\mu - u\,\alpha\|_{TV} + |1-u|$ because $\mu - u\,\alpha = \mu - u\,\alpha + (1-u)\,\alpha$. On the other hand $\|\mu - u\,\alpha\|_{TV} \geq |\mu(\mathbb{X}) - u\,\alpha(\mathbb{X})| = |1-u|$. By combining these two estimates, we conclude that $\|\mu - u\,\alpha\|_{TV} \geq \|\mu - \alpha\|_{TV}/2$.

Let μ be such that $\langle \mu \,|\, 1/h\rangle < \infty$ and define the biased probability distribution:

$$\nu(\mathrm{d}x) := \frac{1/h(x)}{\langle \mu \,|\, 1/h\rangle}\mu(\mathrm{d}x).$$

Exploiting the previous inequality, we deduce that for any $u > 0$:

$$\begin{aligned}\|\mu - u\,\beta\|_{1/h} &= \langle \mu \,|\, 1/h\rangle \cdot \|\nu - (u/\langle \mu \,|\, 1/h\rangle)\,\alpha\|_{TV}\\ &\geq \frac{\langle \mu \,|\, 1/h\rangle}{2} \cdot \|\nu - \alpha\|_{TV} = \frac{\|\mu - \langle \mu \,|\, 1/h\rangle\,\beta\|_{1/h}}{2}.\end{aligned}$$ □

A.2 Definition and properties of the final harvest time U_H^∞

This subsection is dedicated to the proof of Proposition 2.2. For this proof, it is convenient to assume that the process X gets absorbed at ∂ after the extinction time (which may lead to consider the process $\hat{X}$ stopped at the isolated element ∂ as introduced in Sect. 2.1.2, cf also [56], Sect. 11).

Recall that we assume the stopping time U_H to satisfy $\{t_F \wedge \tau_\partial < U_H\} = \{U_H = \infty\}$ for some $t_F > 0$. We defer to Appendix A.3 the discussion on the assumption of Ω to be of "path type", for which the most general definition is given in [56], Definition (23.10). For our purpose, we essentially require the following property (taken *e.g.* from [45], Prop. 8.8 for a canonical space or from [56], Prop. (23.16-ix)), which exploits the notation of $\theta[t]$ for the shift operator with a time-shift of size $t > 0$.

Lemma A.1. *Let Ω be of path type, $(\mathcal{F}_t)$ to be right-continuous and complete, then for any two stopping times S and T, the formula $S + R \circ \theta[S]$ defines a stopping time.*

Recall also that τ_E denotes the hitting time of E, which is closed, that is a stopping time thanks *e.g.* to [56], Section 10. That $\tau_E^F = t_F + \tau_E \circ \theta[t_F]$ is a stopping time can seen as a direct application of Lemma A.1 (or be easily checked directly as in [56], Sect. 10).

We introduce the same sequence $(\tau_E^k)_{k\geq 0}$ as in (3.3) for the proof of Theorem 2.3, where the inductive definition, for $k \geq 1$ can be restated as follows in terms of the shift operator:

$$\tau_E^k = \left(\tau_\partial \wedge \left[\tau_E^{k-1} + \tau_E^F \circ \theta[\tau_E^{k-1}]\right]\right), \qquad \tau_E^0 = 0. \tag{A.1}$$

Since we assumed in this subsection that X is constant to ∂ in the time interval $[\tau_\partial, \infty[$, it holds for any ω that $\tau_E^F \circ \theta[\tau_\partial](\omega) = \infty$. Thanks to Lemma A.1 and by an inductive argument, any element of this sequence clearly defines a stopping time.

Note that the sequence progresses with the second term until it reaches the extinction time and simply stops there. Since $\tau_E^F \geq t_F$, the construction is clearly such that for any $k \geq 1$ on the event $\{\tau_E^k < \tau_\partial\}$, the inequality $\tau_E^k \geq k \cdot t_F$ is satisfied.

We then define U_H^∞ as follows:

$$U_H^\infty = \min\{\tau_E^k + U_H \circ \theta[\tau_E^k];\ k \geq 0\}, \tag{A.2}$$

where we extend the definition of U_H by $U_H = \infty$ for any initial condition not included in E, with notably $U_H \circ \theta[\tau_\partial] = \infty$. These adjustments clearly keep U_H as a stopping time, for any initial condition. Note that it can happen that no harvest occurs before the extinction, in which case $U_H^\infty = \infty$. Also, U_H^∞ is de facto a minimum due to the above inequality on τ_E^k.

Thanks again to Lemma A.1, U_H^∞ is clearly a stopping time as an infimum of a denumerable number of stopping times. For (i), on the event $\{U_H \leq \tau_E^F \wedge \tau_\partial\}$, we have $U_H^\infty \leq U_H = \tau_E^0 + U_H \circ \theta[\tau_E^0]$. Note that $\{t_F \wedge \tau_\partial < U_H\} = \{U_H = \infty\}$ implies on this event that $U_H \leq t_F$. So the reverse inequality comes from the fact that for any $k \geq 1$, either $\tau_E^k \geq t_F \geq U_H$ or $\tau_E^k = \tau_\partial$ (recalling that $U_H \circ \theta[\tau_\partial] = \infty$). On the event $\{\tau_\partial \leq \tau_E^F\} \cap \{\tau_\partial < U_H\}$ on the other hand, we have both $U_H = \infty$ and that for any $k \geq 1$, $\tau_E^k = \tau_\partial$. This concludes that $U_H^\infty = \infty = U_H$ on this event. Because of the decomposition

$$\left\{\tau_\partial \wedge U_H \leq \tau_E^F\right\} = \{U_H \leq \tau_E^F \wedge \tau_\partial\} \cup (\{\tau_\partial \leq \tau_E^F\} \cap \{\tau_\partial < U_H\}),$$

the proof of (i) is completed.

To prove (ii), let ω come from the event $\left\{\tau_E^F < \tau_\partial \wedge U_H\right\}$. Let us prove by induction on $k \geq 1$ that

$$\tau_E^k(\omega) = \tau_E^F(\omega) + \tau_E^{k-1} \circ \theta[\tau_E^F](\omega). \tag{A.3}$$

Since $\tau_E^0 = 0$ and $\tau_E^1(\omega) = \tau_E^F(\omega)$, it clearly holds true for $k = 1$. Note the following elementary property on the shift operator that holds for any stopping times R and S, namely for any $\omega \in \Omega$:

$$\theta[R] \circ \theta[S](\omega) = \theta[S + R \circ \theta[S]](\omega).$$

It directly implies that $\theta[R_1] \circ \theta[S](\omega) + \theta[R_2] \circ \theta[S](\omega) = \theta[R_1 + R_2] \circ \theta[S](\omega)$ holds for any stopping times R_1, R_2, S, and similarly when we replace the sum by the minimum operator. Thanks to the strong Markov property, we have $\tau_\partial(\omega) = \tau_E^F(\omega) + \tau_\partial \circ \theta[\tau_E^F](\omega)$.

Let us assume that (A.3) holds for a given $k \geq 1$. Then, we inject it into (A.1) to deduce the following expression for τ_E^{k+1}, which we simplify thanks to the above properties of the shift operator:

$$\begin{aligned}
\tau_E^{k+1}(\omega) &= \tau_\partial(\omega) \wedge \left[\tau_E^F(\omega) + \tau_E^{k-1} \circ \theta[\tau_E^F](\omega) + \tau_E^F \circ \theta\left[\tau_E^F + \tau_E^{k-1} \circ \theta[\tau_E^F]\right](\omega)\right] \\
&= \tau_E^F(\omega) + \tau_\partial \circ \theta[\tau_E^F](\omega) \wedge \left[\tau_E^{k-1} \circ \theta[\tau_E^F](\omega) + \tau_E^F \circ \theta[\tau_E^{k-1}] \circ \theta[\tau_E^F](\omega)\right] \\
&= \tau_E^F(\omega) + \left(\tau_\partial \wedge [\tau_E^{k-1} + \tau_E^F \circ \theta[\tau_E^{k-1}]]\right) \circ \theta[\tau_E^F](\omega) \\
&= \tau_E^F(\omega) + \tau_E^k \circ \theta[\tau_E^F](\omega).
\end{aligned}$$

By induction, we thus see that (A.3) is satisfied for any $k \geq 1$. Since $U_H(\omega) > \tau_E^F(\omega) \geq t_F$, we deduce $U_H = \infty$ which implies these new expressions for $U_H^\infty(\omega)$:

$$\begin{aligned} U_H^\infty(\omega) &= \min\{\tau_E^k(\omega) + U_H \circ \theta[\tau_E^k](\omega);\ k \geq 1\} \\ &= \min\{\tau_E^F(\omega) + \tau_E^{k-1} \circ \theta[\tau_E^F](\omega) + U_H \circ \theta[\tau_E^{k-1}] \circ \theta[\tau_E^F](\omega);\ k \geq 1\} \\ &= \tau_E^F(\omega) + \min\{\tau_E^\ell + U_H \circ \theta[\tau_E^\ell]; \ell \geq 0\} \circ \theta[\tau_E^F](\omega) \\ &= \tau_E^F(\omega) + U_H^\infty \circ \theta[\tau_E^F](\omega). \end{aligned}$$

This concludes the proof of (ii), and by extension the one of Proposition 2.2. □

A.3 A state space Ω of path type

All the processes involved in our applications can be stated in terms of a path space Ω because all the involved random variables and processes require measurability properties relative to the filtration generated by (a countable family of) right-continuous and left-limited processes. Ω can thus be derived as a canonical space for an extended right-continuous and left-limited process $(\hat{X}_t)$, so that X_t is simply a coordinate of $\hat{X}_t$. Thanks to [45], Proposition 8.8, the conclusion of Lemma A.1 is thus direct in this case, which is sufficient for our purpose.

According to [56], Definition (23.10), a more generic definition of a path space is introduced by means of the following four generic notations: (i) shift operators, that translate paths in time ([56], Def. (2.2)), (ii) splicing maps, that allow to concatenate paths ([56], Def. (22.2)), (iii) killing and (iv) stopping operators, that allow to define an altered path for which, after a given time, the process X is respectively sent to the cemetery ([56], Def. (11.3)) or fixed at its current location ([56], Def. (23.9)). To have good enough measurability properties appears to be a main issue in such an abstract context. It is thus proposed to exploit a specific topology, namely the Ray topology on the augmented state space $\mathbb{X}_\partial := \mathbb{X} \cup \{\partial\}$ ([56], Chap. 17), which makes a rigorous justification quite laborious.

Acknowledgements. I am very grateful to Etienne Pardoux, my PhD supervisor, for his great support all along the redaction of this article. The comments of my reviewers have in addition been particularly helpful to review the writing, the structure of the article and focus more on its specific contribution to the field. Pierre Picco and Pierre-André Zitt also were of great help on this aspect. With the purpose of exploiting this result, Jules Olayé came up to me with insightful questions and comments that contributed to the current form of $(A3_F)$. I express my sincere thanks to them. I would like finally to thank the very inspiring meetings and discussions brought about by the Chair "Modélisation Mathématique et Biodiversité" of VEOLIA-Ecole Polytechnique-MnHn-FX.

References

[1] E. Arjas, E. Nummelin and R.L. Tweedie, Semi-Markov processes on a general state space: α-theory and quasi-stationarity. *J. Austral. Math. Soc. Ser. A* **30** (1980) 187–200.
[2] M. Alfaro, P. Gabriel and O. Kavian, Confining integro-differential equations originating from evolutionary biology: ground states and long time dynamics. *Discrete Cont. Dyn. Syst. B* **28** (2023) 5905–5933.
[3] R. Bürger, Perturbations of positive semigroups and applications to population genetics. *Math. Z.* **197** (1988) 259–272.
[4] R. Bürger and I.M. Bomze, Stationary distributions under mutation-selection balance: structure and properties. *Adv. Appl. Prob.* **28** (1996) 227–251.
[5] Bansaye, B. Cloez and P. Gabriel, Ergodic behavior of non-conservative semigroups via generalized Doeblin's conditions. *Proc. Acta Appl. Math.* **166** (2019) 29–72.
[6] Bansaye, B. Cloez, P. Gabriel and A. Marguet, A non-conservative Harris' ergodic theorem. *J. London Math. Soc.* **106** (2022) 2459–2510.
[7] O. Bonnefon, J. Coville and G. Legendre, Concentration phenomenon in some non-local equation. *Discrete Contin. Dyn. Syst.-B* **22** (2017) 763–781.
[8] H. Berestycki, J. Coville and H.-H. Vo, Persistence criteria for populations with non-local dispersion. *J. Math. Biol.* **72** (2016) 1693–1745.
[9] H. Berestycki, O. Diekmann, C.J. Nagelkerke and P.A. Zegeling, Can a species keep pace with a shifting climate?. *Bull. Math. Biol.* **71** (2009) 399–429.

[10] H. Berestycki, F. Hamel and L. Roques, Analysis of the periodically fragmented environment model. I. Species persistence. *J. Math. Biol.* **51** (2005) 75–113.

[11] H. Berestycki, F. Hamel and L. Roques, Reaction-diffusion equations for population dynamics with forced speed. I. The case of the whole space. *Discrete Contin. Dyn. Syst.-B* **21** (2008) 41–67.

[12] H. Berestycki, F. Hamel and L. Roques, Reaction-diffusion equations for population dynamics with forced speed. II. Cylindrical-type domains. *Discrete Contin. Dyn. Syst.-B* **25** (2009) 19–61.

[13] H. Berestycki, L. Nirenberg and S.R.S. Varadhan, The principal eigenvalue and maximum principle for second-order elliptic operators in general domains. *Commun. Pure Appl. Math.* **XLVII** (1994) 47–92.

[14] T. Britton and E. Pardoux, *Stochastic epidemic models with inference*, Part I. Lecture Notes in Math. **2255** (2019).

[15] J. Bertoin and A.-R. Watson, The strong Malthusian behavior of growth-fragmentation processes. *Annales Henri Lebesgue* **3** (2020) 795–823.

[16] R.S. Cantrell and C. Cosner, On the effects of spatial heterogeneity on the persistence of interacting species. *J. Math. Biol.* **37** (1998) 103–145.

[17] N. Champagnat, K.A. Coulibaly-Pasquier and D. Villemonais, Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions. *Sémin. Probab. XLIX* **2215** (2018) 165–182.

[18] R. Chazottes, P. Collet and S. Méléard, On time scales and quasi-stationary distributions for multitype birth-and-death processes. *Ann. Inst. H. Poincaré Prob. Stat.* **55** (2019) 2249–2294.

[19] N. Champagnat and D. Villemonais, Exponential convergence to quasi-stationary distribution and Q-process. *Prob. Theory Rel. Fields* **164** (2016) 243–283.

[20] N. Champagnat and D. Villemonais, Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing. *ALEA - Lat. Am. J. Prob. Math. Stat.* **XIV** (2017) 177–199.

[21] N. Champagnat and D. Villemonais, Uniform convergence of conditional distributions for absorbed one-dimensional diffusions. *Adv. Appl. Prob.* **50** (2018) 178–203.

[22] N. Champagnat and D. Villemonais, Uniform convergence of penalized time-inhomogeneous Markov processes. *ESAIM: Prob. Stat.* **22** (2018) 129–162.

[23] N. Champagnat and D. Villemonais, Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. *Stoch. Proc. Appl.* **135** (2021) 51–74.

[24] N. Champagnat and D. Villemonais, General criteria for the study of quasi-stationarity. *Electron. J. Probab.* **28** (2023) 1–84.

[25] B. Cloez and P. Gabriel, On an irreducibility type condition for the ergodicity of nonconservative semigroups. *Comptes Rendus. Mathématique* **358** (2020) 733–742.

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

[27] P. Collet, S. Martínez, S. Méléard and J. San Martín, Quasi-stationary distributions for structured birth and death processes with mutations. *Prob. Theory Rel. Fields* **151** (2011) 191–231.

[28] J. Coville, On a simple criterion for the existence of a principal eigenfunction of some nonlocal operators. *J. Differ. Equ.* **249** (2010) 2921–2953.

[29] J. Coville, Singular measure as principal eigenfunction of some nonlocal operators. *Appl. Math. Lett.* **26** (2013) 831–835.

[30] J. Coville, J. Dávila and S. Martínez, Existence and uniqueness of solutions to a nonlocal equation with monostable nonlinearity. *SIAM J. Math. Anal.* **39** (2008) 1693–1709.

[31] J. Coville, J. Dávila and S. Martínez, Pulsating fronts for nonlocal dispersion and KPP nonlinearity. *Ann. Instit. H. Poincare - AN* **30** (2013) 179–223.

[32] J. Coville and F. Hamel, On generalized principal eigenvalues of non-local operators with a drift. *Nonlinear Anal.* **193** (2020) 111569.

[33] B. Cloez, B. de Saporta and T. Roget, Long-time behavior and Darwinian optimality for an asymmetric size-structured branching process. *J. Math. Biol.* **83** (2021).

[34] D.J. Daley and D. Vere-Jones, An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure, 2nd ed. Prob. and Its Appl., Springer (2008).

[35] P. Del Moral, Mean Field Simulation for Monte Carlo Integration. Chapman & Hall/CRC Monographs on Statistics & Applied Probability (2013).

[36] P. Del Moral and L. Miclo, On the stability of nonlinear Feynman-Kac semigroups. *Ann. Fac. Sci. Toulouse Math.* **11** (2002) 135–175.

[37] P. Del Moral and D. Villemonais, Exponential mixing properties for time inhomogeneous diffusion processes with killing. *Bernoulli* **24** (2018) 1010–1032.

[38] C. Dellacherie and P.A. Meyer, Probabilities and Potential. North Holland (2011).

[39] P.A. Ferrari, H. Kesten, S. Martinez and P. Picco, Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Prob.* **23** (1995) 501–521.

[40] G. Ferré, M. Rousset and G. Stoltz, More on the long time stability of Feynman–Kac semigroups. *Stoch. PDE: Anal. Comp.* (2020) 1–44.

[41] J. Garcia-Melian and J.D. Rossi, On the principal eigenvalue of some nonlocal diffusion problems. *J. Differ. Equ.* **246** (2009) 21–38.

[42] R. Hardy and S.C. Harris, A spine approach to branching diffusions with applications to Lp-convergence of martingales. *Sémin. Probab. XLII.* Lect. Notes Math. **1979** (2009) 281–330.

[43] E. Hingant and R. Yvinec, Quasi-stationary distribution and metastability for the stochastic Becker-Döring model. *Electron. Commun. Prob.* **26** (2021) 1–14.
[44] L.I. Ignat, J.D. Rossi and A. San Antolin, Lower and upper bounds for the first eigenvalue of nonlocal diffusion problems in the whole space. *J. Differ. Equ.* **252** (2012) 6429–6447.
[45] O. Kallenberg, Foundations of Modern Probability, 2nd edn. Springer (2002).
[46] M. Kimura, A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. U.S.A.* **54** (1965) 731–736.
[47] I. Kontoyiannis and S.P. Meyn, Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Prob.* **13** (2003) 304–362.
[48] F. Li, J. Coville and X. Wang, On eigenvalue problems arising from nonlocal diffusion models. *Discrete Contin. Dyn. Syst.* **37** (2017) 879–903.
[49] M. Mariani, E. Pardoux and A. Velleret, Metastability between the clicks of the Müller ratchet. preprint arXiv:2007.14715v2 (2022).
[50] S. Méléard and D. Villemonais, Quasi-stationary distributions and population processes. *Prob. Surv.* **9** (2012) 340–410.
[51] S.P. Meyn and R.L. Tweedie, Markov Chains and Stochastic Stability. Springer-Verlag, London (1993).
[52] P. Michel, S. Mischler and B. Perthame, General relative entropy inequality: an illustration on growth models. *J. Math. Pures Appl.* **84** (2005) 1235–1260.
[53] E. Pardoux, Probabilistic Models of Population Evolution : Scaling Limits, Genealogies and Interactions. Springer (2016).
[54] P.K. Pollett, Quasi-stationary distributions: a bibliography. Available at people.smp.uq.edu.au/PhilipPollett/papers/qsds/qsds.html (2015).
[55] L.C.G. Rogers and D. Williams, Diffusions, Markov Processes, and Martingales, Vol. 1. Cambridge Math. Library (2000).
[56] M. Sharpe, General Theory of Markov Processes, Academic Press, Boston (1988).
[57] W. Shen and X. Xie, On principal spectrum points/principal eigenvalues of nonlocal dispersal operators and applications. *Disc. Contin. Dyn. Syst.* **35** (2015) 1665–1696.
[58] D.B. Smith, A sufficient condition for the existence of a principal eigenvalue for nonlocal diffusion equations with applications. *J. Math. Anal. Appl.* **418** (2014) 766–774.
[59] R.L. Tweedie, R-Theory for Markov chains on a general state space I: solidarity properties and R-recurrent chains. *Ann. Prob.* **2** (1974) 840–864.
[60] R.L. Tweedie, Quasi-stationary distributions for Markov chains on a general state space. *J. Appl. Prob.* **11** (1974) 726–741.
[61] N. Torres and D. Salort, Dynamics of neural networks with elapsed time model and learning processes. *Acta Appl. Math.* **170** (2020) 1065–1099.
[62] E.A. van Doorn and P.K. Pollett, Quasi-stationary distributions for discrete-state models. *Eur. J. Oper. Res.* **230** (2013) 1–14.
[63] A. Velleret, Unique quasi-stationary distribution, with a possibly stabilizing extinction. *Stoch. Proc. Appl.* **148** (2022) 98–138.
[64] A. Velleret, Exponential quasi-ergodicity for processes with discontinuous trajectories. second preprint version on ArXiv : arXiv:1902.01441v2 (2019).
[65] A. Velleret, Adaptation of a population to a changing environment under the light of quasi-stationarity. *Adv. Appl. Prob.* (2023) 1–52.
[66] A. Velleret, Two level natural selection with a quasi-stationarity approach. *Disc. Cont. Dyn. Syst.-B* (2023) early access, doi:https://doi.org/10.3934/dcdsb.2023122.
[67] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11** (1971) 155–167.