


## ASYMPTOTICS OF QUASI-STATIONARY DISTRIBUTIONS OF SMALL NOISE STOCHASTIC DYNAMICAL SYSTEMS IN UNBOUNDED DOMAINS

AMARJIT BUDHIRAJA,\*  
NICOLAS FRAIMAN ,\* \*\* AND  
ADAM WATERBURY,\* *University of North Carolina at Chapel Hill*

### Abstract

We consider a collection of Markov chains that model the evolution of multitype biological populations. The state space of the chains is the positive orthant, and the boundary of the orthant is the absorbing state for the Markov chain and represents the extinction states of different population types. We are interested in the long-term behavior of the Markov chain away from extinction, under a small noise scaling. Under this scaling, the trajectory of the Markov process over any compact interval converges in distribution to the solution of an ordinary differential equation (ODE) evolving in the positive orthant. We study the asymptotic behavior of the quasi-stationary distributions (QSD) in this scaling regime. Our main result shows that, under conditions, the limit points of the QSD are supported on the union of interior attractors of the flow determined by the ODE. We also give lower bounds on expected extinction times which scale exponentially with the system size. Results of this type when the deterministic dynamical system obtained under the scaling limit is given by a discrete-time evolution equation and the dynamics are essentially in a compact space (namely, the one-step map is a bounded function) have been studied by Faure and Schreiber (2014). Our results extend these to a setting of an unbounded state space and continuous-time dynamics. The proofs rely on uniform large deviation results for small noise stochastic dynamical systems and methods from the theory of continuous-time dynamical systems.

In general, QSD for Markov chains with absorbing states and unbounded state spaces may not exist. We study one basic family of binomial-Poisson models in the positive orthant where one can use Lyapunov function methods to establish existence of QSD and also to argue the tightness of the QSD of the scaled sequence of Markov chains. The results from the first part are then used to characterize the support of limit points of this sequence of QSD.

*Keywords:* Quasi-stationary distributions; uniform large deviation principles; random perturbations; long-time behavior

2020 Mathematics Subject Classification: Primary 60J10; 34F05  
Secondary 60F10; 92D25

---

Received 15 November 2019; revision received 20 April 2021; accepted 29 April 2021.

\* Postal address: Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC 27599, USA.

\*\* Email address: [fraiman@email.unc.edu](mailto:fraiman@email.unc.edu)

© The Author(s), 2022. Published by Cambridge University Press on behalf of Applied Probability Trust.

## 1. Introduction

In this work we study discrete-time Markov chains with values in the  $d$ -dimensional positive orthant that are absorbed upon hitting the boundary of the orthant. Such processes are well suited to model biological and ecological systems [9, 11] where each coordinate represents the population size of individuals of a given type/species. One of the fundamental issues in mathematical biology is to characterize the conditions for a population of interacting species to coexist, that is, to survive for a long time with no extinctions. Many real-world systems are certain to go extinct eventually, yet appear to be stationary over any reasonable time scale. Generally, the finite nature of the resources available prevents the system from growing without limit. Thus, provided we wait long enough, a sufficiently strong downward fluctuation in population size is bound to occur. We are interested in studying the long-term behavior of such systems away from extinction, under a suitable scaling of the system.

The processes we consider have a natural scaling parameter ( $N$ ) representing the system size. From standard results, as  $N \rightarrow \infty$ , the linearly interpolated trajectory of the state process  $X^N$ , over any compact time interval  $[0, T]$ , converges in distribution in  $C([0, T]; \mathbb{R}_+^d)$  (the space of continuous functions from  $[0, T]$  to  $\mathbb{R}_+^d$ , equipped with the uniform topology) to the solution of an ordinary differential equation (ODE) of the form  $\dot{\varphi}(t) = G(\varphi(t))$ ,  $\varphi(0) = x$  (see (4)). Our goal is to analyze the limiting behavior of the steady states of  $X^N$ , conditioned on non-extinction, as  $N \rightarrow \infty$ , in terms of the properties of the flow determined by the above ODE. The steady state of a Markov chain conditioned on non-extinction is made precise through the notion of a quasi-stationary distribution (QSD) (see Definition 1). We refer the reader to [14] for a comprehensive background and survey of results in the theory of QSD. QSD are important objects in biological models, and discussions of applications in biology can be found in [1, 7, 8, 15, 16].

Our first main result (Theorem 1) studies asymptotics of QSD of  $X^N$  (denoted by  $\mu_N$ ), as  $N \rightarrow \infty$ , provided they exist and the sequence  $\{\mu_N\}$  is tight. Specifically, in Theorem 1 we show that, under Assumptions 1, 2, 3, and 4, any limit point  $\mu$  of the sequence of QSD  $\{\mu_N\}$  is invariant under the flow determined by the ODE (4) and is supported on the union of interior attractors of the flow. We also provide lower bounds on the probability of non-extinction over a fixed time horizon that scale exponentially in system size. These bounds readily give similar lower bounds on expected time to extinction.

In general, Markov chains with absorbing states and an unbounded state space may fail to have a QSD. Conditions for existence of QSD have been studied in [5, 18, 19]; however, these results are not easily applicable to the models considered in this work. We instead make use of the recent work of Champagnat and Villemonais [2], which gives general and broadly applicable Lyapunov-function-based Foster-type criteria for existence of QSD (see Theorem 10). In our second main result we consider a basic family of Markov chains, to which we refer as binomial-Poisson models, where the results of [2] can be applied to give existence of QSD. Using the stability properties of these Markov chains we obtain bounds on exponential moments of certain hitting times that allow us to construct suitable Lyapunov functions (and related objects) for which the conditions in Theorem 10 are satisfied, thus establishing the existence of a QSD  $\mu_N$  for each  $N$ . In fact, this QSD can be characterized as the limit, as  $n \rightarrow \infty$ , of the law of  $X_n^N$ , conditioned on non-extinction, starting from an arbitrary initial condition in the interior. Using this characterization, and similar moment estimates as used in the construction of the Lyapunov functions, we then argue that the sequence of QSD is tight. Finally, from these results and other properties of

the model, we establish our second main result (Theorem 2), which says that the binomial-Poisson model introduced in Section 2.3 satisfies all the conditions in Theorem 1 and therefore provides an important class of Markov chains where the conclusions of Theorem 1 hold.

### 1.1. Approach and idea of proof

We now comment on the proof of Theorem 1. Our results are motivated by the work of Faure and Schrieber [4] (see also the unpublished manuscript of Marmet [13]), which considers analogous problems for a class of Markov chains where the deterministic dynamical system obtained under the scaling limit is given by a discrete-time evolution equation and the dynamics are essentially in a compact space (namely, the one-step map is a bounded function). As in [4], one of the important ingredients in the proof is an analysis of the large deviation behavior of the sequence of small noise Markov chains in Section 2.1. However, because of the continuous-time setting here, one needs to study large deviation principles on suitable path spaces. One of the issues that arise in the large deviation analysis is that transition probabilities of the Markov chain behave in a degenerate manner near the boundaries. For this reason, the associated local rate functions have poor regularity properties, which in turn makes establishing a global large deviation principle on the path space technically challenging. Another issue arises from the unboundedness of the state space. In particular, the moment generating functions of the noise sequences can become arbitrarily large as the system state becomes large. In order to handle these issues, we instead consider large deviation principles for a collection of modified chains in  $\mathbb{R}^d$ . These modified chains behave identically to the original chain until exiting from a given compact set  $K$  in the interior of the orthant; upon exiting, the modified chains change their behavior to a more regular dynamics in an appropriate sense. The large deviation estimates that are needed for our analysis can be obtained by piecing together such large deviation principles associated with all such compact sets  $K$ . A similar approach, in a setting where the state space is compact, has been proposed in [13]. Another important point in the analysis is that one needs large deviation estimates that are uniform in initial condition in compact sets, in the sense of Freidlin and Wentzell [6, Chapter 3.3, pp. 91–92]. For this we use results on uniform Laplace principles for small noise stochastic difference equations that have been developed in [3, Section 6.7]. The recent work [17] shows that a uniform Laplace principle implies a uniform large deviation principle in the sense of Freidlin and Wentzell. These results together allow us to establish uniform probability estimates that are needed in our large deviation analysis (see Section 4).

The proof of Theorem 1, analogous to [4], also requires a detailed analysis of the dynamical system properties of the flow associated with the ODE (4). In particular, a careful understanding of the properties of continuous-time analogues of absorption-preserving pseudo-orbits (in the terminology of [4]) and those of the associated recurrence classes are key to the proof (see Section 3). Although some of the arguments are similar to those of [4], there are new challenges that arise due to the unboundedness of the state space and the continuous-time dynamics. To handle these features we exploit the stability properties of the underlying ODE and develop several a priori estimates for pseudo-orbits that are uniform in time and/or space. The dynamical systems results in Section 3 and the large deviation estimates in Section 4 take us most of the way to the proof of Theorem 1. In particular, in Section 5, using these results, we establish the lower bound on probabilities of non-extinction given in Theorem 1 and also that the limit points  $\mu$  of the QSD are invariant under the flow, they do not charge the boundary, and in fact that they

are supported on the union of absorption-preserving recurrence classes in the interior. The final step is to show that the support in fact lies in the union of the interior attractors. For this, following [4], we reformulate the notion of recurrence in terms of the quasipotential associated with the rate functions in the underlying large deviation principles. Section 6 introduces the quasipotential and this alternative notion of recurrence and proves the equivalence between these two definitions of recurrence classes. The second definition is better suited to the analysis and allows the use of the large deviation estimates of Section 4 in studying the behavior of the stochastic dynamical system in terms of the properties of the recurrence classes. Combining the results of Section 6 with the results of Section 4 and the properties of absorption-preserving pseudo-orbits studied in Section 3, we complete the proof of the main result in Section 7.

### 1.2. Organization

The paper is organized as follows. In Section 2 we introduce the model of interest, state the assumptions, and present the main results of the paper. In Section 3 we introduce some notions from the theory of dynamical systems, and study properties of recurrence points and associated (pseudo-) orbits for the dynamical system associated with the law-of-large-numbers limit of the underlying sequence of scaled Markov chains. In Section 4 we establish some key large deviation estimates. In Section 5 we give some important asymptotic properties of QSD (provided they exist) for the Markov chains considered in this work. In Section 6 we introduce the quasipotential  $V$  that governs the large deviation behavior of the model and study the properties of  $V$ -chain recurrence. In Section 7 we complete the proof of our first main theorem, namely Theorem 1. Finally, in Section 8 we prove our second main result, Theorem 2, which gives an important family of models for which Theorem 1 can be applied.

### 1.3. Notation

Let  $\Delta \doteq \mathbb{R}_+^d$ ,  $\Delta^o \doteq \{x \in \Delta : x > 0\}$ , where inequalities for vectors are interpreted componentwise, and  $\partial\Delta \doteq \Delta \setminus \Delta^o$ . For  $N \in \mathbb{N}$ , let  $\Delta_N \doteq \Delta \cap \frac{1}{N}\mathbb{Z}^d$ ,  $\partial\Delta_N \doteq \partial\Delta \cap \frac{1}{N}\mathbb{Z}^d$ , and  $\Delta_N^o \doteq \Delta^o \cap \frac{1}{N}\mathbb{Z}^d$ . For  $x, y \in \mathbb{R}^d$ ,  $\langle x, y \rangle \doteq \sum_{i=1}^d x_i y_i$ . For  $x \in \mathbb{R}^d$  and  $A \subset \mathbb{R}^d$ ,  $\text{dist}(x, A) \doteq \inf_{y \in A} \|x - y\|$ . We denote by  $\mathcal{N}^\varepsilon(A)$  the  $\varepsilon$ -neighborhood of a set  $A$  in  $\Delta$ , namely  $\mathcal{N}^\varepsilon(A) \doteq \{x \in \Delta : \text{dist}(x, A) < \varepsilon\}$ . For  $r > 0$  and  $x \in \mathbb{R}^d$ ,  $B_r(x)$  will denote the open ball of radius  $r$  centered at  $x$ . Denote by  $\mathcal{P}(S)$  the space of probability measures on a Polish space  $S$ , equipped with the topology of weak convergence. For a  $\mu \in \mathcal{P}(S)$  and  $\mu$ -integrable  $f : S \rightarrow \mathbb{R}$ , we write  $\int f d\mu$  as  $\mu(f)$ . The support of  $\mu \in \mathcal{P}(S)$  will be denoted by  $\text{supp}(\mu)$ . For a signed measure  $\eta$  on  $S$ ,  $\|\eta\|_{TV}$  denotes its total variation norm, namely

$$\|\eta\|_{TV} = \sup_f \left| \int f d\eta \right|,$$

where the supremum is taken over all measurable maps  $f : S \rightarrow \mathbb{R}$  such that  $\sup_{x \in S} |f(x)| \leq 1$ . For a bounded  $F : S \rightarrow \mathbb{R}$ , we denote  $\sup_{x \in S} |F(x)|$  by  $\|F\|_\infty$ . We denote by  $\mathcal{K}$  the collection of all convex compact subsets with a nonempty interior that are contained in  $\Delta^o$ . For  $T < \infty$ , we denote by  $C([0, T] : S)$  the space of continuous functions from  $[0, T]$  to  $S$ , equipped with the uniform topology. For  $\phi \in C([0, T] : \mathbb{R}^d)$ , let  $\|\phi\|_{*,T} \doteq \sup_{0 \leq t \leq T} \|\phi(t)\|$ . Given a metric space  $S_1$  and a Polish space  $S_2$ , a stochastic kernel  $x \mapsto \theta(dy|x)$  on  $S_2$  given  $S_1$  is a measurable map from  $S_1$  to  $\mathcal{P}(S_2)$ .

## 2. Statement of results

### 2.1. The model

Consider the sequence  $\{X_k^N\}_{k \in \mathbb{N}_0}$  of  $\Delta_N$ -valued random variables defined as

$$\begin{aligned} X_{k+1}^N &= X_k^N + \frac{1}{N} \eta_{k+1}^N(X_k^N), & k \in \mathbb{N}_0, \\ X_0^N &= x^N, \end{aligned} \tag{1}$$

where for each  $x \in \Delta_N$ ,  $\eta_k^N(x)$  is a  $\mathbb{Z}^d$ -valued random variable with distribution  $\theta^N(\cdot|x)$  such that  $\text{supp}(\theta^N(\cdot|x)) \subset \prod_{i=1}^d [-Nx_i, \infty)$ .

We will denote by  $\mathbb{P}_\nu^N$  the probability measure under which the Markov chain  $\{X_k^N\}$  has the initial distribution  $\nu$ , namely  $\mathbb{P}_\nu^N(X_0^N \in A) = \nu(A)$ . If  $\nu = \delta_x$ , we write  $\mathbb{P}_\nu^N$  as simply  $\mathbb{P}_x^N$ .

**Definition 1.** A probability measure  $\mu_N$  on  $\Delta_N^o$  is said to be a quasi-stationary distribution (QSD) for the Markov chain  $\{X_k^N\}$  if for every  $n \in \mathbb{N}$

$$\mathbb{P}_{\mu_N}[X_k^N = j | X_k^N \in \Delta_N^o] = \mu_N(j), \quad \text{for all } j \in \Delta_N^o \text{ and } k \in \mathbb{N}.$$

### 2.2. Definitions and assumptions

Consider the continuous-time process  $\hat{X}^N$  obtained from a linear interpolation of  $X^N$ , given as

$$\hat{X}^N(t) = X_n^N + [X_{n+1}^N - X_n^N](Nt - n), \quad t \in [n/N, (n + 1)/N], \quad n \in \mathbb{N}_0, \tag{2}$$

The following assumption on the law-of-large-numbers behavior of  $\hat{X}^N$  will play a central role in our study of asymptotic properties of QSD of  $X^N$ .

**Assumption 1.** *There is a Lipschitz function  $G: \Delta \rightarrow \mathbb{R}^d$  such that for any sequence  $x_N \rightarrow x$  with  $x_N \in \Delta_N$  for every  $N \in \mathbb{N}$ ,*

$$\mathbb{P}_{x_N} \left( \sup_{0 \leq t \leq T} \|\hat{X}^N(t) - \varphi_t(x)\| > \varepsilon \right) \rightarrow 0, \quad \text{as } N \rightarrow \infty, \text{ for every } T \in [0, \infty) \text{ and } \varepsilon > 0, \tag{3}$$

where  $\{\varphi_t(x)\}_{t \geq 0}$  is the solution of the ODE

$$\dot{\varphi}(t) = G(\varphi(t)), \quad \varphi(0) = x. \tag{4}$$

We now introduce the notion of absorption-preserving pseudo-orbits for the flow associated with the ODE (4). Discrete-time analogues of these were introduced in [4].

**Definition 2.** Given  $\delta, T > 0$ , consider a family of points  $\xi = (\xi_0 = x, \dots, \xi_n = y) \in \Delta^{n+1}$  and a collection of times  $T \leq T_1, \dots, T_{n-1}$  such that

- $\|\xi_0 - \xi_1\| < \delta$ ;
- whenever  $\xi_i \in \partial \Delta, \xi_{i+1} \in \partial \Delta$ ;
- $\|\xi_{i+1} - \varphi_{T_i}(\xi_i)\| < \delta$  for  $1 \leq i \leq n - 1$ .

The piecewise continuous path

$$(x, \{\varphi_t(\xi_1): t \in [0, T_1]\}, \{\varphi_t(\xi_2): t \in [0, T_2]\}, \dots, \{\varphi_t(\xi_{n-1}): t \in [0, T_{n-1}]\}, y)$$

is said to be a  $(\delta, T)$  absorption-preserving pseudo-orbit (AP-pseudo-orbit) from  $x$  to  $y$ . Occasionally, we will also refer to the sequence  $\{\xi_i\}_{i=0}^n$  as a  $(\delta, T)$  AP-pseudo-orbit from  $x$  to  $y$ .

**Definition 3.** For two points  $x, y \in \Delta$ , we say that  $x <_{AP} y$  if for all  $\delta, T > 0$  there is a  $(\delta, T)$  AP-pseudo-orbit from  $x$  to  $y$ . If  $x <_{AP} y$  and  $y <_{AP} x$ , we write  $x \sim_{AP} y$ . If  $x \sim_{AP} x$ , then  $x$  is said to be an AP-chain recurrent point. Let  $\mathcal{R}_{AP}$  denote the set of AP-chain recurrent points, and note that  $\sim_{AP}$  is an equivalence relation on  $\mathcal{R}_{AP}$ . For  $x \in \mathcal{R}_{AP}$ , the equivalence class  $[x]_{AP}$  of all  $y \in \mathcal{R}_{AP}$  such that  $y \sim_{AP} x$  is said to be an AP-basic class. Such a class is called maximal if, whenever for some  $y \in \mathcal{R}_{AP}, x <_{AP} y$ , we have  $y \in [x]_{AP}$ . A maximal AP-basic class is called an AP-quasiattractor. We let  $\mathcal{R}_{AP}^* \doteq \mathcal{R}_{AP} \cap \Delta^o$ .

The following will be our main assumptions on the dynamical system  $\{\varphi_t(x)\}$ . Parts (c) and (d) say that the velocity fields decay as the boundaries are approached but not at too fast a rate. Part (e) is our main stability assumption on the dynamics. Parts (a) and (b) are requirements on recurrence classes for the flow that are satisfied quite broadly.

**Assumption 2.**

- (a) *There are a finite number of AP-basic classes contained in  $\Delta^o$ , which are denoted by  $\{K_i\}_{i=1}^v$ . Each  $K_i$  is a closed set. Additionally, for some  $l < v$ ,  $\{K_i\}_{i=1}^l$  are AP-quasiattractors and  $\{K_i\}_{i=l+1}^v$  are non-AP-quasiattractors.*
- (b) *For each  $i = 1, \dots, v$  there is an  $x_i \in K_i$  such that, for every  $T > 0$ ,  $\{\varphi_t(x_i): t \geq T\}$  is dense in  $K_i$ .*
- (c) *There exist  $\varepsilon > 0$  and  $m > 0$  such that for every  $i = 1, \dots, d$ ,  $G_i(x) > mx_i$  whenever  $x \in \Delta^o$  and  $x_i \leq \varepsilon$ .*
- (d) *For every  $i = 1, \dots, d$ , as  $\delta \rightarrow 0$ ,  $\sup_{x \in \Delta: x_i \leq \delta} G_i(x) \rightarrow 0$ .*
- (e) *For some  $\kappa \in (0, \infty)$  and  $M \in (1, \infty)$ ,  $\langle x, G(x) \rangle \leq -\kappa \|x\|^2$  for all  $x \in \Delta$  with  $\|x\| \geq M$ .*

We will need certain assumptions on the moment generating functions of  $\theta^N(\cdot | x)$ .

**Assumption 3.**

- (a) *For every  $N \in \mathbb{N}$ ,  $\zeta \in \mathbb{R}^d$ , and  $x \in \Delta_N^o$*

$$H^N(x, \zeta) \doteq \log \int_{\mathbb{R}^d} \exp\{\langle \zeta, y \rangle\} \theta^N(dy|x) < \infty.$$

- (b) *There exists a stochastic kernel  $\theta(dy|x)$  on  $\mathbb{R}^d$  given  $\Delta^o$  such that the following hold:*
  - (i) *For every  $x \in \Delta^o$ , the convex hull of  $\text{supp}(\theta(\cdot | x))$  equals  $\mathbb{R}^d$ .*
  - (ii) *The map  $x \mapsto \theta(\cdot | x)$  is a continuous map from  $\Delta^o$  to  $\mathcal{P}(\mathbb{R}^d)$ .*
  - (iii) *For every  $\zeta \in \mathbb{R}^d$  and  $K \in \mathcal{K}$ ,  $\sup_{x \in K} H(x, \zeta) < \infty$ , where*

$$H(x, \zeta) \doteq \log \int_{\mathbb{R}^d} \exp\{\langle \zeta, y \rangle\} \theta(dy|x).$$

Furthermore, as  $N \rightarrow \infty$ ,

$$\sup_{x \in K \cap \Delta_N} |H^N(x, \zeta) - H(x, \zeta)| \rightarrow 0.$$

We introduce one final assumption to provide a lower bound on the probability that  $X^N$  is absorbed when its initial state is sufficiently close to  $\partial \Delta$ .

**Assumption 4.**

- i. For each  $N \in \mathbb{N}$  and  $x, y \in \Delta_N^o$ , there is a  $k \in \mathbb{N}$  such that  $\mathbb{P}_y^N(X_k^N = x) > 0$ .
- ii. For every  $\gamma \in (0, \infty)$  and  $T \in \mathbb{N}$ , there is an open neighborhood  $U_\gamma$  of  $\partial \Delta$  in  $\Delta$  such that

$$\liminf_{N \rightarrow \infty} \inf_{x \in U_\gamma \cap \Delta_N} \frac{1}{N} \log \mathbb{P}_x(\hat{X}^N(T) \in \partial \Delta) \geq -\gamma.$$

We now present our main results.

**2.3. Main results**

It is easy to see that under Assumption 2, for all  $x \in \Delta$  and  $t \geq 0$ ,  $\varphi_t(x) \in \Delta$ . In particular,  $\varphi_t$  is a measurable map from  $\Delta$  to itself for every  $t \geq 0$ . We recall the definition of an invariant measure for the flow  $\{\varphi_t\}$ .

**Definition 4.** A probability measure  $\mu$  on  $\Delta$  is  $\{\varphi_t\}$ -invariant if  $\mu(\varphi_t^{-1}(A)) = \mu(A)$  for every measurable  $A \subseteq \Delta$  and  $t > 0$ .

**Theorem 1.** Suppose that for every  $N \in \mathbb{N}$ , there exists a QSD  $\mu_N$  for  $\{X_n^N\}_{n \in \mathbb{N}_0}$ , and the sequence  $\{\mu_N\}$  is relatively compact as a sequence of probability measures on  $\Delta^o$ . Suppose that Assumptions 1, 2, 3, and 4 are satisfied. Then any weak limit point  $\mu$  of this sequence is  $\{\varphi_t\}$ -invariant and is supported on  $\cup_{i=1}^l K_i$ . Moreover, letting

$$\lambda_N \doteq [\mathbb{P}_{\mu_N}(X_1^N \in \Delta^o)]^N, \tag{5}$$

there is a  $c > 0$  and  $N_0 \in \mathbb{N}$  such that  $\lambda_N \geq 1 - e^{-cN}$  for all  $N \geq N_0$ .

We now introduce a basic family of Markov chains, which we refer to as the *binomial-Poisson models*, for which Theorem 1 can be applied.

Consider a population with  $d$  types of particles evolving in discrete time, in which, at each time step, any given particle dies with probability  $1/N$ , and given that the population size at previous time step was  $Nx = (Nx_i)_{i=1}^d$ , the number of particles of type  $i$  that are produced at the next time step follows a Poisson distribution with mean  $F_i(x)$  distribution for some  $F: \Delta \rightarrow \mathbb{R}_+^d$ . Denote the total number of particles of type  $i$  at time  $k$  by  $NX_k^{N,i}$ . The evolution of  $X_k^N = (X_k^{N,1}, \dots, X_k^{N,d})$  is then given by (1), where, for each  $N$ ,  $\theta^N(dy|x) \equiv \theta^{N,*}(dy|x)$  is the distribution of  $U - V$ , where  $U = (U_i)_{i=1}^d$ ,  $V = (V_i)_{i=1}^d$ ,  $\{U_i, V_j, i, j = 1, \dots, d\}$  are mutually independent,  $U_i \sim \text{Poi}(F_i(x))$  (namely, a Poisson random variable with mean  $F_i(x)$ ), and  $V_i \sim \text{Bin}(Nx_i, \frac{1}{N})$  (namely a binomial random variable with  $Nx_i$  trials and probability of success  $1/N$ ).

Define

$$\tau_\partial^N \doteq \inf \{k \in \mathbb{N}_0 : X_k^N \in \partial \Delta_N\}. \tag{6}$$



For a bounded and measurable  $f: \Delta_N \rightarrow \mathbb{R}$ ,

$$P_n^N f(x) \doteq \mathbb{E}_x [f(X_n^N); \tau_\partial^N > n]. \tag{7}$$

**Theorem 2.** *Suppose that, for each  $N$ ,  $X^N$  is given by (1) with  $\theta^N \equiv \theta^{N,*}$ . Further suppose that  $F$  is a bounded Lipschitz map and Parts (a)–(d) of Assumption 2.2 are satisfied with  $G(x) = F(x) - x$ . Then there is a  $\mu_N \in \mathcal{P}(\Delta_N^o)$  such that for every  $N \in \mathbb{N}$  and  $x_N \in \Delta_N^o$ ,*

$$\frac{\delta_{x_N} P_n^N}{\delta_{x_N} P_n^N(1_{\Delta_N^o})}$$

*converges to  $\mu_N$  in the total variation distance as  $n \rightarrow \infty$ . The measure  $\mu_N$  is a QSD for  $\{X^N\}$ . The sequence  $\{\mu_N\}_{N \in \mathbb{N}}$  is relatively compact as a sequence of probability measures on  $\Delta$ , and any weak limit point  $\mu$  of this sequence is  $\{\varphi_t\}$ -invariant and is supported by  $\cup_{i=1}^l K_i$ . Finally, letting  $\lambda_N \doteq [\mathbb{P}_{\mu_N}(X_1^N \in \Delta^o)]^N$ , there is a  $c > 0$  and  $N_0 \in \mathbb{N}$  such that  $\lambda_N \geq 1 - e^{-cN}$  for all  $N \geq N_0$ .*

Theorem 1 is proved in Section 7 while Theorem 2 is established in Section 8.

### 3. Absorption-preserving pseudo-orbits

In this section we present some basic facts on absorption-preserving pseudo-orbits that will be used to prove Theorem 1. Throughout the section we will take Assumptions 1 and 2 to hold.

The proofs of many of these results are similar to those found in [4] for discrete-time flows, but we provide the details for completeness. Recall that the solution of the ODE (4) with initial value  $\varphi(0) = x$  is denoted by  $\{\varphi_t(x)\}_{t \geq 0}$ . The following lemma is a consequence of the stability condition in Assumption 2(e).

**Lemma 1.** *For every  $T > 0$  and compact  $A \subset \Delta$ , there is a  $\delta_0 > 0$  and a compact  $A_1 \subset \Delta$  such that for any  $(\delta_0, T)$  AP-pseudo-orbit  $\{\xi_i\}_{i=0}^{n+1}$  with  $\xi_0 \in A$ , we have  $\xi_i \in A_1$  for all  $i = 0, \dots, n + 1$ .*

*Proof.* For fixed  $x \in \Delta$ ,  $\|\varphi_t(x)\|^2$  solves the ODE

$$\frac{d}{dt} \|\varphi_t(x)\|^2 = 2\langle G(\varphi_t(x)), \varphi_t(x) \rangle.$$

From Assumption 2(e), when  $\|x\| \geq M$

$$2\langle G(x), x \rangle \leq -2\kappa \|x\|^2.$$

This implies the following two facts:

- (a) If for any  $R \geq M$ ,  $x \in B_R \doteq \{z: \|z\| \leq R\}$ , then  $\varphi_t(x) \in B_R$  for every  $t \geq 0$ .
- (b) Given  $T > 0$ , define  $\delta_0 = \delta_0(T) \doteq \frac{\kappa T}{2} \wedge 1$ . Then for any  $\delta \leq \delta_0$ , and any  $R \geq M$ , whenever  $x \in B_{R+\delta}$ , we have that  $\varphi_t(x) \in B_R$  for all  $t \geq T$ .

Now fix  $T > 0$  and a compact  $A \subset \mathbb{R}_+^d$ . Without loss of generality assume that there is an  $R \geq M$  such that  $A \subset B_R$ . Let  $\delta_0 = \delta_0(T)$  be as defined above and consider a  $(\delta_0, T)$  AP-pseudo-orbit  $\{\xi_i\}_{i=0}^{n+1}$  with  $\xi_0 \in A$ . Then the above two facts imply that  $\xi_i \in B_{R+1}$  for all  $i = 0, 1, \dots, n + 1$ . The result follows on taking  $A_1 = B_{R+1}$ . □



As a consequence of Lemma 1 we get the following result on the boundedness of AP–basic classes.

**Lemma 2.** *The AP–basic classes are bounded.*

*Proof.* Fix  $x \in \mathcal{R}_{AP}$  and  $y \in [x]_{AP}$ . Let  $T > 0$  and  $A = \{x\}$ . From Lemma 1, there is a  $\delta_0 > 0$  and a compact  $A_1$  in  $\Delta$  such that for each  $\delta \leq \delta_0$ , any  $(\delta, T)$  AP–pseudo-orbit starting at  $x$  is contained in  $A_1$ . Since  $y \in [x]_{AP}$ , there must exist a  $(\delta, T)$  AP–pseudo-orbit from  $x$  to  $y$ , which means that  $y$  must lie in  $A_1$ . The result follows.  $\square$

For  $x \in \Delta^\circ$ , we denote the forward orbit of  $\varphi$  by  $\gamma^+(x) \doteq \{\varphi_t(x) | t \geq 0\}$ . From Assumption 2(b) and arguments as in Lemma 1 the following result is immediate.

**Lemma 3.** *The following hold:*

- (a) *There exists  $\alpha_0 \in (0, 1)$  such that if for some  $\alpha \in (0, \alpha_0]$  and  $x \in \Delta^\circ$ ,  $\text{dist}(x, \partial\Delta) \geq \alpha$ , then for all  $t \geq 0$ ,  $\text{dist}(\varphi_t(x), \partial\Delta) > \alpha$ .*
- (b) *There exists  $M_0 \in (0, \infty)$  such that if for some  $M \geq M_0$  and  $x \in \Delta^\circ$ ,  $\|x\| \leq M$ , then for all  $t \geq 0$ ,  $\|\varphi_t(x)\| < M$ .*
- (c) *For every  $A \in \mathcal{K}$ , there exist  $T > 0$ ,  $A_1, A_2 \in \mathcal{K}$  such that  $A_1 \supset A$ ,  $A_2 \subset A_1$ ,  $\text{dist}(A_2, \partial A_1) > 0$ , and for all  $x \in A_1$  and  $t \geq T$ ,  $\varphi_t(x) \in A_2$ .*
- (d) *For every  $A_0 \in \mathcal{K}$ , there is an  $A_1 \in \mathcal{K}$  such that for every  $x \in A_0$ , the forward orbit  $\gamma^+(x) \subset A_1$ .*

The proof of the next lemma follows from the observation (a) in the proof of Lemma 1.

**Lemma 4.** *For each compact  $K \subset \Delta$ ,  $\sup_{x \in K} \sup_{t \geq 0} \|\varphi_t(x)\| < \infty$ .*

We say a  $(\delta, T)$  AP–pseudo-orbit described by a collection of points  $\xi = (\xi_0, \dots, \xi_n) \in \Delta^{n+1}$  and a collection of times  $T \leq T_1, \dots, T_{n-1}$  intersects a set  $A \subset \Delta$ , if for some  $j \in \{1, \dots, n-1\}$  and  $t \in [0, T_j]$ ,  $\varphi_t(\xi_j) \in A$ . We say such an orbit lies in  $A$  if its intersection with  $A^c$  is empty. The following lemma shows that for small  $\delta$  and large  $T$ ,  $(\delta, T)$  AP–pseudo-orbits starting from the interior stay away from the boundary.

**Lemma 5.** *Suppose  $A \in \mathcal{K}$ . Then there exist  $\varepsilon_0 > 0$ ,  $T > 0$ ,  $\delta > 0$  such that any  $(\delta, T)$  AP–pseudo-orbit  $\{\xi_k\}_{k=0}^n$  with  $\xi_0 \in A$  does not intersect  $E_{\varepsilon_0} \doteq \{x \in \Delta : x_i \leq \varepsilon_0 \text{ for some } i = 1, \dots, d\}$ . In particular, there is an  $A_1 \in \mathcal{K}$  such that any such AP–pseudo-orbit starting in  $A$  lies in  $A_1$ .*

*Proof.* Let  $\varepsilon_1 \doteq \text{dist}(A, \partial\Delta)$  and let  $\varepsilon$  and  $m$  be as in Assumption 2(c). Let  $\varepsilon_0 \doteq (\varepsilon \wedge \varepsilon_1)/4$ . Note that for any  $x \in \Delta$  and  $i = 1, \dots, d$ ,

$$\frac{d}{dt}([\varphi_t(x)]_i)^2 = 2[\varphi_t(x)]_i G_i(\varphi_t(x)) > 2m([\varphi_t(x)]_i)^2 \quad \text{whenever } [\varphi_t(x)]_i \leq \varepsilon. \tag{8}$$

Since  $m > 0$ , we can choose a  $T > 0$  such that for any  $x \in \Delta$  and  $i = 1, \dots, d$  with  $x_i \geq \varepsilon_0$ , we have  $[\varphi_t(x)]_i > 3\varepsilon_0$  for all  $t \geq T$ . Fix  $\delta \in (0, \varepsilon_0)$ . Consider a  $(\delta, T)$  AP–pseudo-orbit  $\{\xi_k\}_{k=0}^n$  with  $\xi_0 \in A$  and associated time instants  $T \leq T_1, \dots, T_{n-1}$ . Clearly  $\xi_0 \notin E_{2\varepsilon_0}$ , and by (8),  $\varphi_t(\xi_0) \notin E_{2\varepsilon_0}$  for all  $t \in [0, T_1]$ . Also, by our choice of  $T$ ,  $\varphi_{T_1}(\xi_0) \notin E_{3\varepsilon_0}$  and consequently  $\xi_1 \notin E_{2\varepsilon_0}$ .

A recursive argument now shows that the pseudo-orbit has no intersection with  $E_{\varepsilon_0}$ . The result follows.  $\square$

We now recall a definition from the theory of dynamical systems.

**Definition 5.** The  $\omega$ -limit set of  $B \subset \Delta$  is

$$\omega(B) \doteq \{x \in \Delta : \text{there is a sequence } t_n \uparrow \infty \text{ and a sequence } x_n \in B \text{ such that } \varphi_{t_n}(x_n) \rightarrow x\},$$

so for  $x \in \Delta$ ,

$$\omega(x) \doteq \{y \in \Delta : \text{there is a sequence } t_n \uparrow \infty \text{ such that } \varphi_{t_n}(x) \rightarrow y\}.$$

The result below follows from classical arguments and on observing that under Assumption 2(b), if  $x \in \Delta^o$ , then  $\omega(x) \subset \Delta^o$ . For a proof of the lemma in the discrete-time setting see [4]. The proof for the continuous-time setting considered here is similar, so we omit details.

**Lemma 6.** For any  $x \in \Delta$ ,  $\omega(x) \subset \mathcal{R}_{AP}$ .

The following lemma gives a useful property of an AP-quasiattractor.

**Lemma 7.** If  $[x]_{AP}$  is maximal, then  $x <_{AP} z$  if and only if  $z \in [x]_{AP}$ .

*Proof.* Suppose that  $x <_{AP} z$ . In order to show that  $z \in [x]_{AP}$ , it is enough to show that  $z <_{AP} x$ . Note that  $\omega(z)$  is nonempty. Let  $z' \in \omega(z)$ . From Lemma 6,  $z' \in \mathcal{R}_{AP}$ . We now show that  $z <_{AP} z'$ . Since  $z' \in \omega(z)$ , there is a sequence  $T_i \uparrow \infty$  such that  $\varphi_{T_i}(z) \rightarrow z'$ . Fix  $\delta, T > 0$ . Then we can find  $T' > T$  such that  $\|\varphi_{T'}(z) - z'\| < \delta$ . This shows that  $(z, z, z', z')$  is a  $(\delta, T)$  AP-pseudo-orbit from  $z$  to  $z'$ . Since  $\delta, T > 0$  are arbitrary, we have  $z <_{AP} z'$ . Combining this with  $x <_{AP} z$ , we now see that  $x <_{AP} z'$ . Since  $z' \in \mathcal{R}_{AP}$  and  $[x]_{AP}$  is maximal, we must have  $z' \in [x]_{AP}$ , and therefore  $z <_{AP} x$ . This completes the proof of the lemma.  $\square$

The following lemma provides an important invariance property of AP-classes under the flow  $\{\varphi_t\}$ .

**Lemma 8.** Any AP-basic class  $[x]_{AP}$  is positively  $\varphi_t$ -invariant for all  $t \geq 0$ :  $\varphi_t([x]_{AP}) \subset [x]_{AP}$ . Additionally, if  $[x]_{AP} \subset \Delta^o$ , then  $[x]_{AP}$  is  $\varphi_t$ -invariant for all  $t \geq 0$ :  $\varphi_t([x]_{AP}) = [x]_{AP}$ .

*Proof.* Let  $y \in [x]_{AP}$ . To begin, fix  $t, \delta, T > 0$ , and let  $T' > T + t$ . We can find some  $\delta_0 \doteq \delta_0(y) < \delta$  such that if  $\|y - x_0\| < \delta_0$ , then  $\|\varphi_t(y) - \varphi_t(x_0)\| < \delta$ . Since  $y \in \mathcal{R}_{AP}$ , there is a  $(\delta_0, T')$  AP-pseudo-orbit from  $y$  to  $y$ , which we denote by  $\xi = (y, \xi_1, \dots, \xi_{n-1}, y)$ , with corresponding time instants  $(T_1, \dots, T_{n-1})$ . Then  $\tilde{\xi} \doteq (\varphi_t(y), \varphi_t(\xi_1), \xi_2, \dots, \xi_n = y)$  is a  $(\delta, T)$  AP-pseudo-orbit from  $\varphi_t(y)$  to  $y$  with corresponding time instants  $(T_1 - t, T_2, \dots, T_{n-1})$ , since

$$\|\varphi_t(x) - \varphi_t(\xi_1)\| < \delta, \quad \text{and} \quad \|\varphi_{T_1-t}(\varphi_t(\xi_1)) - \xi_2\| = \|\varphi_{T_1}(\xi_1) - \xi_2\| < \delta.$$

Thus  $\varphi_t(y) <_{AP} y$ .

Next, define  $\tilde{\xi} \doteq (y, \xi_1, \dots, \xi_{n-1}, \varphi_t(y))$ , and note that  $\tilde{\xi}$  is a  $(\delta, T)$  AP-pseudo-orbit from  $y$  to  $\varphi_t(y)$  with time instants  $(T_1, \dots, T_{n-2}, T_{n-1} + t)$ , since

$$\|\varphi_{T_{n-1}}(\xi_{n-1}) - y\| < \delta_0,$$

which ensures that

$$\|\varphi_{T_{n-1}+t}(\xi_{n-1}) - \varphi_t(y)\| = \|\varphi_t(\varphi_{T_{n-1}}(\xi_{n-1})) - \varphi_t(y)\| < \delta.$$

We have shown that  $\varphi_t(y) \sim_{AP} y$ , and so  $\varphi_t(y) \in [y]_{AP} = [x]_{AP}$ . Since  $y \in [x]_{AP}$  is arbitrary,  $\varphi_t([x]_{AP}) \subset [x]_{AP}$ . This proves the first part of the lemma.

For the second part, suppose now that  $[x]_{AP} \subset \Delta^o$ . In order to see that  $[x]_{AP} \subset \varphi_t([x]_{AP})$  for each  $t \geq 0$ , let  $y \in [x]_{AP}$  and fix  $t > 0$ . We need to show that there is some  $z \in [x]_{AP}$  such that  $\varphi_t(z) = y$ . Fix a sequence  $(\delta_k, T_k)$  such that  $\delta_k \downarrow 0$  and  $T_k \uparrow \infty$ . Since  $y \in \mathcal{R}_{AP}$ , we can find a sequence of  $(\delta_k, T_k)$  AP-pseudo-orbits with corresponding time instants  $\{T_i^k\}_{i=0}^{n(k)-1}$  from  $y$  to  $y$ , which we denote by  $\xi^k = (\xi_0^k, \dots, \xi_{n(k)}^k)$ . We assume without loss of generality that  $T^k > t$  for all  $k$  and let  $\tilde{T}^k \doteq T_{n(k)-1}^k - t$ . From Lemma 1 there is a compact  $\tilde{K}$  in  $\Delta$  such that for all sufficiently large  $k$ ,  $\xi_i^k \in \tilde{K}$  for all  $i \in \{0, \dots, n(k)\}$ . From Lemma 4 we then have that, for all such  $k$ ,  $\varphi_{\tilde{T}^k}(\xi_{n(k)-1}^k)$  lies in some compact set  $\tilde{K}'$ . Thus (passing to a subsequence) we may assume that  $\varphi_{\tilde{T}^k}(\xi_{n(k)-1}^k) \rightarrow z \in \tilde{K}'$ . Since

$$\varphi_t \left( \varphi_{\tilde{T}^k} \left( \xi_{n(k)-1}^k \right) \right) = \varphi_{T_{n(k)-1}^k} \left( \xi_{n(k)-1}^k \right) \rightarrow y,$$

the continuity of  $\varphi_t$  ensures that  $\varphi_t(z) = y$ . Now we show that  $z \in [x]_{AP}$ . Fix  $\delta, T > 0$ , and let  $k$  be large enough so that  $\delta_k < \delta$ ,  $\|\varphi_{\tilde{T}^k}(\xi_{n(k)-1}^k) - z\| < \delta_k$ ,  $T^k > T$ , and  $\tilde{T}^k > T$ . Then  $(\xi_0^k, \dots, \xi_{n(k)-1}^k, z)$  is a  $(\delta, T)$  AP-pseudo-orbit from  $y$  to  $z$  with corresponding time instants  $(T_1^k, \dots, T_{n(k)-2}^k, \tilde{T}^k)$ , so  $y <_{AP} z$ . Now, fix  $\tilde{t} > \max\{t, T\}$ , and note that

$$\varphi_{\tilde{t}}(z) = \varphi_{\tilde{t}-t}(\varphi_t(z)) = \varphi_{\tilde{t}-t}(y).$$

Since  $y \in [x]_{AP}$ , it follows from the positive  $\varphi_t$ -invariance of  $[x]_{AP}$  that  $\varphi_{\tilde{t}-t}(y) \in [x]_{AP}$ , so there is a  $(\delta, T)$  AP-pseudo-orbit from  $\varphi_{\tilde{t}-t}(y)$  to  $y$ , which we denote by  $(\xi_0, \dots, \xi_n)$ . Denote the corresponding time instants by  $T_1, T_2, \dots, T_{n-1}$ . Then  $\tilde{\xi} \doteq (z, z, \xi_1, \dots, \xi_n)$  is a  $(\delta, T)$  AP-pseudo-orbit from  $z$  to  $y$  with time instants  $(\tilde{t}, T_1, \dots, T_{n-1})$ , so  $z <_{AP} y$  and  $z \in [x]_{AP}$ .  $\square$

We now recall the definition of an attractor for the flow  $\{\varphi_t\}$ .

**Definition 6.** A compact set  $A$  is an attractor for the flow  $\{\varphi_t\}$  if  $\varphi_t(A) = A$  for each  $t \geq 0$  and there is some neighborhood  $U$  of  $A$  such that

$$\lim_{t \rightarrow \infty} \sup_{x \in U} \text{dist}(\varphi_t(x), A) = 0.$$

The neighborhood  $U$  is referred to as a fundamental neighborhood for the attractor  $A$ .

The proof of Corollary 1 follows from the proof of [12, Proposition 4.2].

**Corollary 1.** *If  $[x]_{AP} \subset \Delta^o$  is an AP-quasiattractor, then  $[x]_{AP}$  is an attractor.*

*Proof.* Recall that  $\mathcal{R}_{AP}^*$  denotes the collection of all AP-chain recurrent points in  $\Delta^o$ . Note that, from Assumption 2(a) and Lemma 2, for each  $z \in \mathcal{R}_{AP}^*$ ,  $[z]_{AP}$  is a compact set. Choose  $\delta > 0$  such that  $\mathcal{N}^\delta([x]_{AP})$  is an isolating neighborhood of  $[x]_{AP}$  with closure contained in  $\Delta^o$ . Then, from Lemma 4 and Assumption 2(b), there is a compact  $K_0 \subset \Delta^o$  such that for all  $z \in \mathcal{N}^\delta([x]_{AP})$ ,  $\varphi_t(z) \in K_0$  for all  $t \geq 0$ . Let

$$\varepsilon^* \doteq \inf_{y \in \mathcal{R}_{AP}^* \setminus [x]_{AP}} \text{dist}([y]_{AP}, [x]_{AP}).$$

For  $\varepsilon \leq \varepsilon^*$ , let

$$K^\varepsilon \doteq K_0 \setminus \left( \bigcup_{y \in \mathcal{R}_{AP}^* \setminus [x]_{AP}} \mathcal{N}^\varepsilon([y]_{AP}) \right).$$

We claim that there exist  $\varepsilon \leq \varepsilon^*$  and  $\delta_0 \leq \delta$  such that

$$\text{for all } z \in \mathcal{N}^{\delta_0}([x]_{\text{AP}}), \quad \varphi_t(z) \in K^\varepsilon \text{ for all } t \geq 0. \tag{9}$$

We argue via contradiction. Suppose the claim is false; then, since there are finitely many AP–basic classes in  $\mathcal{R}_{\text{AP}}^*$ , there exist  $\delta_n \downarrow 0$ ,  $\varepsilon_n \downarrow 0$ ,  $z_n \in \mathcal{N}^{\delta_n}([x]_{\text{AP}})$ ,  $t_n \geq 0$ ,  $y \in \mathcal{R}_{\text{AP}}^* \setminus [x]_{\text{AP}}$ , such that  $\varphi_{t_n}(z_n) \in \mathcal{N}^{\varepsilon_n}([y]_{\text{AP}})$ . Passing to a subsequence we may assume that  $z_n \rightarrow z$  and  $\varphi_{t_n}(z_n) \rightarrow u$ . Then  $z \in [x]_{\text{AP}}$  and  $u \in [y]_{\text{AP}}$ . We consider two cases: (I) along a further subsequence  $t_n$  converges to some  $t^* < \infty$ ; (II)  $t_n \rightarrow \infty$ . In Case I,  $u = \varphi_{t^*}(z)$  and so by Lemma 8  $u \in [x]_{\text{AP}}$ . But this is a contradiction since  $y \notin [x]_{\text{AP}}$ . In Case II, for every  $\delta, T > 0$ , there is a  $(\delta, T)$  AP–pseudo-orbit from  $z$  to  $u$ , which means that  $z <_{\text{AP}} u$ . Since  $[x]_{\text{AP}}$  is a quasiattractor, from Lemma 7,  $u \in [x]_{\text{AP}}$ , which is once more a contradiction to the fact that  $y \notin [x]_{\text{AP}}$ . Thus we have the claim. Now fix  $\delta_0 \leq \delta^*$  and  $\varepsilon \leq \varepsilon^*$  so that (9) holds.

We now argue that

$$\text{for some } \delta_1 \in (0, \delta_0), \text{ whenever } y \in \mathcal{N}^{\delta_1}([x]_{\text{AP}}), \text{ we have } \varphi_t(y) \in \mathcal{N}^{\delta_0}([x]_{\text{AP}}) \text{ for all } t \geq 0. \tag{10}$$

Once more we proceed via contradiction. Suppose the statement is false. Then there exist  $\delta_n \downarrow 0$ ,  $y_n \in \mathcal{N}^{\delta_n}([x]_{\text{AP}})$ ,  $t_n \geq 0$  such that  $\varphi_{t_n}(y_n) \in (\mathcal{N}^{\delta_0}([x]_{\text{AP}}))^c$ . We can find a subsequence along which  $y_n \rightarrow y$  and  $\varphi_{t_n}(y_n) \rightarrow u$ . We must have  $y \in [x]_{\text{AP}}$  and  $u \in (\mathcal{N}^{\delta_0}([x]_{\text{AP}}))^c$ . Once again we consider two cases as above. In Case I,  $u = \varphi_{t^*}(y) \in [x]_{\text{AP}}$ , which contradicts the fact that  $u \in (\mathcal{N}^{\delta_0}([x]_{\text{AP}}))^c$ . In Case II,  $y <_{\text{AP}} u$ , and so as before,  $u \in [x]_{\text{AP}}$ . Once more this is a contradiction. Thus we have shown (10). Now fix  $\delta_1 \in (0, \delta_0)$  such that (10) holds. Let  $U_0 \doteq \mathcal{N}^{\delta_0}([x]_{\text{AP}})$  and  $U_1 \doteq \mathcal{N}^{\delta_1}([x]_{\text{AP}})$ .

We will now show that

$$\lim_{t \rightarrow \infty} \sup_{y \in U_1} \text{dist}(\varphi_t(y), [x]_{\text{AP}}) = 0. \tag{11}$$

Together with Lemma 8 we will then have that  $[x]_{\text{AP}}$  is an attractor, completing the proof of the result. In order to show (11) we will show that for each open neighborhood  $O$  of  $[x]_{\text{AP}}$ ,  $O \subset U_1$ , there is some  $t(O) < \infty$  such that  $\varphi_t(U_1) \subset O$  for all  $t \geq t(O)$ . For any such  $O$ , let  $O_1 \subset\subset O$  be an open neighborhood of  $[x]_{\text{AP}}$  such that for all  $y \in O_1$ ,  $\varphi_t(y) \in O$  for all  $t \geq 0$ . Here, for open sets  $G_1, G_2$ , we write  $G_1 \subset\subset G_2$  if  $\bar{G}_1 \subset G_2$ . Existence of such an  $O_1$  is shown in a similar manner as (10). It suffices to show that

$$t(O) \doteq \inf\{t: \varphi_t(U_0) \subset O\} < \infty,$$

since then for each  $t \geq t(O)$ ,

$$\varphi_t(U_1) = \varphi_{t(O)}(\varphi_{t-t(O)}(U_1)) \subset \varphi_{t(O)}(U_0) \subset O,$$

which will complete the proof.

In order to see that  $t(O) < \infty$  for each such  $O$ , we argue by contradiction. Suppose that there is some  $O$  (with the associated  $O_1$ ) such that  $t(O) = \infty$ . Then we can find sequences  $\{z_n\} \subset U_0$  and  $T_n \uparrow \infty$  such that  $\varphi_{T_n}(z_n) \in O^c$ . From the definition of  $O_1$ , this says that  $\varphi_t(z_n) \in O_1^c$  for all  $0 \leq t \leq T_n$ . Suppose that  $z_n \rightarrow z$  along a subsequence. Then  $\varphi_t(z) \in O_1^c$  for all  $t > 0$ . Also, since  $z \in \mathcal{N}^{\delta_0}([x]_{\text{AP}})$ , by (9),  $\varphi_t(z) \in K^\varepsilon$  for all  $t \geq 0$ . Thus we have  $\omega(z) \subset K^\varepsilon \setminus O_1$ . The final statement of Lemma 1 implies that for each  $x \in \Delta^o$ ,  $\omega(x) \neq \emptyset$  and therefore  $\omega(z)$  is a nonempty subset of  $\mathcal{R}_{\text{AP}}^*$ . Thus we have that  $(K^\varepsilon \setminus O_1) \cap \mathcal{R}_{\text{AP}}^*$  is nonempty, which contradicts the definition of  $K^\varepsilon$  and  $O_1$ . Thus we have that  $t(O) < \infty$ , and the result follows.  $\square$

The following lemma shows that suitable AP–pseudo-orbits come arbitrarily close to AP–recurrence classes.

**Lemma 9.**

- (a) For each  $\delta > 0$  and compact  $A \in \Delta$ , there is a  $\delta_0 \in (0, 1]$  and  $T_A \in (0, \infty)$  such that any  $(\delta_0, T_A)$  AP–pseudo-orbit that starts in  $A$  intersects  $N^\delta(\mathcal{R}_{AP})$ .
- (b) For each  $\delta > 0$  and  $A \in \mathcal{K}$ , there is a  $T_A^* \in (0, \infty)$  such that for every  $x \in A$ , there is a  $t_0 \in [0, T_A^*]$  with  $\varphi_{t_0}(x) \in N^\delta(\mathcal{R}_{AP}^*)$ .

*Proof.* Consider first Part (a). Fix  $\delta > 0$  and a compact  $A \in \Delta$ , and let  $T = 1$ . With this choice of  $A$  and  $T$ , let  $\delta_0$  and  $A_1$  be as given in Lemma 1. For  $x \in \Delta$ , let  $T^\delta(x) \doteq \inf\{t \geq 0: \varphi_t(x) \in N^\delta(\mathcal{R}_{AP})\}$ . Since  $\omega(x)$  is a nonempty subset of  $\mathcal{R}_{AP}$ ,  $T^\delta(x) < \infty$  for each  $x \in \Delta$ . We now claim that  $T^\delta$  is an upper semicontinuous function on  $\Delta$ . For this it suffices to argue that for each  $\alpha > 0$ , the level set  $L_\alpha \doteq \{x \in \Delta: T^\delta(x) \geq \alpha\}$  is closed. Let  $\{x_n\} \subset L_\alpha$  be a sequence converging to some  $x \in \Delta$ , and note that for each  $t \geq 0$ ,  $\lim_{n \rightarrow \infty} \varphi_t(x_n) = \varphi_t(x)$ . For  $t < \alpha$ ,  $\varphi_t(x_n) \in (N^\delta(\mathcal{R}_{AP}))^c$ , which is closed, so  $\varphi_t(x) \in (N^\delta(\mathcal{R}_{AP}))^c$ . Since this holds for all  $t < \alpha$ , we have that  $x \in L_\alpha$ . This shows that the level sets of  $T^\delta$  are closed and thus establishes the claim. Since an upper semicontinuous function achieves its supremum over any compact set,  $T_1 = \sup_{x \in A_1} T^\delta(x) < \infty$ . Let  $T_A \doteq T_1 \vee 1$ . Then, from Lemma 1, any  $(\delta_0, T_A)$  AP–pseudo-orbit given by a collection of points  $\xi = (\xi_0 = x, \dots, \xi_n = y) \in \Delta^{n+1}$  and a collection of times  $T_A \leq T_1, \dots, T_{n-1}$ , with  $x \in A$ , must satisfy  $\xi_i \in A_1$  for every  $i \in \{0, \dots, n\}$ . Also, by the definition of  $T_A$ , we must have that for each  $i \in \{1, \dots, n-1\}$ , there is a  $t \in [0, T_i]$  such that  $\varphi_t(\xi_i) \in N^\delta(\mathcal{R}_{AP})$ . The result in Part (a) follows.

The proof of Part (b) can be completed in a similar manner on observing that from Lemma 3, for every  $x \in A$ , the forward orbit  $\gamma^+(x)$  is contained in a compact subset of  $\Delta^o$ . We omit the details. □

The following lemma gives key properties of pseudo-orbits in relation to their visits to neighborhoods of AP–quasiattractors and non-quasiattractors.

**Lemma 10.**

- (a) For every  $\theta > 0$ , there are  $\delta = \delta(\theta) < \theta$  and  $T = T(\theta) > 0$  with the property that if there is a  $(\delta, T)$  AP–pseudo-orbit  $\xi \doteq (\xi_0, \dots, \xi_n)$  with

$$\xi_0 \in N^\delta(K_i), \xi_n \in N^\delta(K_{i'}), \text{ and } \xi_j \in (N^\theta(K_i))^c \text{ for some } j \in \{1, \dots, n-1\}, \quad (12)$$

then we must have  $i \neq i'$ .

- (b) There exist  $\delta, T > 0$  such that if for some  $i, i' \in \{1, \dots, v\}$  there is a  $(\delta, T)$  AP–pseudo-orbit  $\xi \doteq (\xi_0, \dots, \xi_n)$  such that  $\xi_0 \in N^\delta(K_i)$  and  $\xi_n \in N^\delta(K_{i'})$ , then we must have that  $K_i \leq_{AP} K_{i'}$ .

*Proof.* For the first statement in the lemma we will argue via contradiction. By Lemma 5 we can choose  $\bar{\delta} > 0$ ,  $\bar{T} > 0$ , and  $\bar{K} \in \mathcal{K}$  such that any  $(\bar{\delta}, \bar{T})$  AP–pseudo-orbit starting from  $N^{\bar{\delta}}(K_i)$  lies in  $\bar{K}$  for every  $i = 1, \dots, v$ . Henceforth we only consider  $(\delta, T)$  AP–pseudo-orbits with  $\delta < \bar{\delta}$  and  $T > \bar{T}$ . Fix  $\theta > 0$  and suppose that there is a sequence  $\theta > \delta_k \downarrow 0$  and  $T_k \uparrow \infty$ , such that for every  $k$  there is a  $(\delta_k, T_k)$  AP–pseudo-orbit  $\xi^k \doteq (\xi_0^k, \dots, \xi_{n(k)}^k)$  that satisfies (12) (with

$\xi, \delta, n$  replaced by  $(\xi_k, \delta_k, n(k))$ , with  $i = i'$ . Let  $j(k) \in \{1, \dots, n(k) - 1\}$  be such that  $\xi_{j(k)}^k \in (N^\theta(K_i))^c$ . By passing to a subsequence if necessary, we can find  $x, y \in K_i$  and  $z \in (N^\theta(K_i))^c \cap \bar{K}$  such that  $\xi_0^k \rightarrow x, \xi_{n(k)}^k \rightarrow y$ , and  $\xi_{j(k)}^k \rightarrow z$ .

In order to see that  $x \leq_{AP} z$ , fix  $\delta, T > 0$  and let  $k$  be large enough so that

$$\delta_k < \frac{\delta}{2}, \quad T_k > T, \quad \|x - \xi_0^k\| < \frac{\delta}{2}, \quad \text{and} \quad \|\xi_{j(k)}^k - z\| < \frac{\delta}{2}.$$

Then  $\|x - \xi_1^k\| \leq \|x - \xi_0^k\| + \|\xi_0^k - \xi_1^k\| < \delta$ , and

$$\left\| \varphi_{T_{j(k)-1}^k} \left( \xi_{j(k)-1}^k \right) - z \right\| \leq \left\| \varphi_{T_{j(k)-1}^k} \left( \xi_{j(k)-1}^k \right) - \xi_{j(k)}^k \right\| + \left\| \xi_{j(k)}^k - z \right\| < \delta,$$

and so  $\tilde{\xi} \doteq (x, \xi_1^k, \dots, \xi_{j(k)-1}^k, z)$  is a  $(\delta, T)$  AP-pseudo-orbit from  $x$  to  $z$ . Thus  $x <_{AP} z$ . Similarly,  $z <_{AP} y$ , which shows that  $z \in K_i$ . However, since  $z \in (N^\theta(K_i))^c$ , this is a contradiction. This proves (a).

Now consider Part (b). Fix  $i, i' \in \{1, \dots, v\}$  and suppose that for each  $\delta, T > 0$  there is some  $(\delta, T)$  AP-pseudo-orbit  $\xi \doteq (\xi_0, \dots, \xi_n)$  such that  $\xi_0 \in N^\delta(K_i)$  and  $\xi_n \in N^\delta(K_{i'})$ . Let  $\delta_k \downarrow 0$  and  $T_k \uparrow \infty$  and let  $\xi^k \doteq (\xi_0^k, \dots, \xi_{n(k)}^k)$  be a  $(\delta_k, T_k)$  AP-pseudo-orbit such that  $\xi_0^k \in N^{\delta_k}(K_i)$  and  $\xi_{n(k)}^k \in N^{\delta_k}(K_{i'})$ . Passing to subsequences if necessary, we can find  $x \in K_i$  and  $y \in K_{i'}$  such that  $\xi_0^k \rightarrow x$  and  $\xi_{n(k)}^k \rightarrow y$ . Thus, for any fixed  $\delta, T > 0$ , when  $k$  is sufficiently large,  $\tilde{\xi} \doteq (x, \xi_1^k, \dots, \xi_{n(k)-1}^k, y)$  is a  $(\delta, T)$  AP-pseudo-orbit from  $K_i$  to  $K_{i'}$ , showing that  $K_i \leq_{AP} K_{i'}$ . So if for some  $i, i', K_i \leq_{AP} K_{i'}$  does not hold, there must exist  $\bar{\delta} = \delta(i, i') > 0$  and  $\bar{T} = T(i, i') < \infty$  such that there is no  $(\bar{\delta}, \bar{T})$  AP-pseudo-orbit  $\xi \doteq (\xi_0, \dots, \xi_n)$  with the property that  $\xi_0 \in N^{\bar{\delta}}(K_i)$  and  $\xi_n \in N^{\bar{\delta}}(K_{i'})$ . Define  $\delta = \min_{(i,i')} \delta(i, i')$  and  $T = \max_{(i,i')} T(i, i')$ . Clearly, the statement in Part (b) holds with this choice of  $(\delta, T)$ . □

The final result of this section is a consequence of Lemma 9 and Lemma 10. It summarizes key properties of AP-pseudo-orbits in relation to AP-recurrent classes. This result will be used in Section 7 in the proof of Theorem 1.

**Lemma 11.** *For each  $\delta, T > 0$  and compact set  $A \subset \Delta^\circ$ , there is a collection of open neighborhoods  $\{V_i\}_{i=1}^v$  of  $\{K_i\}_{i=1}^v$ , with  $\bar{V}_i \subset N^\delta(K_i) \cap \Delta^\circ$ , along with  $\delta_0 \in (0, \delta)$ ,  $T_0 \in (T, \infty)$ , and  $n \in \mathbb{N}$ , such that the following hold:*

1.  $\overline{N^{\delta_0}(K_i)} \subset V_i$  for each  $i \in \{1, \dots, v\}$ .
2. For each  $i \in \{1, \dots, l\}$ , if  $\xi \doteq (\xi_0, \dots, \xi_n)$  is a  $(\delta_0, T_0)$  AP-pseudo-orbit with  $\xi_0 \in V_i$ , then  $\xi_j \in V_i$  for all  $j \in \{1, \dots, n\}$ .
3. If  $\xi \doteq (\xi_0, \dots, \xi_n)$  is a  $(\delta_0, T_0)$  AP-pseudo-orbit with corresponding time instants  $(T_1, \dots, T_{n-1})$  such that  $\xi_0 \in N^{\delta_0}(K_i)$  and  $\xi_n \in N^{\delta_0}(K_j)$  for some  $i, j \in \{1, \dots, v\}$ , and there is  $m \in \{1, \dots, n - 1\}$  such that  $\xi_m \in V_i^c$ , then  $i \neq j$  and  $K_i \leq_{AP} K_j$ .
4. If  $\xi \doteq (\xi_0, \dots, \xi_n)$  is a  $(\delta_0, T_0)$  AP-pseudo-orbit with  $\xi_0 \in A$ , then there is some  $k \in \{1, \dots, n - 1\}$  and  $t \in [0, T_k]$  such that  $\varphi_t(\xi_k) \in N^\delta(\mathcal{R}_{AP} \cap \Delta^\circ)$ .

*Proof.* Fix  $\delta, T \in (0, \infty)$  and a compact  $A \in \Delta^\circ$ . Since  $K_i$  is an attractor for each  $i \in \{1, \dots, l\}$ , there is a bounded open neighborhood  $O_i$  of  $K_i$ , with  $\bar{O}_i \subset N^\delta(K_i) \cap \Delta^\circ$ , such that

$$\lim_{t \rightarrow \infty} \sup_{x \in O_i} \text{dist}(\varphi_t(x), K_i) = 0. \tag{13}$$

For each  $i \in \{l + 1, \dots, v\}$ , let  $O_i$  be an arbitrary bounded, open, and isolating neighborhood of  $K_i$  such that  $\bar{O}_i \subset N^\delta(K_i) \cap \Delta^o$ . Denote the  $(\delta, T)$  given by Lemma 10(b) by  $(\delta_1^*, T_1^*)$ , and denote the  $(\delta_0, T_A)$  given by Lemma 9(a) by  $(\delta_2^*, T_2^*)$ . Let  $\theta > 0$  be small enough so that  $N^\theta(K_i) \subset O_i$  for each  $i \in \{1, \dots, v\}$ . From Lemma 10 we can find  $\delta_1 < \min\{\theta, \delta_1^*, \delta_2^*\}$  and  $T_1 > \max\{T, T_1^*, T_2^*\}$  such that if  $\xi \doteq (\xi_0, \dots, \xi_n)$  is a  $(\delta_1, T_1)$  AP-pseudo-orbit with  $\xi_0 \in N^{\delta_1}(K_i)$  and  $\xi_n \in N^{\delta_1}(K_j)$  such that  $\xi_m \in (N^\theta(K_i))^c$  for some  $m \in \{1, \dots, n - 1\}$ , then  $i \neq j$  and  $K_i \leq_{AP} K_j$ .

Now, let  $V_i \doteq N^{\theta+\varepsilon}(K_i)$ , where  $\varepsilon > 0$  is small enough so that  $\bar{V}_i \subset O_i$  for all  $i \in \{1, \dots, v\}$ , and let  $\delta_2 < \delta_1$  be small enough so that  $\bar{N}^{\delta_2}(V_i) \subset O_i$ . Thus, for every  $i \in \{1, \dots, v\}$

$$K_i \subset N^{\delta_2}(K_i) \subset N^\theta(K_i) \subset V_i \subset N^{\delta_2}(V_i) \subset O_i,$$

where, as before, for open sets  $G_1, G_2$ , we write  $G_1 \subset\subset G_2$  if  $\bar{G}_1 \subset G_2$ .

From (13), there is some  $T_2 > T_1$  such that if  $t \geq T_2$ , then for each  $i \in \{1, \dots, l\}$ ,

$$\sup_{u \in O_i} \text{dist}(\varphi_t(u), K_i) < \delta_2.$$

Then Parts 1 and 2 hold when  $\delta_0 \doteq \delta_2$  and  $T_0 \doteq T_2$ . Additionally, Part 3 holds from the property of  $(\delta_1, T_1)$  AP-pseudo-orbits noted above, since  $V_i^c \subset (N^\theta(K_i))^c$  for each  $i \in \{1, \dots, v\}$ . Finally, since  $T_0 \geq T_2^*$  and  $\delta_0 \leq \delta_2^*$ , from Lemma 9, Part 4 holds as well.  $\square$

#### 4. Large deviation estimates

Throughout this section we will assume that Assumption 3 is satisfied. We will give some key uniform large deviation bounds that will be used in Sections 5, 6, and 7.

For  $\alpha \in (0, 1)$  let  $\mathcal{V}_\alpha \doteq \{x \in \Delta^o : \text{dist}(x, \partial\Delta) > \alpha\}$ . For each compact  $K \in \mathcal{K}$ , let  $\mathcal{V}_{\alpha,K} \doteq \mathcal{V}_\alpha \cap K$ , and let  $\pi_{\alpha,K}$  denote the projection map from  $\mathbb{R}^d$  to  $\bar{\mathcal{V}}_{\alpha,K}$ , defined as

$$\pi_{\alpha,K}(x) \doteq \arg \min_y \{\|y - x\| : y \in \bar{\mathcal{V}}_{\alpha,K}\}.$$

Similarly, denote by  $\pi_{\alpha,K}^N$  the projection map from  $\mathbb{R}^d$  to  $\bar{\mathcal{V}}_{\alpha,K} \cap \Delta_N$ . Let  $\theta^{N,\alpha,K}$  be a transition probability kernel on  $\mathbb{R}^d$  defined by

$$\theta^{N,\alpha,K}(\cdot | x) \doteq \theta^N(\cdot | \pi_{\alpha,K}^N(x)).$$

Let  $\{X_n^{N,\alpha,K}\}$  be an  $\mathbb{R}^d$ -valued chain defined as in (1) but with  $\theta^N$  replaced with  $\theta^{N,\alpha,K}$ . We consider continuous-time processes  $\hat{X}^{N,\alpha,K}$  associated with  $\{X_n^{N,\alpha,K}\}$  as

$$\hat{X}^{N,\alpha,K}(t) = X_n^{N,\alpha,K} + [X_{n+1}^{N,\alpha,K} - X_n^{N,\alpha,K}](Nt - n), \quad t \in [n/N, (n + 1)/N], \quad n \in \mathbb{N}_0.$$

We now present a basic large deviation result for  $\hat{X}^{N,\alpha,K}$ . Recall the stochastic kernel  $\theta(dy|x)$  from Assumption 3(b). For  $x, \zeta \in \mathbb{R}^d$ , define

$$H_{\alpha,K}(x, \zeta) \doteq \log \int_{\mathbb{R}^d} \exp\{\langle \zeta, y \rangle\} \theta(dy | \pi_{\alpha,K}(x)),$$

and let

$$L_{\alpha,K}(x, \beta) \doteq \sup_{\zeta \in \mathbb{R}^d} \{\langle \zeta, \beta \rangle - H_{\alpha,K}(x, \zeta)\}.$$



We note that for every  $\beta, \zeta \in \mathbb{R}^d$ ,  $H_{\alpha,K}(x, \zeta) = H_{\alpha',K'}(x, \zeta)$  and  $L_{\alpha,K}(x, \beta) = L_{\alpha',K'}(x, \beta)$  whenever  $\pi_{\alpha,K}(x) = x = \pi_{\alpha',K'}(x)$ . For  $x \in \Delta^o$  and  $\beta, \zeta \in \mathbb{R}^d$ , define

$$L(x, \beta) \doteq L_{\alpha,K}(x, \beta), \quad H(x, \zeta) \doteq H_{\alpha,K}(x, \zeta) \text{ if } x \in \mathcal{V}_{\alpha,K}.$$

For  $\alpha > 0, x \in \mathbb{R}^d, K \in \mathcal{K}, T \in (0, \infty)$ , and  $\phi \in C([0, T]; \mathbb{R}^d)$ , define

$$S_{\alpha,K}(x, T, \phi) \doteq \begin{cases} \int_0^T L_{\alpha,K}(\phi(t), \dot{\phi}(t)) dt & \text{if } \phi \text{ is absolutely continuous,} \\ \infty & \text{otherwise.} \end{cases}$$

Note that if, for  $\alpha, \alpha' > 0$  and  $K, K' \in \mathcal{K}$ ,  $\phi \in C([0, T]; \bar{\mathcal{V}}_{\alpha,K}) \cap C([0, T]; \bar{\mathcal{V}}_{\alpha',K'})$ , then  $S_{\alpha,K}(\phi(0), T, \phi) = S_{\alpha',K'}(\phi(0), T, \phi)$ . Thus, for  $\phi \in C([0, T]; \mathbb{R}^d)$  that satisfies  $\phi(0) = x$  and  $\phi(t) \in \Delta^o$  for all  $t \in [0, T]$ , we define

$$S(x, T, \phi) = S_{\alpha,K}(x, T, \phi) \quad \text{if } \phi \in C([0, T]; \bar{\mathcal{V}}_{\alpha,K}) \text{ for some } \alpha > 0 \text{ and } K \in \mathcal{K}. \tag{14}$$

The following uniform large deviation principle will be used several times in this work.

**Theorem 3.** *Suppose Assumption 3 is satisfied. Fix  $T \in (0, \infty), \alpha > 0$ , and  $K, K' \in \mathcal{K}$ . For each  $a \in (0, \infty)$ , let*

$$\Phi_{x,\alpha,K',T}(a) \doteq \left\{ \phi \in C([0, T]; \mathbb{R}^d) : S_{\alpha,K'}(x, T, \phi) \leq a \right\}.$$

(a) (Compact level sets.) *For every  $a \in (0, \infty)$ , the set  $\bigcup_{x \in K} \Phi_{x,\alpha,K',T}(a)$  is compact.*

(b) (Upper bound.) *Given  $\delta, \gamma \in (0, 1)$  and  $L \in (0, \infty)$ , there is some  $N < \infty$  such that*

$$\mathbb{P}_x \left( \left\| \hat{X}^{n,\alpha,K'} - \phi \right\|_{*,T} < \delta \right) \geq \exp \left( -n(S_{\alpha,K'}(x, T, \phi) + \gamma) \right)$$

*for all  $n \geq N, x \in K \cap \Delta_N$ , and  $\phi \in \Phi_{x,\alpha,K',T}(L)$ .*

(c) (Lower bound.) *Given  $\delta, \gamma \in (0, 1)$  and  $L \in (0, \infty)$ , there is some  $N < \infty$  such that*

$$\mathbb{P}_x \left( d \left( \hat{X}^{n,\alpha,K'}, \Phi_{x,\alpha,K',T}(l) \right) \geq \delta \right) \leq \exp \left( -n(l - \gamma) \right)$$

*for all  $n \geq N, x \in K \cap \Delta_N$ , and  $l \in [0, L]$ .*

*Proof.* We will apply [3, Theorem 6.7.5]. For  $x, \zeta \in \mathbb{R}^d$ , let

$$H_{\alpha,K}^N(x, \zeta) \doteq \log \int_{\mathbb{R}^d} \exp\{\langle \zeta, y \rangle\} \theta^N(dy | \pi_{\alpha,K}^N(x)).$$

By Assumption 3(b)(iii), for each compact  $A \subset \mathbb{R}^d$  and  $\zeta \in \mathbb{R}^d$ ,

$$\sup_{N \in \mathbb{N}} \sup_{x \in \mathbb{R}^d} H_{\alpha,K}^N(x, \zeta) < \infty, \quad \sup_{x \in \mathbb{R}^d} H_{\alpha,K}(x, \zeta) < \infty, \tag{15}$$

and

$$\sup_{x \in A} \left| H_{\alpha,K}^N(x, \zeta) - H_{\alpha,K}(x, \zeta) \right| \rightarrow 0 \quad \text{as } N \rightarrow \infty. \tag{16}$$

Furthermore, from Assumption 3(b)(ii),  $x \mapsto \theta(dy | \pi_{\alpha,K}(x))$  is a continuous map from  $\mathbb{R}^d$  to  $\mathcal{P}(\mathbb{R}^d)$ . Thus, Conditions 6.2.1 and 6.7.2 of [3] are satisfied. Next, since from Assumption

3(b)(i) the convex hull of the support of  $\theta(dy|\pi_{\alpha,K}(x))$  is all of  $\mathbb{R}^d$ , [3, Condition 6.7.4] is satisfied as well. Thus, from [3, Theorem 6.7.5] we have that, for every  $T \in (0, \infty)$ ,  $\{\hat{X}^{N,\alpha,K}\}_{N \in \mathbb{N}}$  satisfies a Laplace principle, uniformly on compact subsets of  $\mathbb{R}^d$ , in the sense of [3, Definition 1.2.6], with rate function  $S_{\alpha,K}(x, T, \cdot)$ . It is shown in [17, Theorem 4.3] that a uniform Laplace principle of the form given in [3, Theorem 6.7.5] implies a uniform large deviation principle in the sense of Freidlin and Wentzell [6], which means that Parts (a)–(c) of the theorem hold. The result follows.  $\square$

**Lemma 12.** For every  $\alpha \in (0, 1)$  and a compact  $K \in \Delta^o$ ,  $(x, \beta) \mapsto L_{\alpha,K}(x, \beta)$  is a continuous map on  $\mathbb{R}^d \times \mathbb{R}^d$ .

*Proof.* The proof follows from [3, Lemma 6.5.2] on noting that, by Assumption 3(b) for every  $x \in \mathbb{R}^d$ , the convex hull of the support of  $\theta(dy|\pi_{\alpha,K}(x))$  is  $\mathbb{R}^d$ , and  $\sup_{x \in \mathbb{R}^d} H_{\alpha,K}(x, \zeta) < \infty$  for every  $\zeta \in \mathbb{R}^d$ .  $\square$

An important consequence of the above uniform large deviation principle is the following uniform upper bound for closed sets  $F$  in  $C([0, T]; \mathbb{R}^d)$ .

**Theorem 4.** Fix  $T \in (0, \infty)$ ,  $\alpha > 0$ , and  $K, K' \in \mathcal{K}$ . Then, for every closed set  $F$  in  $C([0, T]; \mathbb{R}^d)$ ,

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \sup_{x \in K \cap \Delta_N} \mathbb{P}_x(\hat{X}^{N,\alpha,K'} \in F) \leq - \inf_{x \in K} \inf_{\phi \in F} S_{\alpha,K'}(x, T, \phi).$$

*Proof.* Fix  $T, \alpha, K, K'$  as in the statement of the theorem. We begin by showing that for each  $s \geq 0$  and  $\delta > 0$  there is some  $\varepsilon \doteq \varepsilon(\delta) \in (0, 1)$  such that for all  $x, y \in K$  with  $\|x - y\| \leq \varepsilon$ ,

$$\left\{ \phi \in C([0, T]; \mathbb{R}^d) : d\left(\phi, \Phi_{x,\alpha,K',T}(s)\right) \leq \delta \right\} \supseteq \left\{ \phi \in C([0, T]; \mathbb{R}^d) : d\left(\phi, \Phi_{y,\alpha,K',T}\left(s - \frac{\delta}{4}\right)\right) \leq \frac{\delta}{2} \right\}. \tag{17}$$

Let  $\kappa_0 \doteq 1 + \sup_{x \in K, \|\beta\| \leq 1} L(x, \beta)$ . From Lemma 12,  $\kappa_0 < \infty$ . Since

$$\cup_{y \in K} \Phi_{y,\alpha,K',T}\left(s - \frac{\delta}{4}\right)$$

is a compact set, we can find  $\varepsilon \in \left(0, \frac{\delta}{8\kappa_0}\right)$  such that for all

$$\psi \in \cup_{y \in K} \Phi_{y,\alpha,K',T}\left(s - \frac{\delta}{4}\right)$$

and  $0 \leq t_1 \leq t_2 \leq T$  with  $|t_1 - t_2| \leq \varepsilon$ , we have  $\|\psi(t_2) - \psi(t_1)\| \leq \frac{\delta}{8}$ .

Fix  $y \in K$  and  $\phi$  in the set on the right side of (17). Then there is a  $\psi_1 \in \Phi_{y,\alpha,K',T}\left(s - \frac{\delta}{4}\right)$  such that

$$\|\phi - \psi_1\|_{*,T} \leq \frac{\delta}{2} + \frac{\delta}{8} = \frac{5\delta}{8}.$$

Note in particular that  $\psi_1(0) = y$ . Fix  $x \in K$  such that  $\|y - x\| \leq \varepsilon$ . Let  $t_0 \doteq \|x - y\|$  and define the function  $\eta_{x,y} : [0, t_0] \rightarrow \mathbb{R}^d$  as

$$\eta_{x,y}(t) \doteq x + \frac{(y - x)}{\|y - x\|} t. \tag{18}$$

Define  $\psi_2: [0, T] \rightarrow \mathbb{R}^d$  as

$$\psi_2(s) \doteq \eta_{x,y}(s)1_{[0,t_0]}(s) + \psi_1(s - t_0)1_{(t_0,T]}(t).$$

Note that  $\psi_2(0) = x$  and

$$\begin{aligned} S_{\alpha,K'}(x, T, \psi_2) &= \int_0^{t_0} L_{\alpha,K'}(\psi_2(t), \dot{\psi}_2(t))dt + \int_{t_0}^T L_{\alpha,K'}(\psi_2(t), \dot{\psi}_2(t))dt \\ &\leq \varepsilon\kappa_0 + s - \frac{\delta}{4} = \frac{\delta}{8} + s - \frac{\delta}{4} \leq s. \end{aligned}$$

Thus  $\psi_2 \in \Phi_{x,\alpha,K',T}(s)$ . Furthermore,

$$\|\phi - \psi_2\|_{*,T} \leq \|\phi - \psi_1\|_{*,T} + \|\psi_1 - \psi_2\|_{*,T} \leq \frac{5\delta}{8} + \|\psi_1 - \psi_2\|_{*,T}.$$

Also, for  $t \in (t_0, T]$ ,

$$\|\psi_1(t) - \psi_2(t)\| = \|\psi_1(t) - \psi_1(t - t_0)\| \leq \frac{\delta}{8},$$

and for  $t \in [0, t_0]$ ,

$$\|\psi_1(t) - \psi_2(t)\| \leq \|\psi_2(t) - y\| + \|\psi_1(t) - \psi_1(0)\| \leq \varepsilon + \frac{\delta}{8} \leq \frac{\delta}{8} + \frac{\delta}{8} = \frac{\delta}{4}.$$

Thus

$$\|\phi - \psi_2\|_{*,T} \leq \frac{5\delta}{8} + \frac{\delta}{4} \leq \delta.$$

Since  $\psi_2 \in \Phi_{x,\alpha,K',T}(s)$ , we have  $d(\phi, \Phi_{x,\alpha,K',T}(s)) \leq \delta$ , and thus  $\phi$  is in the set on the left side of (17). This proves the inclusion in (17).

Now fix a closed set  $F$  in  $C([0, T]:\mathbb{R}^d)$ . If

$$\inf_{x \in K} \inf_{\phi \in F} S_{\alpha,K'}(x, T, \phi) = 0,$$

then the result clearly holds, so we assume that

$$\bar{S} \doteq \inf_{x \in K} \inf_{\phi \in F} S_{\alpha,K'}(x, T, \phi) > 0.$$

Fix  $s \in (0, \bar{S})$ , and let  $\{x_n\} \subset K$  and  $\varepsilon \downarrow 0$ . Since  $K$  is compact, we may pass to a subsequence and assume that  $x_n \rightarrow \tilde{x}$  for some  $\tilde{x} \in K$ . Since

$$\inf_{\phi \in F} S_{\alpha,K'}(\tilde{x}, T, \phi) > s,$$

$F \cap \Phi_{\tilde{x},\alpha,K',T}(s) = \emptyset$ . This, along with the facts that  $\Phi_{\tilde{x},\alpha,K',T}(s)$  is compact and  $F$  is closed, ensures that there is some  $\delta \in (0, s)$  such that

$$F \subset \left\{ \phi \in C([0, T]:\mathbb{R}^d) : d(\phi, \Phi_{\tilde{x},\alpha,K'}(s)) > \delta \right\}.$$

Let  $\varepsilon = \varepsilon(\delta) > 0$  be chosen as above (17). Without loss of generality we assume that  $\|\tilde{x} - x_n\| \leq \varepsilon$  for all  $n$ . Then, for every  $n \in \mathbb{N}$ ,

$$F \subset \left\{ \phi \in C\left([0, T]: \mathbb{R}^d: d\left(\phi, \Phi_{\tilde{x}, \alpha, K', T}(s)\right) > \delta \right\}$$

$$\subset \left\{ \phi \in C\left([0, T]: \mathbb{R}^d: d\left(\phi, \Phi_{x_n, \alpha, K', T}\left(s - \frac{\delta}{4}\right)\right) > \frac{\delta}{2} \right\}.$$

From the upper bound in Theorem 3(c) we see that

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_{x_N} \left( \hat{X}^{N, \alpha, K'} \in F \right) \leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_{x_N} \left( d\left(\hat{X}^{N, \alpha, K'}, \Phi_{x_N, \alpha, K', T}\left(s - \frac{\delta}{4}\right)\right) > \frac{\delta}{2} \right)$$

$$\leq -\left(s - \frac{\delta}{4}\right).$$

The result follows from letting  $\delta \rightarrow 0$  and  $s \rightarrow \bar{S}$ . □

### 5. Asymptotic behavior of QSD

In this section we assume that Assumptions 1, 2, 3, and 4 are satisfied. Using these assumptions we will provide several exponential probability estimates and use them to deduce some asymptotic properties of the QSD  $\{\mu_N\}$  (when they exist). For  $N \in \mathbb{N}$  and  $T \in (0, \infty)$ , let

$$D_T^N \doteq \sup_{0 \leq t \leq T} \left\| \hat{X}^N(t) - \varphi_t(X_0^N) \right\| = \left\| \hat{X}^N - \varphi_\cdot(X_0^N) \right\|_{*, T}.$$

The estimates obtained in Lemma 13 and Lemma 14 are the key steps in the proof of Theorem 5, which gives the asymptotics of  $\lambda_N \doteq \left[ \mathbb{P}_{\mu_N}(X_1^N \in \Delta^o) \right]^N$ , where  $\mu_N$  is a QSD for  $\{X^N\}$ . Recall the definition of  $\mathcal{V}_\alpha$  from Section 4.

**Lemma 13.** *For each  $\alpha > 0$ , compact set  $K \subset \mathcal{V}_\alpha$ ,  $\varepsilon > 0$ , and  $T > 0$ , there is a  $c \in (0, \infty)$  and  $N_0 \in \mathbb{N}$  such that for every  $N \geq N_0$ ,*

$$\sup_{x \in K \cap \Delta_N} \mathbb{P}_x \left[ D_T^N \geq \varepsilon \right] \leq \exp(-Nc).$$

*Proof.* Let  $\alpha > 0$  and let  $K \subset \mathcal{V}_\alpha$  be compact. For each  $\varepsilon \in (0, \alpha)$ , let

$$F_\varepsilon = \left\{ \psi \in C([0, T]: \mathbb{R}^d): \sup_{0 \leq t \leq T} \|\psi(t) - \varphi_t(\psi(0))\| \geq \varepsilon \right\}.$$

Using Lemma 3 we can (and will) assume without loss of generality that  $\varepsilon$  is small enough so that the compact set

$$K' \doteq \overline{(\varphi_{[0, \infty)}(K))} \subset \Delta^o.$$

Note that

$$\sup_{x \in K \cap \Delta_N} \mathbb{P}_x \left[ D_T^N \geq \varepsilon \right] = \sup_{x \in K \cap \Delta_N} \mathbb{P}_x \left[ \sup_{0 \leq t \leq T} \left\| \hat{X}^{N, \alpha, K'}(t) - \varphi_t\left(\hat{X}^{N, \alpha, K'}(0)\right) \right\| \geq \varepsilon \right].$$

Since  $F_\varepsilon$  is closed, Theorem 4 says that for each  $\delta > 0$ , there is an  $N_\delta \in \mathbb{N}$  such that for all  $N \geq N_\delta$ ,

$$\begin{aligned} \log \sup_{x \in K \cap \Delta_N} \mathbb{P}_x \left[ \sup_{0 \leq t \leq T} \|\hat{X}^{N, \alpha, K'}(t) - \varphi_t(x)\| \geq \varepsilon \right] &= \log \sup_{x \in K \cap \Delta_N} \mathbb{P}_x \left( \hat{X}^{N, \alpha, K'} \in F_\varepsilon \right) \\ &\leq -N \left[ \inf_{x \in K} \inf_{\psi \in F_\varepsilon} S_{\alpha, K'}(x, T, \psi) - \delta \right]. \end{aligned}$$

To prove the result, it suffices to show that

$$\inf_{x \in K} \inf_{\psi \in F_\varepsilon} S_{\alpha, K'}(x, T, \psi) > 0.$$

Arguing by contradiction, suppose that this infimum is 0. Then there are sequences  $\{x_n\} \subset K$  and  $\{\psi_n\} \subset C([0, T]:\mathbb{R}^d)$  such that  $\psi_n \in F_\varepsilon$  for each  $n$  and

$$\lim_{n \rightarrow \infty} S_{\alpha, K'}(x_n, T, \psi_n) = 0.$$

Since  $S_{\alpha, K'}(x, T, \phi) < \infty$  if and only if  $\phi(0) = x$ , we can assume without loss of generality that  $x_n = \psi_n(0)$  for every  $n$ . For each  $\varepsilon' > 0$ ,

$$\psi_n \in \left\{ \phi \in C([0, T]:\mathbb{R}^d) : S_{\alpha, K'}(y, T, \phi) \leq \varepsilon' \text{ for some } y \in K \right\}$$

whenever  $n$  is sufficiently large. Since  $K$  is compact, Theorem 3 ensures that  $\{\psi_n\}$  is pre-compact in  $C([0, T]:\mathbb{R}^d)$ , and so there is a convergent subsequence of  $\{\psi_n\}$ . Denoting this subsequence by  $\{\psi_{n_k}\}$  and its limit by  $\psi$ , we have that

$$\lim_{k \rightarrow \infty} (\psi_{n_k}, x_{n_k}) = \lim_{k \rightarrow \infty} (\psi_{n_k}, \psi_{n_k}(0)) = (\psi, \psi(0)).$$

Since  $\phi \mapsto S_{\alpha, K'}(\phi(0), T, \phi)$  is lower semicontinuous, it follows that

$$S_{\alpha, K'}(\psi(0), T, \psi) \leq \lim_{k \rightarrow \infty} S_{\alpha, K'}(\psi_{n_k}(0), T, \psi_{n_k}) = 0,$$

which says that  $\psi(t) = \varphi_t(\psi(0))$ . However, this is a contradiction, since  $\psi \in F_\varepsilon$ . The result follows. □

**Lemma 14.** *Let  $U$  be a fundamental neighborhood of an attractor  $A \subset \Delta^o$  such that  $\bar{U} \subset \Delta^o$ . Then for every  $T_0 \in (0, \infty)$ , there are  $c_0 \in (0, \infty)$ ,  $T \in (T_0, \infty)$ , and  $N_0 \in \mathbb{N}$  such that*

$$\sup_{x \in U \cap \Delta_N} \mathbb{P}_x \left( X_{[NT]}^N \in U^c \right) \leq \exp(-c_0 N)$$

for all  $N \geq N_0$ .

*Proof.* Let  $\alpha \doteq \text{dist}(A, U^c)$ . Since  $U$  is a fundamental neighborhood of the attractor  $A$ , we can find  $T > T_0$  such that

$$\sup_{t \geq T} \sup_{x \in U} \text{dist}(\varphi_t(x), A) < \alpha/2.$$

Let  $K \in \mathcal{K}$  be a compact set containing  $U$ . From Lemma 3 there exists a  $\sigma \in (0, \alpha/4)$  and a  $K' \in \mathcal{K}$  such that  $N^\sigma(\gamma^+(U)) \subset K'$ , where  $\gamma^+(U) = \cup_{x \in U} \gamma^+(x)$ . Then for each  $x \in U \cap \Delta_N$ , we have

$$\begin{aligned}
 \mathbb{P}_x \left( X_{[NT]}^N \in U^c \right) &\leq \mathbb{P}_x \left( \text{dist} \left( X_{[NT]}^N, A \right) > \alpha \right) \leq \mathbb{P}_x \left( \left\| X_{[NT]}^N - \varphi_T(x) \right\| > \alpha/2 \right) \\
 &\leq \mathbb{P}_x \left( \left\| X_{[NT]}^N - \hat{X}^N(T) \right\| + \left\| \hat{X}^N(T) - \varphi_T(x) \right\| > \alpha/2 \right) \\
 &\leq \mathbb{P}_x \left( D_T^N > \sigma \right) + \mathbb{P}_x \left( \left\| X_{[NT]}^N - \hat{X}^N(T) \right\| > \alpha/4, D_T^N \leq \sigma \right).
 \end{aligned}
 \tag{19}$$

Using the Markov property, we have

$$\begin{aligned}
 \mathbb{P}_x \left( \left\| X_{[NT]}^N - \hat{X}^N(T) \right\| > \alpha/4, D_T^N \leq \sigma \right) &\leq \mathbb{P}_x \left( \left\| X_{[NT]}^N - X_{[NT]+1}^N \right\| > \alpha/4, D_T^N \leq \sigma \right) \\
 &\leq \sup_{x \in K' \cap \Delta_N} \mathbb{P}_x \left( \left\| X_1^N - x \right\| > \alpha/4 \right).
 \end{aligned}$$

From Assumption 3 we have that for every  $\lambda > 0$ ,

$$C(\lambda) \doteq \sup_{N \in \mathbb{N}} \sup_{x \in K' \cap \Delta_N} \mathbb{E}_x \left( e^{\lambda N \left\| X_1^N - x \right\|} \right) < \infty.$$

Thus, for any  $\lambda > 0$ ,

$$\sup_{x \in K' \cap \Delta_N} \mathbb{P}_x \left( \left\| X_1^N - x \right\| > \alpha/4 \right) \leq c(\lambda) e^{-\lambda N \alpha/4}.$$

The result follows on using the above estimate and Lemma 13 in (19). □

The following lemma says that for every open  $U \subset \Delta^\circ$ , the support of  $\mu_N$  (when it exists) has a nonempty intersection with  $U$  when  $N$  is sufficiently large.

**Lemma 15.** *Suppose that for each  $N \in \mathbb{N}$ ,  $X^N$  has a QSD  $\mu_N$ . Then for each open  $U \subset \Delta^\circ$ , there is some  $N_0 \in \mathbb{N}$  such that  $\mu_N(U) > 0$  for all  $N \geq N_0$ .*

*Proof.* Let  $N_0$  be large enough so that  $U \cap \Delta_N^\circ$  is nonempty for all  $N \geq N_0$ . Fix  $N \geq N_0$ ,  $x \in U \cap \Delta_N^\circ$ , and  $w \in \Delta_N^\circ$  with  $\mu_N(w) > 0$ . From Assumption 4(a), there is a  $k \in \mathbb{N}$  such that  $\mathbb{P}_w(X_k^N = x) > 0$ . Then

$$\mu_N(U) \geq \mu_N(x) = \frac{\sum_{y \in \Delta_N^\circ} \mu_N(y) \mathbb{P}_y \left( X_k^N = x \right)}{\sum_{z \in \Delta_N^\circ} \sum_{y \in \Delta_N^\circ} \mu_N(y) \mathbb{P}_y \left( X_k^N = z \right)} \geq \frac{\mu_N(w) \mathbb{P}_w \left( X_k^N = x \right)}{\sum_{z \in \Delta_N^\circ} \sum_{y \in \Delta_N^\circ} \mu(y) \mathbb{P}_y \left( X_k^N = z \right)} > 0.$$

□

The following lemma quantifies the asymptotic behavior of the sequence  $\{\lambda_N\}$  introduced in (5).

**Theorem 5.** *Suppose that for each  $N \in \mathbb{N}$ ,  $X^N$  has a QSD  $\mu_N$ . Then there exist  $c, c' \in (0, \infty)$  such that for all  $N \in \mathbb{N}$ ,*

$$0 \leq 1 - \lambda_N \leq c' \exp(-cN).$$

*Proof.* From Assumption 2 and Corollary 1 there exists an attractor  $A$  in  $\Delta^\circ$ . Let  $U \subset \Delta^\circ$  be a fundamental neighborhood of  $A$ . From Lemma 14 there are  $c_0 \in (0, \infty)$  and  $T, N_1 \in \mathbb{N}$  such that for all  $N \geq N_1$ ,

$$\sup_{x \in U \cap \Delta_N} \mathbb{P}_x \left( X_{NT}^N \in U^c \right) \leq \exp(-c_0 N). \tag{20}$$

From Lemma 15 there is an  $N_2 \in \mathbb{N}$  such that for all  $N \geq N_2$ ,  $\mu_N(U) > 0$ . Fixing  $N \geq N_1 \vee N_2$ , we have

$$\begin{aligned} \lambda_N^T \mu_N(U) &= \sum_{x \in \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in U) \mu_N(x) \\ &\geq \sum_{x \in U \cap \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in U) \mu_N(x) \\ &\geq \inf_{x \in U \cap \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in U) \sum_{x \in U \cap \Delta_N^o} \mu_N(x) \\ &= \inf_{x \in U \cap \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in U) \mu_N(U). \end{aligned}$$

Thus, for all  $N \geq N_1 \vee N_2$ ,

$$\lambda_N \geq \lambda_N^T \geq \inf_{x \in U \cap \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in U) = 1 - \sup_{x \in U \cap \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in U^c) \geq 1 - \exp(-c_0 N),$$

where the last inequality uses (20). The result follows. □

For  $\delta > 0$ ,  $T \in \mathbb{N}$ , and  $K \in \mathcal{K}$ , let

$$\beta_{\delta,K}^N(T) \doteq \sup_{x \in \Delta_N \cap K} \mathbb{P}_x [\|\hat{X}^N - \varphi_*(x)\|_{*,T} \geq \delta]. \tag{21}$$

The following lemma gives a different lower bound on  $\lambda_N$ . This bound will be needed in the proof of Theorem 6 below.

**Lemma 16.** *Suppose that for each  $N \in \mathbb{N}$ ,  $X^N$  has a QSD  $\mu_N$ . Let  $A$  be an attractor in  $\Delta^o$ ,  $\tilde{U} \subset \Delta^o$  an open set containing  $A$ , and  $K \in \mathcal{K}$  such that  $\tilde{U} \subset K$ . Then there exist  $\delta > 0$  and  $T, N_0 \in \mathbb{N}$  such that  $\lambda_N^T \geq 1 - \beta_{\delta,K}^N(T)$  for each  $N \geq N_0$ .*

*Proof.* Since  $A$  is an attractor, there is a fundamental neighborhood  $U$  of  $A$  contained in  $\tilde{U}$ . Thus we can find a  $\delta > 0$  and  $T \in \mathbb{N}$  such that  $N^\delta(\varphi_T(\bar{U})) \subset U$ . From Lemma 15 we can find an  $N_0 \in \mathbb{N}$  such that  $\mu_N(U) > 0$  for all  $N \geq N_0$ . Following the proof of Theorem 5, we see that

$$\lambda_N^T \mu_N(U) \geq \left( 1 - \sup_{x \in U \cap \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in U^c) \right) \mu_N(U).$$

From our choice of  $U$  and  $\delta$  it now follows that

$$\lambda_N^T \geq 1 - \sup_{x \in U \cap \Delta_N^o} \mathbb{P}_x (X_{NT}^N \in (N^\delta(\varphi_T(\bar{U})))^c) \geq 1 - \beta_{\delta,K}^N(T).$$

□

A key consequence of the following theorem is that the support of any weak limit point of  $\mu_N$  is contained in  $\Delta^o$ . This, along with a further characterization of the support of such weak limit points given in Corollary 2, is a key element in the proof of Theorem 1.



**Theorem 6.** *Suppose that for each  $N \in \mathbb{N}$ ,  $X^N$  has a QSD  $\mu_N$ . Then, for every  $\delta > 0$ ,  $T \in \mathbb{N}$ , and  $K \in \mathcal{K}$ , there exists an open neighborhood  $V_K$  of  $\partial\Delta$  in  $\Delta$  such that*

$$\limsup_{N \rightarrow \infty} \mu_N(V_K) \leq \limsup_{N \rightarrow \infty} \frac{\beta_{\delta,K}^N(T)}{\inf_{x \in V_K \cap \Delta_N} \mathbb{P}_x[\hat{X}^N(T) \in \partial\Delta]} = 0. \tag{22}$$

*Suppose in addition that  $\mu_N$  converges along some subsequence to some probability measure  $\mu$  on  $\Delta$ . Then there is an open neighborhood  $V_0$  of  $\partial\Delta$  in  $\Delta$  such that  $\mu(V_0) = 0$ .*

*Proof.* Fix  $\delta, T, K$  as in the statement of the theorem. Let

$$\delta_0 \doteq \frac{1}{2} \inf_{t \in [0, T]} \inf_{x \in K} \text{dist}(\varphi_t(x), \partial\Delta)$$

and let  $K' \doteq \overline{N^{\delta_0}(\varphi_{[0, T]}(K))}$ , and consider the closed set

$$F \doteq \left\{ \phi \in C([0, T]; \mathbb{R}^d) : \|\phi - \varphi.(\phi(0))\|_{*, T} \geq \delta_0 \right\}.$$

Fix  $\alpha \in (0, \delta_0)$  and  $K_1 \in \mathcal{K}$  that contains some open neighborhood of  $K'$ . Then from Theorem 4,

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \sup_{x \in K \cap \Delta_N} \mathbb{P}_x(\hat{X}^N \in F) &= \limsup_{N \rightarrow \infty} \frac{1}{N} \log \sup_{x \in K \cap \Delta_N} \mathbb{P}_x(\hat{X}^{N, \alpha, K_1} \in F) \\ &\leq - \inf_{x \in K} \inf_{\phi \in F} S_{\alpha, K_1}(x, T, \phi) \doteq -c(K). \end{aligned}$$

Clearly  $c(K) > 0$ . From Assumption 4(b) we can find an open neighborhood  $V_K$  of  $\partial\Delta$  such that

$$\liminf_{N \rightarrow \infty} \inf_{x \in V_K \cap \Delta_N} \frac{1}{N} \log \mathbb{P}_x(\hat{X}^N(T) \in \partial\Delta) \geq -c(K)/4.$$

Combining last two displays, we can find an  $N_1 \in \mathbb{N}$  such that for all  $N \geq N_1$ ,

$$\frac{\beta_{\delta,K}^N(T)}{\inf_{x \in V_K \cap \Delta_N} \mathbb{P}_x[\hat{X}^N(T) \in \partial\Delta]} = \frac{\sup_{x \in \Delta_N \cap K} \mathbb{P}_x \left( \|\hat{X}^N - \varphi.(x)\|_{*, T} \geq \delta \right)}{\inf_{x \in V_K \cap \Delta_N} \mathbb{P}_x[\hat{X}^N(T) \in \partial\Delta]} \leq \exp(-Nc(K)/2),$$

which converges to 0 as  $N \rightarrow \infty$ . This proves the last equality in (22).

Next, from Assumption 2 and Corollary 1 there exists an attractor  $A$  in  $\Delta^o$ . Let  $\tilde{U} \in \Delta^o$  be an open set containing  $A$ , and let  $K \in \mathcal{K}$  be such that  $\tilde{U} \subset K$ . Then from Lemma 16 there exist  $\delta > 0$  and  $T, N_0 \in \mathbb{N}$  such that

$$\lambda_N^T \geq 1 - \beta_{\delta,K}^N(T) \quad \text{for each } N \geq N_0.$$

Since  $\mu_N(\Delta_N^o) = 1$ , we have, with  $V_K$  given as in the first part of the theorem,

$$\begin{aligned} 1 - \beta_{\delta,K}^N(T) &\leq \lambda_N^T \mu_N(\Delta_N^o) \\ &= \sum_{x \in \Delta_N^o} (1 - \mathbb{P}_x(X_{NT}^N \in \partial\Delta)) \mu_N(x) \\ &= \sum_{x \in V_K \cap \Delta_N^o} (1 - \mathbb{P}_x(\hat{X}^N(T) \in \partial\Delta)) \mu_N(x) + \sum_{x \in \Delta_N^o \setminus V_K} (1 - \mathbb{P}_x(\hat{X}^N(T) \in \partial\Delta)) \mu_N(x) \end{aligned}$$

$$\begin{aligned} &\leq \left( 1 - \inf_{x \in V_K \cap \Delta_N^o} \mathbb{P}_x(\hat{X}^N(T) \in \partial \Delta) \right) \mu_N(V_K) + \mu_N(\Delta_N^o \setminus V_K) \\ &= 1 - \inf_{x \in V_K \cap \Delta_N^o} \mathbb{P}_x(\hat{X}^N(T) \in \partial \Delta) \mu_N(V_K). \end{aligned}$$

Rearranging the previous inequality, we obtain

$$\mu_N(V_K) \leq \frac{\beta_{\delta,K}^N(T)}{\inf_{x \in V_K \cap \Delta_N^o} \mathbb{P}_x(\hat{X}^N(T) \in \partial \Delta)}.$$

This proves the first inequality in (22).

Finally, let  $V_0$  be an open neighborhood of  $\partial \Delta$  such that  $\bar{V}_0 \subset V_K$ . From the first part of the theorem, taking the limit along the convergent subsequence,

$$\mu(V_0) \leq \liminf_{N \rightarrow \infty} \mu_N(V_0) \leq \liminf_{N \rightarrow \infty} \mu_N(V_K) = 0.$$

The result follows. □

The following theorem proves the invariance of  $\mu$  under the flow  $\{\varphi_t\}$ .

**Theorem 7.** *Suppose that for each  $N \in \mathbb{N}$ ,  $X^N$  has a QSD  $\mu_N$ , and suppose that  $\mu_N$  converges along some subsequence to some probability measure  $\mu$ . Then  $\mu$  is invariant under  $\{\varphi_t\}$ . In particular,  $\mu(\varphi_t^{-1}(B)) = \mu(B)$  for each measurable set  $B \subset \Delta$  and  $t \geq 0$ .*

*Proof.* From Corollary 1, for each  $i \in \{1, \dots, l\}$ ,  $K_i$  is an attractor. Fix  $1 \leq i \leq l$ ,  $\delta > 0$ , and  $K \in \mathcal{K}$  such that  $N^\gamma(K_i) \subset K$  for some  $\gamma > 0$ . Let  $\beta_{\delta,K}^N$  be as in (21). It suffices to show that for any continuous and bounded  $f: \Delta \rightarrow \mathbb{R}$  and  $t > 0$ ,  $\mu(f) = \mu(f \circ \varphi_t)$ . Fix  $f$  and  $t$  as above and let  $\varepsilon > 0$  be arbitrary. Using the fact that  $\{\mu_N\}$  (considered along the convergent subsequence) is tight, we can assume that the  $K$  chosen above satisfies

$$\sup_{N \geq 1} \mu_N(K^c) \leq \frac{\varepsilon}{2\|f\|_\infty}.$$

Let  $t_N = \lfloor Nt \rfloor / N$ . Note that  $t_N \rightarrow t$  as  $N \rightarrow \infty$ , and from Theorem 5,  $\lambda_N^{t_N} \rightarrow 1$  as  $N \rightarrow \infty$ . For a bounded  $g: \Delta^o \rightarrow \mathbb{R}$  and  $k \in \mathbb{N}$ , let

$$\mathcal{P}_k^N f(x) \doteq \mathbb{E}_x [f(X_k^N); \tau_\partial^N > k], \quad x \in \Delta_N^o,$$

and

$$\mathcal{P}_k^{N,*} f(x) \doteq \mathbb{E}_x [f(X_k^N)], \quad x \in \Delta_N^o.$$

Then

$$\mu_N(f) = \lambda_N^{-t_N} \mu_N \left( \mathcal{P}_{\lfloor Nt \rfloor}^N f \right).$$

In particular, as  $N \rightarrow \infty$ ,

$$\left| \mu_N(f) - \mu_N \left( \mathcal{P}_{\lfloor Nt \rfloor}^N f \right) \right| \leq \|f\|_\infty \left| 1 - \lambda_N^{t_N} \right| \rightarrow 0.$$

Also,

$$\left| \mu_N \left( \mathcal{P}_{\lfloor Nt \rfloor}^N f \right) - \mu_N(f \circ \varphi_t) \right| \leq \frac{\varepsilon}{2\|f\|_\infty} 2\|f\|_\infty + \sup_{x \in K} \left| \mathcal{P}_{\lfloor Nt \rfloor}^N f(x) - f \circ \varphi_t(x) \right|.$$

For each  $x \in K \cap \Delta_N$ ,

$$\begin{aligned} \left| \mathcal{P}_{\lfloor Nt \rfloor}^N f(x) - f \circ \varphi_t(x) \right| &\leq \left| \mathcal{P}_{\lfloor Nt \rfloor}^{N,*} f(x) - \mathcal{P}_{\lfloor Nt \rfloor}^N f(x) \right| + \left| \mathcal{P}_{\lfloor Nt \rfloor}^{N,*} f(x) - f \circ \varphi_t(x) \right| \\ &\leq \|f\|_\infty \mathbb{P}_x(\tau_\delta^N \leq \lfloor Nt \rfloor) + \left| \mathcal{P}_{\lfloor Nt \rfloor}^{N,*} f(x) - f \circ \varphi_t(x) \right|, \end{aligned}$$

and Assumption 1 ensures that as  $N \rightarrow \infty$ ,

$$\sup_{x \in K} \left| \mathcal{P}_{\lfloor Nt \rfloor}^{N,*} f(x) - f \circ \varphi_t(x) \right| \rightarrow 0.$$

Let  $\tilde{\delta} \doteq \inf_{x \in K, 0 \leq s \leq t} \text{dist}(\varphi_s(x), \partial \Delta) > 0$ , and note that Assumption 1 ensures that as  $N \rightarrow \infty$ ,

$$\sup_{x \in K} \mathbb{P}_x(\tau_\delta^N \leq \lfloor Nt \rfloor) \leq \sup_{x \in K} \mathbb{P}_x(\|X^N - \varphi_*(x)\|_{*,t} > \tilde{\delta}) \rightarrow 0.$$

Combining the two previous convergence properties, we see that as  $N \rightarrow \infty$ ,

$$\left| \mathcal{P}_{\lfloor Nt \rfloor}^N f(x) - f \circ \varphi_t(x) \right| \rightarrow 0,$$

and therefore that

$$|\mu(f) - \mu(f \circ \varphi_t)| \leq \limsup_{N \rightarrow \infty} |\mu_N(f) - \mu_N(f \circ \varphi_t)| \leq \varepsilon.$$

Since  $\varepsilon > 0$  is arbitrary, the result follows. □

We now recall the definition of the Birkhoff center of  $\{\varphi_t\}$ .

**Definition 7.** The Birkhoff center of  $\{\varphi_t : t \geq 0\}$  is

$$BC(\varphi) \doteq \overline{\{x \in \Delta : x \in \omega(x)\}}.$$

**Lemma 17.** *The Birkhoff center of  $\{\varphi_t : t \geq 0\}$  is contained in the closure of  $\mathcal{R}_{AP}$ . Furthermore,  $BC(\varphi) \cap \Delta^o \subset \mathcal{R}_{AP}^*$ .*

*Proof.* Let  $\delta, T > 0$  and suppose that  $x \in \omega(x)$ . There is a sequence of time instants  $t_i \uparrow \infty$  such that  $\varphi_{t_i}(x) \rightarrow x$ , so if we let

$$j = \min\{i : t_i > T \text{ and } \|\varphi_{t_i}(x) - x\| < \delta\},$$

then  $(x, x, \varphi_{t_j}(x), x, x)$  is a  $(\delta, T)$  AP-pseudo-orbit from  $x$  to  $x$ . Since  $\delta, T$  are arbitrary,  $x \in \mathcal{R}_{AP}$ . This proves the first part of the lemma. The second part is now immediate on using Assumption 2(a). □

We will use the Poincaré recurrence theorem given below. For a proof see [10, Theorem 4.1.19].

**Theorem 8.** *Let  $\nu$  be a measure which is invariant under  $\{\varphi_t\}$ . Then for each measurable  $B \subset \Delta$  and  $T > 0$ ,*

$$\nu(\{x \in B : \{\varphi_t(x)\}_{t \geq T} \subset \Delta \setminus B\}) = 0.$$

The next result is a consequence of Lemma 17 and Theorem 8. It shows that the support of  $\mu$  is contained in  $\mathcal{R}_{AP}^*$ .

**Corollary 2.** *Suppose that for each  $N \in \mathbb{N}$ ,  $X^N$  has a QSD  $\mu_N$ , and suppose that  $\mu_N$  converges along some subsequence to some probability measure  $\mu$ . Then  $\text{supp}(\mu) \subset \mathcal{R}_{AP}^*$ .*

*Proof.* From Theorem 7,  $\mu$  is invariant under  $\{\varphi_t\}$ . Enumerate the  $d$ -dimensional rationals in  $\Delta$  as  $\mathbb{Q}^d \doteq \{q_1, q_2, \dots\}$ , and for  $m, n \in \mathbb{N}$ , denote the ball of radius  $n^{-1}$  centered at  $q_m$  by  $B(q_m, n^{-1})$ . Then for each  $m, n \in \mathbb{N}$ , Theorem 8 says that

$$\mu(\tilde{B}(q_m, n^{-1})) = \mu(B(q_m, n^{-1})),$$

where

$$\tilde{B}(q_m, n^{-1}) \doteq \{x \in B(q_m, n^{-1}) : \text{there exist } t_k \uparrow \infty \text{ with } \varphi_{t_k}(x) \in B(q_m, n^{-1}) \text{ for all } k \in \mathbb{N}\}.$$

Let  $R \doteq \bigcap_{n=1}^\infty \bigcup_{m=1}^\infty \tilde{B}(q_m, n^{-1})$ ; then

$$1 = \mu\left(\bigcap_{n=1}^\infty \bigcup_{m=1}^\infty B(q_m, n^{-1})\right) = \mu\left(\bigcap_{n=1}^\infty \bigcup_{m=1}^\infty \tilde{B}(q_m, n^{-1})\right) = \mu(R),$$

which together with Theorem 6 implies that  $\text{supp}(\mu) \subseteq \bar{R} \cap \Delta^o$ . Furthermore, if  $x \in R$ , then  $x \in \omega(x)$ , so  $R \subseteq \text{BC}(\varphi)$ , and consequently  $\bar{R} \cap \Delta^o \subseteq \text{BC}(\varphi) \cap \Delta^o$ . It now follows from Lemma 17 that

$$\text{supp}(\mu) \subseteq \bar{R} \cap \Delta^o \subseteq \text{BC}(\varphi) \cap \Delta^o \subseteq \mathcal{R}_{AP}^*.$$

□

Combining the results of Corollary 2, Theorem 7, and Theorem 5, we have most of Theorem 1. In particular, we have the lower bound on probabilities of non-extinction given in Theorem 1, and we also have that the limit points  $\mu$  of the QSD are invariant under the flow and are supported on the union of absorption-preserving recurrence classes in the interior. The final step is to show that the support in fact lies in the union of the interior attractors. For this we will introduce another notion of recurrence which is given in terms of the quasipotential associated with the rate functions in the underlying large deviation principles.

### 6. Quasipotential and chain recurrence

In this section we suppose that Assumptions 1, 2, and 3 are satisfied. Recall the rate function  $S$  introduced in (14). For  $x, y \in \Delta^o$ , let  $\mathcal{C}(x, y, T) \doteq \{\phi \in C([0, T]: \Delta^o) : \phi(0) = x, \phi(T) = y\}$ , and define

$$V(x, y) \doteq \liminf_{T \rightarrow \infty} \inf_{\phi \in \mathcal{C}(x, y, T)} S(x, T, \phi). \tag{23}$$

For  $x, y \in \Delta^o$ , we say  $x$  leads to  $y$  in  $\Delta^o$  if  $V(x, y) = 0$ , and we write  $x <_V y$ . We say  $x \in \Delta^o$  is  $V$ -chain recurrent if  $x <_V x$ . The collection of  $V$ -chain recurrent points is denoted by  $\mathcal{R}_V$ . For  $x, y \in \Delta^o$  we say  $x \sim_V y$  if  $x <_V y$  and  $y <_V x$ . Equivalence classes under  $\sim_V$  will be called  $V$ -basic classes, and the equivalence class associated with an  $x \in \mathcal{R}_V$  will be denoted by  $[x]_V$ . For  $x, y \in \mathcal{R}_V$  we say  $[x]_V < [y]_V$  if  $x <_V y$ . A  $V$ -basic class  $[x]_V$  is said to be maximal if, whenever  $y \in \mathcal{R}_V$  satisfies  $[x]_V < [y]_V$ , we have that  $y \in [x]_V$ . A maximal  $V$ -basic class is called a  $V$ -quasiattractor. The following is the main result of this section.

**Theorem 9.** *We have  $\mathcal{R}_{AP}^* = \mathcal{R}_V$ , and for each  $x \in \mathcal{R}_V$ ,  $[x]_V = [x]_{AP}$ . In particular there are finitely many  $V$ -chain recurrent points, and for every  $x \in \mathcal{R}_V$ ,  $[x]_V$  is a closed set. Furthermore,  $K_i$  for  $i = 1, \dots, l$  is a  $V$ -quasiattractor, while  $K_i$  for  $i = l + 1, \dots, v$  is not a  $V$ -quasiattractor.*

Before proving Theorem 9 we will establish some basic results regarding  $V(\cdot, \cdot)$  and  $\mathcal{R}_V$ . The following lemma is a consequence of the stability properties of the ODE (4) studied in Lemma 3 and the property that low-cost trajectories closely follow the solution of the ODE.

**Lemma 18.** *Let  $\alpha \in (0, \infty)$  and  $K \in \mathcal{K}$ . Let  $T_0 \in (0, \infty)$  and suppose  $T_n \in [T_0, \infty)$  for all  $n \in \mathbb{N}$ . Let  $\phi_n \in C([0, T_n]; \Delta^o)$  be such that  $\phi_n(0) \in \mathcal{V}_{\alpha, K}$  for each  $n \geq 1$ . Suppose that  $S(\phi_n(0), T_n, \phi_n) \rightarrow 0$  as  $n \rightarrow \infty$ . With  $\alpha_0$  and  $M_0$  as in Lemma 3, let  $\alpha_1 = \frac{\alpha}{2} \wedge \alpha_0$  and  $K_1 = \overline{B_{M_1}(0)}$ , where  $M_1 = 1 + (M_0 \vee \sup_{x \in K} \|x\|)$ . Then, for some  $k \geq 1$ ,  $\phi_n(t) \in \mathcal{V}_{\alpha_1, K_1}$  for all  $n \geq k$  and  $t \in [0, T_n]$ .*

*Proof.* For  $n \geq 1$ , let

$$\tau(\phi_n) \doteq \inf\{t \in [0, T_n] : \text{dist}(\phi_n(t), \partial\Delta) \leq \alpha_1 \text{ or } \|\phi_n(t)\| \geq M_1\},$$

where the infimum is taken to be  $T_n$  if the above set is empty. Note that the result holds trivially if the above set is empty for all but finitely many  $n$ . Now, arguing by contradiction, suppose the set is nonempty for infinitely many  $n$ . Consider the subsequence along which the above sets are nonempty, and denote the subsequence once more by  $\{n\}$ . Also assume without loss of generality that  $\gamma_n \doteq S(\phi_n(0), T_n, \phi_n) \leq 1$  for every  $n$ .

We claim that there is a  $\delta > 0$  and  $k_0 \in \mathbb{N}$  such that  $\tau(\phi_n) \geq \delta$  for all  $n \geq k_0$ . Indeed, otherwise, by passing to a further subsequence (once more denoted by  $\{n\}$ ) we can find a sequence  $\delta_n \rightarrow 0$  such that for every  $n$ ,

$$\phi_n(0) \in \mathcal{V}_{\alpha, K}, \phi_n(\delta_n) \in [\mathcal{V}_{\alpha_1, K_1}]^c.$$

Since  $S(\phi_n(0), T_n, \phi_n) \leq \gamma_n \leq 1$ , we must have from the compactness-of-level-sets property in Theorem 3 that  $\phi_n(0)$  and  $\phi_n(\delta_n)$  converge along a subsequence to the same limit, which is a contradiction.

Let  $\delta > 0$  be such that  $\tau(\phi_n) \geq \delta$  for all sufficiently large  $n$ . For each such  $n$  let  $\hat{t}_n \doteq \tau(\phi_n) - \delta$ , and define  $\phi_n^* : [0, \delta] \rightarrow \Delta^o$  as  $\phi_n^*(t) = \phi_n(t + \hat{t}_n)$ ,  $t \in [0, \delta]$ . Then  $\phi_n^* \in C([0, \delta]; \mathcal{V}_{\alpha_1, K_1})$ . Also,

$$\begin{aligned} \int_0^\delta L_{\alpha_1, K_1}(\phi_n^*(t), \dot{\phi}_n^*(t)) dt &= \int_0^\delta L(\phi_n^*(t), \dot{\phi}_n^*(t)) dt \\ &\leq \int_0^{T_n} L(\phi_n(t), \dot{\phi}_n(t)) dt = \gamma_n \leq 1. \end{aligned} \tag{24}$$

In particular,

$$\{\phi_n^*\} \subset \cup_{x \in K_1} \Phi_{x, \alpha_1, K_1, \delta}(1).$$

From Theorem 3, the latter set is compact, and so, along some subsequence,  $\phi_n^*$  converges to some  $\phi^*$  in  $C([0, \delta]; \mathcal{V}_{\alpha_1, K_1})$ . Using the compactness of level sets again, we have from (24) and the fact that  $\gamma_n \rightarrow 0$  that

$$\int_0^\delta L_{\alpha_1, K_1}(\phi^*(t), \dot{\phi}^*(t)) dt = 0.$$

In particular,  $\phi^*(t)$  solves the ODE (4), namely  $\phi^*(t) = \varphi_t(\phi^*(0))$  for  $t \in [0, \delta]$ . Since  $\phi^*(0) \in \mathcal{V}_{\alpha_1, K_1}$ , in view of Lemma 3, we must have that  $\|\phi^*(\delta)\| < M_1$  and  $\text{dist}(\phi^*(\delta), \partial\Delta) > \alpha_1$ . However, from the definition of  $\tau(\phi_n)$ , we have that for each  $n$ ,  $\phi_n^*(\delta)$  satisfies either  $\text{dist}(\phi_n^*(\delta), \partial\Delta) \leq \alpha_1$  or  $\|\phi_n^*(\delta)\| \geq M_1$ . This is a contradiction, since  $\phi_n^*$  converges to  $\phi^*$  (along some subsequence). The result follows.  $\square$

**Corollary 3.** *Let  $K \in \mathcal{K}$  and  $T_0 > 0$ . Then there exist a  $\gamma > 0$  and an  $A_1 \in \mathcal{K}$  such that whenever for some  $x \in K$  we have  $T^x \in [T_0, \infty)$  and  $\phi^x \in C([0, T^x]; \Delta^o)$  with  $\phi^x(0) = x$  and  $S(x, T^x, \phi^x) \leq \gamma$ , we have  $\phi^x(t) \in A_1$  for all  $t \in [0, T^x]$ .*

*Proof.* Let  $\alpha \in (0, \infty)$  be such that  $K = \mathcal{V}_{\alpha, K}$ . Let  $\alpha_1, K_1$  be as in Lemma 18. We argue by contradiction. Suppose the statement in the corollary is false. Then there are sequences  $\gamma_n \downarrow 0$ ,  $x_n \in K$ , time instants  $T^{x_n} \in [T_0, \infty)$ , trajectories  $\phi^{x_n} \in C([0, T^{x_n}]; \Delta^o)$ , and sets  $B_n = \{x \in \Delta^o : \|x\| \leq n, \text{dist}(x, \partial \Delta) \geq 1/n\}$  such that  $S(\phi^{x_n}(0), T^{x_n}, \phi^{x_n}) \leq \gamma_n$  and  $\phi^{x_n}(t_n) \in B_n^c$  for some  $t_n \in [0, T^{x_n}]$ . However, from Lemma 18, there exists a  $k \in \mathbb{N}$  such that  $\phi^{x_n}(t) \in \mathcal{V}_{\alpha_1, K_1}$  for all  $n \geq k$  and all  $t \in [0, T^{x_n}]$ , which is clearly a contradiction since we can find an  $n_0 > k$  such that  $\mathcal{V}_{\alpha_1, K_1} \subset B_n$  for all  $n \geq n_0$ .  $\square$

The following continuity property of  $V$ , which is a consequence of the continuity of  $L_{\alpha, K}$  shown in Lemma 12, will be needed in the proof of Theorem 9.

**Lemma 19.** *Suppose  $x_n, x \in \Delta^o$  are such that  $x_n \rightarrow x$  as  $n \rightarrow \infty$ . Then for every  $y \in \Delta^o$ ,  $V(x_n, y) \rightarrow V(x, y)$  and  $V(y, x_n) \rightarrow V(y, x)$ .*

*Proof.* Fix  $x \in \Delta^o$  and let  $G \subset \Delta^o$  be a bounded open ball containing  $x$  such that  $\bar{G} \subset \Delta^o$ . Without loss of generality assume that  $x_n \in G$  for every  $n$ . Choose  $\alpha \in (0, 1)$  and  $K \in \mathcal{K}$  such that  $\mathcal{V}_{\alpha, K} \supset \bar{G}$ . Since  $\bar{G}$  is compact, from Lemma 12, we have that

$$\sup_{z \in \bar{G}, \|\beta\| \leq 1} L_{\alpha, K}(z, \beta) \doteq \kappa_0 < \infty,$$

where  $B_1(0)$  is the unit ball in  $\mathbb{R}^d$ . Let  $\varepsilon \in (0, \infty)$  be arbitrary. Take  $x_1, x_2 \in G$  such that  $x_1 \neq x_2$  and  $\|x_1 - x_2\| \leq \varepsilon/(2\kappa_0)$ . Also, fix  $y \in \Delta^o$ . From the definition of  $V(x_2, y)$  we can find a sequence  $T_k \rightarrow \infty$  and  $\phi_k \in C([0, T_k]; \Delta^o)$  such that for all  $k$ ,  $\phi_k(0) = x_2$ ,  $\phi_k(T_k) = y$ , and

$$S(x_2, T_k, \phi_k) \leq V(x_2, y) + \varepsilon/2.$$

Let  $\delta = \|x_1 - x_2\|$ ,  $\beta \doteq \frac{(x_2 - x_1)}{\|x_2 - x_1\|}$ ,  $\tilde{T}_k \doteq T_k + \delta$ , and for  $t \leq \tilde{T}_k$  define

$$\tilde{\phi}_k(t) = \begin{cases} x_1 + \beta t, & t \leq \delta, \\ \phi_k(t - \delta), & t \geq \delta. \end{cases}$$

Then

$$\begin{aligned} S(x_1, \tilde{T}_k, \tilde{\phi}_k) &= \int_0^{\tilde{T}_k} L(\tilde{\phi}_k(t), \dot{\tilde{\phi}}_k(t)) dt \\ &= \int_0^\delta L(\tilde{\phi}_k(t), \dot{\tilde{\phi}}_k(t)) dt + S(x_2, T_k, \phi_k) \\ &= \int_0^\delta L_{\alpha, K}(x_1 + \beta t, \beta) dt + S(x_2, T_k, \phi_k) \\ &\leq \kappa_0 \frac{\varepsilon}{2\kappa_0} + V(x_2, y) + \frac{\varepsilon}{2} = V(x_2, y) + \varepsilon. \end{aligned}$$

Thus  $V(x_1, y) \leq V(x_2, y) + \varepsilon$ , which proves the convergence  $V(x_n, y) \rightarrow V(x, y)$ . The proof of  $V(y, x_n) \rightarrow V(y, x)$  is similar and is omitted.  $\square$

The following result is a consequence of the compactness-of-level-sets property in Theorem 3 and the uniqueness of the path where the rate function vanishes.

**Lemma 20.** Fix  $T \in (0, \infty)$  and a  $K \in \mathcal{K}$ . For each  $\delta > 0$ , there is some  $\varepsilon \doteq \varepsilon(K, T, \delta) > 0$  such that for any  $\phi \in C([0, T]; \mathcal{V}_{\alpha, K})$  and  $x \in K$ , if  $S_{\alpha, K}(x, T, \phi) \leq \varepsilon$ , then

$$\sup_{0 \leq t \leq T} \|\phi(t) - \varphi_t(x)\| < \delta.$$

*Proof.* Arguing via contradiction, suppose that there is some  $\delta > 0$  such that for all  $\varepsilon > 0$ , there exist  $x \in K$  and  $\phi_\varepsilon \in C([0, T]; \mathcal{V}_{\alpha, K})$  such that  $S_{\alpha, K}(x, T, \phi_\varepsilon) < \varepsilon$  but  $\|\phi_\varepsilon(t) - \varphi_x(t)\| \geq \delta$  for some  $t \in [0, T]$ . Using the compactness-of-level-sets property in Theorem 3(a) and recalling that  $S_{\alpha, K}(x, T, \phi) = 0$  if and only if  $\phi(t) = \varphi_t(x)$  for  $t \in [0, T]$ , we see that

$$c \doteq \inf\{S_{\alpha, K}(x, T, \phi) : x \in K, \sup_{t \in [0, T]} \|\phi(t) - \varphi_t(x)\| \geq \delta\} > 0.$$

Thus  $c \leq S_{\alpha, K}(x, T, \phi_\varepsilon) < \varepsilon$  for all  $\varepsilon > 0$ . Letting  $\varepsilon \downarrow 0$ , we obtain  $c = 0$ , which is a contradiction.  $\square$

As an intermediate step we now prove a somewhat weaker statement than that in Theorem 9.

**Lemma 21.** Suppose that  $x \in \mathcal{R}_V$ . Then  $x \in \mathcal{R}_{AP}^*$  and  $[x]_V \subset [x]_{AP}$ .

*Proof.* Let  $y \in [x]_V$ . Then there exist time instants  $T_n \uparrow \infty$  and  $\phi_n \in C([0, T_n]; \Delta^o)$  such that for all  $n \geq 1$ ,  $\phi_n(0) = x$ ,  $\phi_n(T_n) = y$ , and  $S(x, T_n, \phi_n) < \frac{1}{n}$ . From Lemma 18 there exist  $k \in \mathbb{N}$ ,  $\alpha_1 > 0$ , and  $K_1 \in \mathcal{K}$  such that, for all  $n \geq k$ ,  $\phi_n \in C([0, T_n]; \mathcal{V}_{\alpha_1, K_1})$ .

Now fix  $T, \delta > 0$ . From Lemma 20 there is an  $\varepsilon > 0$  such that, with  $T^* = T$  and  $T^* = 2T$ , the following holds:

$$\begin{aligned} \text{if, for some } \phi \in C([0, T^*], \mathcal{V}_{\alpha_1, K_1}) \text{ and } z \in \mathcal{V}_{\alpha_1, K_1}, \quad S_{\alpha_1, K_1}(z, T^*, \phi) \leq \varepsilon, \\ \text{then } \|\phi - \varphi \cdot (z)\|_{*, T^*} < \delta. \end{aligned} \tag{25}$$

Choose  $n_0$  such that  $1/n_0 \leq \varepsilon$  and  $T_{n_0} \geq T$ . Write  $T_{n_0} = mT + t_0$ , where  $m \in \mathbb{N}$  and  $t_0 \in [0, T)$ . Then from (25), with  $\phi = \phi_{n_0}$ ,

$$\|\phi(jT) - \varphi_T(\phi((j-1)T))\| < \delta \quad \text{for } j = 1, \dots, m-1,$$

and

$$\|\phi(mT + t_0) - \varphi_{T+t_0}(\phi((m-1)T))\| < \delta.$$

Thus, with  $\xi_0 = \xi_1 = x$ ,  $\xi_2 = \phi(T)$ ,  $\dots$ ,  $\xi_m = \phi((m-1)T)$ ,  $\xi_{m+1} = \phi(T_{n_0})$ , the sequence  $\xi = (\xi_0, \dots, \xi_{m+1})$  along with time instants  $(T, T, \dots, T + t_0)$  defines a  $(\delta, T)$  AP-pseudo-orbit from  $x$  to  $y$ . Since  $\delta, T > 0$  are arbitrary,  $x <_{AP} y$ . Similarly,  $y <_{AP} x$ , showing that  $x \in \mathcal{R}_{AP}$  and  $y \in [x]_{AP}$ . This shows that  $[x]_V \subset [x]_{AP}$  and completes the proof.  $\square$

From Lemma 21 and Assumption 2 (see also Lemma 2), the closure of  $\mathcal{R}_V$  is a compact set in  $\Delta^o$ .

We now complete the proof of Theorem 9 by establishing the reverse inclusion from the one established in Lemma 21.

*Proof of Theorem 9.* From Lemma 21, if  $x \in \mathcal{R}_V$ , then  $x \in \mathcal{R}_{AP}^*$  and  $[x]_V \subset [x]_{AP}$ . Now suppose that  $x \in \mathcal{R}_{AP}^*$ . From Assumption 2 there is an  $x^* \in [x]_{AP}$  such that  $\{\varphi_t(x^*); t \geq T\}$  is dense



in  $[x]_{AP}$  for every  $T > 0$ . Fix  $y \in [x]$ . For  $n \in \mathbb{N}$ , let  $t_n, \tilde{t}_n \in (0, \infty)$  be such that  $t_n \uparrow \infty, \tilde{t}_n \uparrow \infty$  as  $n \rightarrow \infty$ , and for every  $n$ ,

$$\|\varphi_{t_n}(x^*) - x\| \leq 1/n, \quad \|\varphi_{t_n+\tilde{t}_n}(x^*) - y\| \leq 1/n.$$

Using Lemma 19 it follows that  $x <_V y$ . This shows that  $x \in \mathcal{R}_V$  and that  $[x]_{AP} \subset [x]_V$ . We thus have that  $\mathcal{R}_{AP}^* = \mathcal{R}_V$ , and for all  $x \in \mathcal{R}_V = \mathcal{R}_{AP}$ ,  $[x]_{AP} = [x]_V$ . Similar arguments show that  $[x]$  is a  $V$ -quasiattractor if and only if it is an AP-quasiattractor. The result follows.

In view of Theorem 9, henceforth we will use the qualifiers ‘ $V$ ’ and ‘AP’ interchangeably when referring to recurrence classes and quasiattractors in  $\Delta^o$ .

### 7. Proof of Theorem 1

In this section we assume that Assumptions 1, 2, 3, and 4 are satisfied. The following lemma shows that there are low-cost trajectories that take any given point in a recurrence class to any other point in the same class.

**Lemma 22.** *For any  $\gamma > 0$  and  $K \in \mathcal{R}_V$ , there is a  $T \in (1, \infty)$  such that for all  $x, y \in K$ , there exist  $T_{x,y} \in (1, T)$  and  $\phi_{x,y} \in C([0, T_{x,y}]; \Delta^o)$  with*

$$S(x, T_{x,y}, \phi_{x,y}) \leq \gamma, \quad \phi_{x,y}(0) = x, \quad \phi_{x,y}(T_{x,y}) = y.$$

*Proof.* Fix  $\gamma \in (0, 1)$  and  $K \in \mathcal{R}_V$ . Let  $\gamma_0 \doteq \sup_{z \in K, \|\beta\| \leq 1} L(z, \beta)$ . Let  $k \in \mathbb{N}$  and  $v_1, \dots, v_k \in K$  be such that for any  $x \in K$ , there exists  $1 \leq i \leq k$  with  $\|x - v_i\| \leq \gamma/(4\kappa_0)$ . For  $i, j \in \{1, \dots, k\}$ , let  $\tilde{T}_{i,j} \in (1, \infty)$  and  $\psi_{i,j} \in C([0, \tilde{T}_{i,j}]; \Delta^o)$  be such that  $\psi_{i,j}(0) = v_i, \psi_{i,j}(\tilde{T}_{i,j}) = v_j$  and  $S(v_i, \tilde{T}_{i,j}, \psi_{i,j}) \leq \gamma/2$ . Let  $x, y \in K$  be arbitrary and select  $i, j \in \{1, \dots, k\}$  such that  $\|x - v_i\| \leq \gamma/(4\kappa_0)$  and  $\|y - v_j\| \leq \gamma/(4\kappa_0)$ . Consider the continuous trajectory  $\phi_{x,y}$  in  $\Delta^o$  defined over the time interval of length  $T_{x,y} = \|x - v_i\| + \tilde{T}_{i,j} + \|y - v_j\|$  as follows:

$$x \xrightarrow[\|v_i - x\|]{\text{lin}} v_i \xrightarrow[\tilde{T}_{i,j}]{\psi_{i,j}} v_j \xrightarrow[\|v_j - y\|]{\text{lin}} y \tag{26}$$

In the above display, for a term of the form  $a \xrightarrow[c]{d} b$ , the trajectory connects the points  $a$  and  $b$  in time length  $d$  in a manner described by  $c$ . When  $c = \text{lin}$ , the trajectory is just a linear path connecting  $a$  and  $b$ ; when  $c = \psi_{i,j}$ , the trajectory is defined by  $\psi_{i,j}$  introduced above. Clearly  $S(x, T_{x,y}, \phi_{x,y}) \leq \gamma, \phi_{x,y}(0) = x$  and  $\phi_{x,y}(T_{x,y}) = y$ . Also,  $T_{x,y} \leq \max_{1 \leq i, j \leq k} \tilde{T}_{i,j} + 2 \doteq T$ . The result follows.  $\square$

Recall that for a set  $B \subset \Delta, \tau_B^N \doteq \inf\{t \geq 0: \hat{X}^N(t) \notin B\}$ . The following lemma gives an upper bound on the probabilities of long residence times of the Markov chain near non-quasiattractors.

**Lemma 23.** *Suppose that  $K_j \in \mathcal{R}_V$  is not a quasiattractor. Then we can find some  $\lambda > 0$  such that for all  $\gamma > 0$ , there is some  $N_0 \doteq N_0(\gamma)$  and  $\zeta \doteq \zeta_\gamma: \mathbb{N} \rightarrow \mathbb{R}$  satisfying  $\lim_{n \rightarrow \infty} \zeta_\gamma(n) = 0$  such that*

$$\sup_{x \in N^\lambda(K_j)} \mathbb{P}_x \left( \tau_{N^\lambda(K_j)}^N > \exp(N\gamma) \right) \leq \zeta_\gamma(N)$$

for all  $N \geq N_0$ .

*Proof.* Since  $K_j$  is not a quasiattractor, there exist  $\lambda_0 \in (0, 1), u_1 \in K_j, y_1 \in \Delta^o \cap [N^{2\lambda_0}(K_j)]^c$  such that  $u_1 <_V y_1$ . Choose  $\lambda_1 \in (0, \lambda_0)$  such that, for some  $A_0 \in \mathcal{K}, y_1 \in A_0,$

$\bigcup_{k=1}^v \overline{N^{\lambda_1}(K_k)} \subset A_0$ , and for each  $i, k \in \{1, \dots, v\}$  such that  $i \neq k$ ,  $\text{dist}(N^{\lambda_1}(K_k), N^{\lambda_1}(K_i)) \geq \lambda_1$ . From Lemma 3 we can find  $A_1 \in \mathcal{K}$  such that the forward orbit  $\gamma^+(x) \subset A_1$  for every  $x \in A_0$ . Let

$$\sup_{z \in A_1, \|\beta\| \leq 1} L_{\alpha, K}(z, \beta) \doteq \kappa_0.$$

Let  $\gamma > 0$  be given and let  $\gamma_0 = \gamma/6$ . Fix  $\delta \in \left(0, \lambda_1 \wedge \frac{\gamma_0}{\kappa_0}\right)$ . Then, denoting by  $\eta_{x,y}$  the linear trajectory from  $x$  to  $y$ ,

$$\text{for } x^*, y^* \in A_1 \text{ with } \|x^* - y^*\| \leq \delta, \quad S(x^*, \|y^* - x^*\|, \eta_{x^*, y^*}) \leq \gamma_0.$$

With  $\delta$  as above, choose  $T_{A_1}^*$  as in Lemma 9(b) (with  $A$  replaced by  $A_1$ ). Then, in view of Theorem 9, for every  $x \in A_1$ , there exists a  $t_0 \in [0, T_{A_1}^*]$  such that  $\varphi_{t_0}(x) \in N^\delta(\mathcal{R}_V)$ .

Define for  $x \in N^{\lambda_1}(K_j)$  the continuous trajectory  $\phi_x^\gamma(\cdot)$  according to the following two cases: Case I,  $\varphi_{t_0}(x) \in N^\delta(K_i)$  for some  $i \neq j$ ; Case II,  $\varphi_{t_0}(x) \in N^\delta(K_j)$ .

In Case I, we simply take  $\phi_x^\gamma(t) = \varphi_t(x)$  for  $t \in [0, t_0]$ . In particular,  $T_x^\gamma \doteq t_0$  is the length of the time interval over which the trajectory is defined.

For Case II we proceed as follows. Taking  $K = A_0$  and  $T_0 = 1$  in Corollary 6, denote by  $(\gamma^*, A^*)$  the  $(\gamma, A_1)$  given by the corollary. Let  $u_0 \in K_j$  be such that  $\|u_0 - \varphi_{t_0}(x)\| \leq \delta$ . Then  $u_0 <_V u_1 <_V y_1$ . Let  $t_1(x) \in [1, \infty)$  and  $\phi_1 \in C([0, t_1(x)]; \Delta^o)$  be such that  $\phi_1(0) = u_0$ ,  $\phi_1(t_1(x)) = y_1$ , and  $S(u_0, t_1(x), \phi_1) \leq \gamma^* \wedge \gamma/3$ . Using Lemma 22 we can assume without loss of generality that  $\sup_{w \in N^\lambda(K_j)} t_1(w) \doteq \bar{t}_1 < \infty$ . From Corollary 3,  $\phi_1(t) \in A^*$  for all  $t \in [0, t_1(x)]$ . Consider the continuous trajectory  $\phi_x^\gamma$  in  $\Delta^o$  that connects  $x$  and  $y_1$  in the following manner:

$$x \xrightarrow{\text{flow}}_{t_0} \varphi_{t_0}(x) \xrightarrow{\|\varphi_{t_0}(x) - u_0\|}_{\frac{\phi_1}{t_1(x)}} u_0 \xrightarrow{y_1}$$

The above display is interpreted similarly to (26), with a term of the form  $a \xrightarrow{c} b$ , when  $c = \text{flow}$ , representing the segment of  $\varphi_t(a)$  until it reaches  $b$ . In this case let  $T_x^\gamma = t_0 + \|\varphi_{t_0}(x) - u_0\| + t_1(x)$  denote the length of the time interval over which  $\phi_x^\gamma$  is defined.

Note that in both cases,

$$T^\gamma \doteq \sup_{x \in N^{\lambda_1}(K_j)} T_x^\gamma \leq t_0 + 1 + \bar{t}_1 < \infty.$$

Also, in both cases,  $\phi_x^\gamma(t) \in A_1 \cup A^* \doteq A_2$  for all  $t \in [0, T_x^\gamma]$ . Furthermore, in Case II,

$$S(x, \phi_x^\gamma, T_x^\gamma) \leq 0 + \gamma_0 + \gamma/3 = \gamma/2,$$

and in Case I the cost on the left side of the above display is 0.

Let  $\lambda \in (0, \lambda_1)$ ,  $\alpha' > 0$  be such that  $K' \doteq \overline{N^\lambda(A_2)} \subset \Delta^o$  and  $K' = \mathcal{V}_{\alpha', K'}$ . Extend the trajectory  $\phi_x^\gamma$  from  $[0, T_x^\gamma]$  to  $[0, T^\gamma]$  by defining  $\phi_x^\gamma(t + T_x^\gamma) \doteq \varphi_t(\phi_x^\gamma(T_x^\gamma))$  for  $t \in (T_x^\gamma, T^\gamma]$ . The bound from Theorem 3(b) ensures that for each  $\tilde{\delta} \in (0, 1)$  there is some  $N_0(\tilde{\delta}) \in \mathbb{N}$  such that, whenever  $N \geq N_0(\tilde{\delta})$ ,

$$\begin{aligned} \mathbb{P}_x \left( \left\| \phi_x^\gamma - \hat{X}^N \right\|_{*, T_x^\gamma} < \lambda \right) &= \mathbb{P}_x \left( \left\| \phi_x^\gamma - \hat{X}^{N, \alpha, K'} \right\|_{*, T_x^\gamma} < \lambda \right) \geq \mathbb{P}_x \left( \left\| \phi_x^\gamma - \hat{X}^{N, \alpha, K'} \right\|_{*, T^\gamma} < \lambda \right) \\ &\geq \exp \left( -N \left( S(x, T^\gamma, \phi_x^\gamma) + \tilde{\delta}/4 \right) \right) = \exp \left( -N \left( S(x, T_x^\gamma, \phi_x^\gamma) + \tilde{\delta}/4 \right) \right) \\ &\geq \exp \left( -N \left( \gamma/2 + \tilde{\delta}/4 \right) \right) \end{aligned}$$

for all  $x \in N^\lambda(K_j)$ .

It follows that for each  $x \in N^\lambda(K_j)$ , if  $N \geq N_0(\gamma)$ , then

$$\begin{aligned} \mathbb{P}_x \left( \tau_{N^\lambda(K_j)}^N > T^\gamma \right) &\leq 1 - \mathbb{P}_x \left( \|\phi_x^\gamma - \hat{X}^N\|_{*,T_x^\gamma} < \lambda \right) \\ &\leq 1 - \exp(-N(\gamma/2 + \gamma/4)). \end{aligned}$$

Using the Markov property we see that, if  $N \geq N_0(\gamma)$  and  $x \in N^\lambda(K_j)$ , then

$$\begin{aligned} \mathbb{P}_x \left( \tau_{N^\lambda(K_j)}^N > \exp(N\gamma) \right) &\leq \mathbb{P}_x \left( \tau_{N^\lambda(K_j)}^N > \left\lfloor \frac{\exp(N\gamma)}{T^\gamma} \right\rfloor T^\gamma \right) \\ &\leq (1 - \exp(-3N\gamma/4))^{\left\lfloor \frac{\exp(N\gamma)}{T^\gamma} \right\rfloor}. \end{aligned}$$

We can assume without loss of generality that  $N(\gamma)$  is large enough so that

$$\left\lfloor \frac{\exp(N\gamma)}{T^\gamma} \right\rfloor > \frac{\exp(N\gamma)}{2T^\gamma}.$$

Then for all  $N \geq N_0(\gamma)$ ,

$$\begin{aligned} \sup_{x \in N^\lambda(K_j)} \mathbb{P}_x \left( \tau_{N^\lambda(K_j)}^N > \exp(N\gamma) \right) &\leq (1 - \exp(-3N\gamma/4))^{\frac{\exp(N\gamma)}{2T^\gamma}} \\ &= \exp \left( \log(1 - \exp(-3N\gamma/4)) \frac{\exp(N\gamma)}{2T^\gamma} \right) \\ &\leq \exp \left( -\frac{\exp(N\gamma)}{2T^\gamma} \exp(-3N\gamma/4) \right) \\ &= \exp \left( -\frac{\exp(N\gamma/4)}{2T^\gamma} \right). \end{aligned}$$

The result follows from taking

$$\zeta_\gamma(N) \doteq \exp \left( -\frac{\exp(N\gamma/4)}{2T^\gamma} \right).$$

□

*Proof of Theorem 1.* Recall that we assume that Assumptions 1, 2, 3, and 4 are satisfied. Also, by assumption, for every  $N \in \mathbb{N}$ , there exists a QSD  $\mu_N$  for  $\{X^N\}$ , and the sequence  $\{\mu_N\}$  is relatively compact. From Theorem 5 there are  $a_1, c_1 \in (0, \infty)$  such that

$$\lambda_N \geq 1 - a_1 e^{-c_1 N} \quad \text{for all } N \in \mathbb{N}.$$

Let  $\mu$  be a limit point of  $\mu_N$ . From Theorem 7,  $\mu$  is invariant under the flow  $\{\varphi_t\}$ . From Corollary 2,  $\text{supp}(\mu) \subset \mathcal{R}_{\text{AP}}^*$ . Thus, to finish the proof, it suffices to show that for every  $j \in \{l + 1, \dots, v\}$ , there is a neighborhood  $V_j$  of  $K_j$  such that  $\mu(V_j) = 0$ . Fix  $\varepsilon > 0$  and choose  $F_0 \in \mathcal{K}$  such that  $\mu_N(F_0^c) < \varepsilon$  for every  $N \in \mathbb{N}$ . This can be done in view of Theorem 6 and our assumption that the sequence  $\{\mu_N\}$  is relatively compact.

Using Lemma 3(c) we can assume that  $F_0$  is large enough so that for some  $T_1 \in (0, \infty)$  and  $\hat{\delta} > 0$ ,  $\varphi_t(x) \in F_1$  for all  $t \geq T$  and  $x \in F_0$ , where  $F_1 \subset F_0$  is such that  $\text{dist}(F_1, \partial F_0) > \hat{\delta}$ .

Let  $\lambda$  be as in Lemma 23. Fix  $\delta = \lambda \wedge \hat{\delta}$ . From Lemma 11, we can choose  $\delta_0 \in (0, \delta)$ , an integer  $T_0 > T_1$ , and open sets  $V_i$  with  $\bar{V}_i \subset N^{\delta_0}(K_i) \cap \Delta^o$  such that Parts 1–3 of Lemma 11 hold.

Consider

$$\beta_{\delta_0, T_0, F_0}^N \doteq \sup_{x \in \Delta_N \cap F_0} \mathbb{P}_x \left[ \|\hat{X}^N - \varphi_{\cdot}(x)\|_{*, T_0} \geq \delta_0 \right].$$

Then from Lemma 13 there exist  $c_2 > 0$  and  $a_2 \in (0, \infty)$  such that

$$\beta_{\delta_0, T_0, F_0}^N \leq a_2 e^{-Nc_2}.$$

Define  $c^* = \min\{1, c_1, c_2\}$ . With  $\gamma = c^*/8$ , let  $\zeta(N, \gamma) \doteq \zeta^*(N)$  be as in Lemma 23. Then, for some  $a_3 \in (0, \infty)$ ,

$$\sup_{x \in N^{\lambda}(K_j)} \mathbb{P}_x \left( \tau_{N^{\lambda}(K_j)}^N > \exp(Nc^*/8) \right) \leq a_3 \zeta^*(N) \quad \text{for all } N \in \mathbb{N}.$$

Define  $m_N = \exp(Nc^*/2)$  and  $m'_N = \exp(Nc^*/4)$ .

Define the events

$$\mathcal{E}_N = \left\{ (\hat{X}^N(0), \hat{X}^N(T_0), \hat{X}^N(2T_0), \dots, \hat{X}^N(m_N T_0)), (T_0, T_0, \dots, T_0), \right. \\ \left. \text{defines a } (\delta_0, T_0)\text{AP-pseudo-orbit} \right\}$$

and

$$\mathcal{E}'_N = \left\{ \text{for any } i \in \{l+1, \dots, v\} \text{ and any } q \geq m'_N, \text{ and } p \geq 0, \right. \\ \left. \text{if } \hat{X}^N(pT_0) \in N^{\delta_0}(K_i), \text{ then } \hat{X}^N((p+q)T_0) \notin N^{\delta_0}(K_i) \right\}.$$

Without loss of generality we can assume that  $m_N > (b+2)(m'_N+1)$ . Then, for  $x \in \Delta^o$ ,

$$\mathbb{P}_x(\hat{X}^N(m_N T_0) \in V_i) \leq \mathbb{P}_x(\hat{X}^N(m_N T_0) \in V_i, \mathcal{E}_N, \mathcal{E}'_N) + \mathbb{P}_x(\mathcal{E}_N, (\mathcal{E}'_N)^c) + \mathbb{P}_x((\mathcal{E}_N)^c).$$

For  $\alpha = 1, \dots, b+1$ , define  $t_\alpha^N = \lfloor \alpha m_N / (b+2) \rfloor$ . Then, from Part 2 of Corollary 11 and the definition of  $\mathcal{E}'_N$ , with  $K = \cup_{j=1}^v K_j$ ,

$$\mathbb{P}_x \left( \hat{X}^N(m_N T_0) \in V_i, \mathcal{E}_N, \mathcal{E}'_N \right) \leq \sum_{\alpha=1}^{b+1} \mathbb{P}_x \left( \hat{X}^N(t_\alpha^N T_0) \in [N^{\delta_0}(K)]^c \cap \Delta^o \right).$$

Using Part 3 of Lemma 11, for every  $x \in \Delta^o$ ,

$$\mathbb{P}_x(\mathcal{E}_N, (\mathcal{E}'_N)^c) \leq \sum_{i=l+1}^v \sup_{x \in N^{\delta_0}(K_i)} \mathbb{P}_x \left( \tau_{V_i}^N > T_0 m'_N \right) \\ \leq \sum_{i=l+1}^v \sup_{x \in N^{\delta_0}(K_i)} \mathbb{P}_x \left( \tau_{N^{\delta_0}(K_i)}^N > \exp(Nc^*/4) \right) \leq b \zeta^*(N).$$

From our choice of  $\delta_0, T_0$  we see that if for some  $k$ ,  $\hat{X}^N((k-1)T_0) \in F_0$ , then  $\varphi_{T_0}(\hat{X}^N((k-1)T_0)) \in F_1$ , and if in addition

$$\|\hat{X}^N(kT_0) - \varphi_{T_0}(\hat{X}^N((k-1)T_0))\| \leq \delta_0,$$

then  $\hat{X}^N(kT_0) \in F_0$ . Using this observation, we see that, with

$$k^* \doteq \min \{ 1 \leq k \leq m_N : \|\hat{X}^N(kT_0) - \varphi_{T_0}(\hat{X}^N((k-1)T_0))\| > \delta_0 \},$$

for every  $x \in F_0$ ,

$$\begin{aligned} \mathbb{P}_x((\mathcal{E}_N)^c) &= \mathbb{P}_x(\|\hat{X}^N(kT_0) - \varphi_{T_0}(\hat{X}^N((k-1)T_0))\| \geq \delta_0 \text{ for some } k = 1, \dots, m_N) \\ &= \mathbb{P}_x(k^* \leq m_N) \\ &\leq \sum_{k=1}^{m_N} \mathbb{P}_x(\|\hat{X}^N(kT_0) - \varphi_{T_0}(\hat{X}^N((k-1)T_0))\| \geq \delta_0, \hat{X}^N((k-1)T_0) \in F_0) \\ &\leq m_N \sup_{x \in F_0} \mathbb{P}_x(\|\hat{X}^N(T_0) - \varphi_{T_0}(\hat{X}^N(0))\| \geq \delta_0) \\ &\leq m_N \beta_{\delta_0, T_0, F_0}^N \leq a_2 \exp(Nc^*/2) \exp(-Nc^*) = a_2 \exp(-Nc^*/2). \end{aligned}$$

Thus, from our choice of  $F_0$ ,

$$\begin{aligned} \lambda_N^{m_N T_0} \mu_N(V_j) &= \int \mu_N(dx) \mathbb{P}_x(\hat{X}^N(m_N T_0) \in V_j) \\ &\leq \int \mu_N(dx) \mathbb{P}_x(\mathcal{E}_N, (\mathcal{E}'_N)^c) + \sum_{\alpha=1}^{b+1} \int \mu_N(dx) \mathbb{P}_x(\hat{X}^N(t_\alpha^N T_0) \in [N^{\delta_0}(K)]^c \cap \Delta^o) \\ &\quad + \int_{F_0} \mu_N(dx) \mathbb{P}_x((\mathcal{E}_N)^c) + \varepsilon \\ &\leq b\zeta^*(N) + (b+1)\mu_N([N^{\delta_0}(K)]^c) + a_2 \exp(-Nc^*/2) + \varepsilon. \end{aligned}$$

Note that  $\mu_N([N^{\delta_0}(K)]^c) \rightarrow 0$ , in view of Theorem 6 and Corollary 2. Since  $\lambda_N \geq 1 - a_1 e^{-c_1 N}$  and  $m_N e^{-c_1 N} \leq e^{-c_* N/2} \rightarrow 0$ ,  $\lambda_N^{m_N T_0} \rightarrow 1$ . Thus, sending  $N \rightarrow \infty$  in the above display, we have  $\mu(V_j) \leq \varepsilon$ . Since  $\varepsilon > 0$  is arbitrary, the result follows.  $\square$

### 8. Proof of Theorem 2

In this section we prove Theorem 2. For this we first show that when  $\theta^N = \theta^{N,*}$ , under the conditions of the theorem, Assumptions 1, 2, 3, and 4 are satisfied. These assumptions are verified in Sections 8.1, 8.2, 8.3, and 8.4, respectively. We then argue in Section 8.5 that, for every  $N$ ,  $X^N$  has a QSD  $\mu_N$  of the form in the statement of Theorem 2. In Section 8.6 we show that the sequence  $\{\mu_N\}$  is tight. Finally, in Section 8.7 we combine the results of previous sections to complete the proof of Theorem 2.

#### 8.1. Verification of Assumption 2.1

We need to show that if  $\theta^N = \theta^{N,*}$ , and  $x^N \rightarrow x$ , then (3) holds. The proof follows by a standard application of Grönwall’s lemma and from moment formulas of Poisson and binomial random variables, and thus we only give a sketch. First, using the relation (1) and the discrete-time Grönwall inequality it is easy to verify that for every  $T < \infty$ ,

$$\sup_{N \in \mathbb{N}} \mathbb{E}_{x^N} \max_{0 \leq k \leq \lfloor NT \rfloor} \|X_k^N\|^2 < \infty. \tag{27}$$

Next, using the relation

$$\eta_{k+1}^N(x) = G(x) + [\eta_{k+1}^N(x) - \mathbb{E}(\eta_{k+1}^N(x))], \quad x \in \Delta,$$

and the Lipschitz property of  $G$ , it can be checked that

$$\hat{X}^N(t) = x^N + \int_0^t G(\hat{X}^N(s))ds + M^N(t) + R^N(t), \quad t \in [0, T], N \in \mathbb{N}, \tag{28}$$

where  $M^N$  is a martingale and  $\sup_{0 \leq t \leq T} \|R^N(t)\|$  converges to 0 in probability as  $N \rightarrow \infty$ . Standard moment estimates show that  $\mathbb{E}(\sup_{0 \leq t \leq T} \|M^N(t)\|^2) \rightarrow 0$  as  $N \rightarrow \infty$ . Next, using the moment bound (27) and the convergence properties noted above, it can be checked that  $\hat{X}^N$  is tight in  $C([0, T]:\Delta)$ . Finally, if  $\hat{X}^N$  converges in distribution along a subsequence to  $\hat{X}$ , then from (28) it follows that  $\hat{X}$  must satisfy

$$\hat{X}(t) = x + \int_0^t G(\hat{X}(s))ds, \quad t \in [0, T].$$

From the unique solvability of the ODE in (4), which is a consequence of the Lipschitz property of  $G$ , it now follows that  $\hat{X}(t) = \varphi_t(x)$  for all  $t \in [0, T]$ , almost surely. This proves the convergence in (3). □

**8.2. Verification of Assumption 2**

Parts (a)–(d) hold by assumption. We now verify Part (e). Since  $G(x) = F(x) - x$ , for each  $x \in \Delta$   $\langle x, G(x) \rangle = \langle x, F(x) \rangle - \|x\|^2$ . As  $F$  is bounded, taking  $M \doteq 2\|F\|_\infty$ , we see that  $\|F(x)\| \leq \|x\|/2$  for all  $\|x\| \geq M$ . Thus

$$\langle x, G(x) \rangle \leq \frac{1}{2}\|x\|^2 - \|x\|^2 = -\frac{1}{2}\|x\|^2 \quad \text{for all } x \in \Delta \text{ with } \|x\| \geq M.$$

Thus Assumption 2(e) holds with  $\kappa = 1/2$  and  $M$  as above. □

**8.3. Verification of Assumption 3**

Part (a) of the assumption is immediate from the fact that for  $x \in \Delta^o$ ,  $\theta^{N,*}(\cdot | x)$  is the probability law of  $U^N - V^N$ , where  $U^N = (U_i^N)_{i=1}^d$  and  $V^N = (V_j^N)_{j=1}^d$  are  $d$ -dimensional random variables such that  $\{U_i^N, V_j^N, i, j = 1, \dots, d\}$  are mutually independent and  $U_i^N \sim \text{Poi}(F_i(x))$ ,  $V_j^N \sim \text{Bin}(Nx_j, 1/N)$ , for  $i, j = 1, \dots, d$ .

To verify Part (b), for  $x \in \Delta^o$  define  $\theta(\cdot | x)$  as the probability law of  $U - V$ , where  $U = (U_i)_{i=1}^d$  and  $V = (V_j)_{j=1}^d$  are  $d$ -dimensional random variables such that  $\{U_i, V_j, i, j = 1, \dots, d\}$  are mutually independent and  $U_i \sim \text{Poi}(F_i(x))$ ,  $V_j \sim \text{Poi}(x_j)$ , for  $i, j = 1, \dots, d$ . Then with this choice of  $\theta$ , Parts (i) and (ii) of Assumption 3(b) are clearly satisfied. Finally, Part (iii) is a consequence of the observation that if  $z_N \rightarrow z \in (0, \infty)$ , then for every  $\lambda \in \mathbb{R}$ , as  $N \rightarrow \infty$ ,

$$\left[ \left(1 - \frac{1}{N}\right) + \frac{1}{N}e^\lambda \right]^{Nz_N} \rightarrow e^{z(e^\lambda - 1)}.$$

□

**8.4. Verification of Assumption 4**

Part (a) of the assumption is clearly satisfied (in fact with  $k = 1$ ). Part (b) is verified in the following lemma.

**Lemma 24.** *Suppose that  $\theta^N = \theta^{N,*}$ . Then, for every  $\gamma \in (0, \infty)$  and  $T \in \mathbb{N}$ , there is an open neighborhood  $U_\gamma$  of  $\partial\Delta$  in  $\Delta$  such that*

$$\liminf_{N \rightarrow \infty} \inf_{x \in U_\gamma \cap \Delta_N} \frac{1}{N} \log \mathbb{P}_x(\hat{X}^N(T) \in \partial\Delta) \geq -\gamma.$$

*Proof.* For  $x \in \Delta^o$ , let

$$i_x \doteq \arg \min_{1 \leq i \leq d} x_i.$$

From Assumption 2(d) we can find  $\delta_0 > 0$  such that

$$\sup_{y \in N^{\delta_0}(\partial\Delta)} F(y)_{i_y} < \frac{\gamma}{2T}.$$

Let

$$\delta_1 \doteq \frac{\gamma}{2(T - \log(e^T - 1))},$$

$\delta \doteq \min\{\delta_0, \delta_1\}$ , and  $U_\gamma \doteq N^\delta(\partial\Delta)$ . Fix  $x \in U_\gamma \cap \Delta_N^o$ , and note that, under  $\mathbb{P}_x$ ,

$$\hat{X}^N(T) = x + \frac{1}{N} \sum_{j=1}^{NT} \eta_j^N(X_{j-1}^N),$$

where  $\eta_j^N(X_{j-1}^N) = U^j - V^j$  and the conditional distribution of  $(U^j - V^j)$  given that  $X_{j-1}^N = x$  is that of  $(U^N, V^N)$  as in Section 8.3. Thus

$$\begin{aligned} \mathbb{P}_x(\hat{X}^N(T) \in \partial\Delta) &\geq \mathbb{P}_x\left(x_{i_x} + \frac{1}{N} \sum_{j=1}^{NT} (U_{i_x}^j - V_{i_x}^j) = 0\right) \\ &\geq \mathbb{P}_x\left(U_{i_x}^1 = \dots = U_{i_x}^{NT} = 0, \sum_{j=1}^{NT} V_{i_x}^j = Nx_{i_x}\right). \end{aligned}$$

Let  $\tilde{U}^1, \dots, \tilde{U}^{NT}$  be independent and identically distributed (i.i.d.) Poisson random variables with mean  $\gamma/2T$ , and let  $\tilde{V}^1, \tilde{V}^2, \dots, \tilde{V}^{NT}$  be i.i.d. geometric random variables with probability of success  $1/N$  such that  $\{\tilde{U}^j, \tilde{V}^k, j, k\}$  are mutually independent. Then

$$\begin{aligned} &\mathbb{P}_x\left(U_{i_x}^1 = \dots = U_{i_x}^N = 0, \sum_{j=1}^N V_{i_x}^j = Nx_{i_x}\right) \\ &= \mathbb{P}_x\left(U_{i_x}^1 = \dots = U_{i_x}^N = 0, \text{ by time instant } NT \text{ all initial } Nx_{i_x} \text{ type } i_x \text{ particles die}\right) \\ &\geq \mathbb{P}_x\left(\tilde{U}^1 = \dots = \tilde{U}^{NT} = 0, \tilde{V}^1 \leq NT, \tilde{V}^2 \leq NT, \dots, \tilde{V}^{Nx_{i_x}} \leq NT\right) \\ &= [\mathbb{P}(\tilde{U}^1 = 0)]^{NT} \left(1 - \left(1 - \frac{1}{N}\right)^{NT}\right)^{Nx_{i_x}} \\ &= \exp\left(-NT \frac{\gamma}{2T}\right) \left(1 - \left(1 - \frac{1}{N}\right)^{NT}\right)^{Nx_{i_x}}. \end{aligned}$$



Combining the last two displays,

$$\begin{aligned} \frac{1}{N} \log \mathbb{P}_x(\hat{X}^N(T) \in \partial\Delta) &\geq \frac{1}{N} \left( \log \left( \exp \left( -N \frac{\gamma}{2} \right) \right) + \log \left( \left( 1 - \left( 1 - \frac{1}{N} \right)^{NT} \right)^{Nx_{i_x}} \right) \right) \\ &= -\frac{\gamma}{2} + x_{i_x} \log \left( 1 - \left( 1 - \frac{1}{N} \right)^{NT} \right) \\ &\geq -\frac{\gamma}{2} + \delta \log \left( 1 - \left( 1 - \frac{1}{N} \right)^{NT} \right), \end{aligned}$$

and thus from our choice of  $\delta$ ,

$$\liminf_{N \rightarrow \infty} \inf_{x \in U_\gamma \cap \Delta_N} \frac{1}{N} \log \mathbb{P}_x(\hat{X}^N(T) \in \partial\Delta) \geq -\frac{\gamma}{2} + \delta (-T + \log(e^T - 1)) \geq -\gamma.$$

□

### 8.5. Existence of quasi-stationary distributions

In this section we prove the existence of a QSD  $\mu_N$  for the Markov chain  $\{X_n^N\}$ , for each  $N \in \mathbb{N}$ , and show that the sequence  $\{\mu_N\}$  of QSD is relatively compact in  $\mathcal{P}(\Delta)$ . For some uniform bounds needed for the tightness proof in Section 8.6, it will be convenient to consider the  $N$ -step processes  $\{\tilde{X}_n^N\}_{n \in \mathbb{N}_0}$ , where

$$\tilde{X}_n^N \doteq X_{nN}^N, \quad n \in \mathbb{N}_0, \quad N \in \mathbb{N}. \tag{29}$$

Recall the definition of  $\tau_\delta^N$  and  $P_n^N$  from (6) and (7).

For existence of QSD, we will use the following result from [2].

**Theorem 10.** ([2, Theorem 2.1, Proposition 3.1].) *Fix  $N \in \mathbb{N}$ . Suppose that there are  $\theta_1, \theta_2, c_1 \in (0, \infty)$ , functions  $\varphi_1, \varphi_2: \Delta_N^o \rightarrow \mathbb{R}_+$ , and a measurable subset  $K \subset \Delta_N^o$  such that the following hold:*

(B1) *For each  $x \in K$ , for some  $n_2(x) \in \mathbb{N}$ ,*

$$\mathbb{P}_x(X_{n_2(x)}^N \in K) > 0 \quad \text{for all } x \in K \text{ and } n \geq n_2(x).$$

(B2) *We have  $\theta_1 < \theta_2$ , and*

- (a)  $\inf_{x \in \Delta_N^o} \varphi_1(x) \geq 1, \sup_{x \in K} \varphi_1(x) < \infty$ ;
- (b)  $\inf_{x \in K} \varphi_2(x) > 0, \sup_{x \in \Delta_N^o} \varphi_2(x) \leq 1$ ;
- (c)  $P_1^N \varphi_1(x) \leq \theta_1 \varphi_1(x) + c_1 1_K(x)$  for all  $x \in \Delta_N^o$ ;
- (d)  $P_1^N \varphi_2(x) \geq \theta_2 \varphi_2(x)$  for all  $x \in \Delta_N^o$ .

Suppose also that there exist  $C \in (0, \infty)$  and  $n_0, m_0 \in \mathbb{N}$  such that  $n_0 \leq m_0$  and

$$\mathbb{P}_x(X_{n_0}^N \in \cdot \cap K) \leq C \mathbb{P}_y(X_{m_0}^N \in \cdot) \quad \text{for all } x \in \Delta_N^o \text{ and } y \in K. \tag{30}$$

Then there exist  $C_1 \in (0, \infty)$ ,  $\alpha \in (0, 1)$ , and a probability measure  $\mu_N$  on  $\Delta_N^o$  such that for all  $n \in \mathbb{N}$ ,

$$\left\| \frac{\mu P_n^N}{\mu P_n^N(1_{\Delta_N^o})} - \mu_N \right\|_{TV} \leq C\alpha^n \frac{\mu(\varphi_1)}{\mu(\varphi_2)}$$

for all probability measures  $\mu$  on  $\Delta_N^o$  which satisfy  $\mu(\varphi_1) < \infty$  and  $\mu(\varphi_2) > 0$ . Moreover,  $\mu_N$  is the unique QSD of  $\{X^N\}$  that satisfies  $\mu_N(\varphi_1) < \infty$  and  $\mu_N(\varphi_2) > 0$ . Additionally,  $\mu_N(K) > 0$ .

**Remark 1.** The above theorem combines two different results from [2]. Proposition 3.1 of [2] shows that under the assumptions of Theorem 10, for some  $c_2 \in (0, \infty)$ ,  $n_1 \in \mathbb{N}$ , and a probability measure  $\nu$  supported on  $K$ , we have

$$\mathbb{P}_x(X_{n_1}^N \in \cdot) \geq c_2\nu(\cdot \cap K) \quad \text{for all } x \in K.$$

This proposition also shows that for some  $c_3 \in (0, \infty)$ ,

$$\sup_{y \in K} \mathbb{P}_y(n < \tau_\partial^N) \leq c_3 \cdot \sup_{n \in \mathbb{N}_0} \frac{\inf_{y \in K} \mathbb{P}_y(n < \tau_\partial^N)}{\inf_{y \in K} \mathbb{P}_y(n < \tau_\partial^N)}$$

Using these facts, it then follows that, under the assumptions of Theorem 10, all the conditions of Theorem 2.1 in [2] are satisfied, which gives the existence of QSD  $\mu_N$  with the properties stated in the above theorem.

In Lemma 11 we use the above result to establish existence of a QSD for the sequence  $\{X_n^N\}$  considered in this work, for each  $N \in \mathbb{N}$ . We begin with some preliminary estimates.

For  $r \in \mathbb{N}$ , consider

$$K_r \doteq \{x \in \Delta^o : x \cdot 1 \leq r\}, \quad K_r^N \doteq K_r \cap \Delta_N^o, \tag{31}$$

and let

$$\sigma_\partial^N \doteq \inf \{k \in \mathbb{N}_0 : \tilde{X}_k^N \in \partial \Delta_N\},$$

$$\tau_r^N \doteq \inf \{k \in \mathbb{N}_0 : X_k^N \in K_r^N\}, \quad \sigma_r^N = \inf \{k : \tilde{X}_k^N \in K_r^N\},$$

and

$$\hat{\tau}_r^N \doteq \tau_r^N \wedge \tau_\partial^N, \quad \hat{\sigma}_r^N \doteq \sigma_r^N \wedge \sigma_\partial^N.$$

**Lemma 25.** Fix  $\lambda_0 \in (0, \infty)$ . There exist  $c(\lambda_0) \in (0, \infty)$  and  $r_0 > 0$  such that for all  $r \geq r_0$  and  $\lambda \leq \lambda_0$ ,

$$\mathbb{E}_x \left( e^{\lambda \hat{\sigma}_r^N} \right) \leq e^{x \cdot 1} c(\lambda_0) \quad \text{for all } x \in \Delta_N^o \text{ and } N \in \mathbb{N}.$$

Furthermore, if  $r \geq r_0$ , then

$$\mathbb{E}_x \left( e^{\frac{\lambda}{N} \hat{\tau}_r^N} \right) \leq e^{x \cdot 1} c(\lambda_0) \quad \text{for all } x \in \Delta_N^o \text{ and } N \in \mathbb{N}.$$

*Proof.* Let  $a = \max_i \|F_i\|_\infty$ . Given  $u \in N^{-1}\mathbb{N}$ , consider the random variable  $V_u$  that represents the number of particles among  $Nu$  initial particles that die in  $N$  steps, when at each step

any particle can die independently of the remaining particles with probability  $1/N$ . Note that  $V_u \sim \text{Bin}(Nu, \gamma(N))$ , where

$$\gamma(N) = 1 - \left(1 - \frac{1}{N}\right)^N.$$

Let  $U \sim \text{Poi}(Nad)$  be independent of  $V_x$ . Then, under  $\mathbb{P}_x$ ,

$$(\tilde{X}_1^N - x) \cdot 1 \leq_d \frac{1}{N}(U - V_{x,1}),$$

where for two real random variables  $Z_1, Z_2$ , we write  $Z_1 \leq_d Z_2$  if  $\mathbb{P}(Z_2 \geq u) \geq \mathbb{P}(Z_1 \geq u)$  for all  $u \in \mathbb{R}$ . Also,

$$\mathbb{E}_x[e^{\frac{1}{N}(U - V_{x,1})}] = C_N(1)e^{-V_0^N(1)x \cdot 1},$$

where  $C_N(1) = \mathbb{E}\left(\exp\left\{\frac{1}{N}U\right\}\right)$  and

$$\mathbb{E} \exp\left\{-\frac{1}{N}V_{x,1}\right\} = e^{-V_0^N(1)x \cdot 1}.$$

Note that for  $x \in (K_r^N \cup \partial \Delta_N)^c$ ,

$$\begin{aligned} \mathbb{P}_x(\hat{\sigma}_r^N > 1) &\leq \mathbb{E}_x\left(e^{\tilde{X}_1^N \cdot 1} 1_{\hat{\sigma}_r^N > 1}\right) = e^{x \cdot 1} \mathbb{E}_x\left(e^{(\tilde{X}_1^N - x) \cdot 1} 1_{\hat{\sigma}_r^N > 1}\right) \\ &= e^{x \cdot 1} C_N(1)e^{-V_0^N(1)x \cdot 1} \leq e^{x \cdot 1} C_N(1)e^{-V_0^N(1)r}. \end{aligned}$$

By a recursive argument, for  $n \in \mathbb{N}$ ,

$$\mathbb{P}_x(\hat{\sigma}_r^N > n) \leq e^{x \cdot 1} e^{-n(rV_0^N(1) - \log C_N(1))}. \tag{32}$$

Note that

$$\log C_N(1) = Nad(e^{1/N} - 1).$$

Also,

$$\mathbb{E} \exp\left\{-\frac{1}{N}V_{x,1}\right\} = \left[1 - \gamma(N)(1 - e^{-1/N})\right]^{N(x \cdot 1)},$$

and thus

$$V_0^N(1) = -N \log \left[1 - \gamma(N)(1 - e^{-1/N})\right].$$

Combining the above observations, we have

$$\frac{\log C_N(1)}{V_0^N(1)} = \frac{Nad(e^{1/N} - 1)}{-N \log \left[1 - \gamma(N)(1 - e^{-1/N})\right]} \leq \frac{ad(e^{1/N} - 1)}{\gamma(N)(1 - e^{-1/N})} = ad \frac{e^{1/N}}{\gamma(N)}.$$

Since  $\gamma(N) \rightarrow (1 - e^{-1})$ , we can assume without loss of generality that for all  $N \in \mathbb{N}$ ,

$$\frac{\log C_N(1)}{V_0^N(1)} \leq \frac{2ade^2}{e - 1} \doteq \vartheta, \quad V_0^N(1) \geq \frac{1}{2}(1 - e^{-1}) \doteq \varsigma.$$

Thus, for  $r \geq r_0 \doteq \left(\frac{\lambda}{\varsigma} + \vartheta\right)$ ,

$$e^{-n(rV_0^N(1) - \log C_N(1))} \leq e^{-nV_0^N(1)\left(\frac{\lambda}{\varsigma} + \vartheta - \frac{\log C_N(1)}{V_0^N(1)}\right)} \leq e^{-nV_0^N(1)\frac{\lambda}{\varsigma}} \leq e^{-n\lambda}.$$

Combining this with (32), for all  $N \in \mathbb{N}$ ,  $x \in (K_r^N \cup \partial)^c$  and  $\lambda_0 < \lambda$ ,

$$\mathbb{E}_x \left( e^{\lambda_0 \hat{\sigma}_r^N} \right) \leq e^{x \cdot 1} \frac{e^{\lambda_0} - e^{(\lambda_0 - \lambda)}}{1 - e^{(\lambda_0 - \lambda)}}.$$

This proves the first statement in the lemma. The second statement follows on noting that  $\hat{\tau}_r^N \leq N \hat{\sigma}_r^N$  for each  $r \in \mathbb{R}_+$  and  $N \in \mathbb{N}$ . □

**Lemma 26.** Fix  $\lambda_0 \in (0, \infty)$  and let  $r_0$  be as in Lemma 25. Then for each  $\lambda \in (0, \lambda_0)$  and  $r \geq r_0$ ,

$$\sup_{N \in \mathbb{N}} \sup_{y \in K_r^N} \mathbb{E}_y \left( \mathbb{E}_{\tilde{X}_1^N} \left( e^{\lambda \hat{\sigma}_r^N} \right) 1_{1 < \sigma_\theta^N} \right) < \infty.$$

Furthermore, for every  $N \in \mathbb{N}$ ,

$$\sup_{y \in K_r^N} \mathbb{E}_y \left( \mathbb{E}_{X_1^N} \left( e^{\frac{\lambda}{N} \hat{\tau}_r^N} \right) 1_{1 < \tau_\theta^N} \right) < \infty.$$

*Proof.* We only prove the first statement. The second statement is shown in a similar manner. Fix  $r \geq r_0$  and  $\lambda < \lambda_0$ . For notational simplicity, denote  $K_r^N$  by  $K$ . Then, for  $y \in K$ ,

$$\begin{aligned} \mathbb{E}_y \left( \mathbb{E}_{\tilde{X}_1^N} \left( e^{\lambda \hat{\sigma}_r^N} \right) 1_{1 < \sigma_\theta^N} \right) &= \mathbb{E}_y \left( \mathbb{E}_{\tilde{X}_1^N} \left( e^{\lambda \hat{\sigma}_r^N} \right) 1_{1 < \sigma_\theta^N} \left( 1_{\tilde{X}_1^N \in K} + 1_{\tilde{X}_1^N \in K^c} \right) \right) \\ &\leq 1 + \mathbb{E}_y \left( \mathbb{E}_{\tilde{X}_1^N} \left( e^{\lambda \hat{\sigma}_r^N} \right) 1_{1 < \sigma_\theta^N} 1_{\tilde{X}_1^N \in K^c} \right). \end{aligned}$$

Let  $a \doteq \max_i \|F_i\|_\infty$  and  $U \sim \text{Poi}(Na)$ . Then with  $c(\lambda_0)$  as in Lemma 27 we have

$$\begin{aligned} \mathbb{E}_y \left( \mathbb{E}_{\tilde{X}_1^N} \left( e^{\lambda \hat{\sigma}_r^N} \right) 1_{1 < \sigma_\theta^N} 1_{\tilde{X}_1^N \in K^c} \right) &\leq \mathbb{E}_y \left( c(\lambda_0) e^{\tilde{X}_1^N \cdot 1} 1_{1 < \sigma_\theta^N} 1_{\tilde{X}_1^N \in K^c} \right) \\ &\leq c(\lambda_0) \mathbb{E}_y \left( e^{y \cdot 1 + \frac{1}{N} U} \right) \\ &= e^{y \cdot 1} c(\lambda_0) e^{dNa} (e^{\frac{1}{N}} - 1). \end{aligned}$$

Since  $\sup_{N \in \mathbb{N}} N(e^{1/N} - 1) \leq e$ , the result follows.

The following lemma will be used to verify the condition (B2)(d) of Theorem 10.

**Lemma 27.** There exists  $r_1 \in (0, \infty)$  such that

$$\theta_2 \doteq \inf_{r \geq r_1} \inf_{N \in \mathbb{N}} \inf_{x \in K_r^N} \mathbb{P}_x \left( \tilde{X}_1^N \in K_r^N; \sigma_\theta^N > 1 \right) > 0.$$

Furthermore, for each  $N \in \mathbb{N}$ , there exists  $r_1 \in (0, \infty)$  such that

$$\theta_2(N) \doteq \inf_{r \geq r_1} \inf_{x \in K_r^N} \mathbb{P}_x \left( X_1^N \in K_r^N; \tau_\theta^N > 1 \right) > 0.$$

*Proof.* Once again, we only prove the first statement. Consider, for  $z \in \mathbb{N}/N$ , a collection of  $Nz$  particles of a single type, where each particle, independently of all other particles, has a  $1/N$  chance of dying at each time step. Then the probability that all  $Nz$  particles are dead in  $N$  time steps is

$$p(z, N) \doteq \left( 1 - \left( 1 - \frac{1}{N} \right)^N \right)^{Nz}.$$

Note that for any  $K \subset \Delta^o$ ,  $x \in K \cap \Delta_N$ , and  $N \geq 1$ ,  $\min_{1 \leq i \leq d} x_i \geq \frac{1}{N}$ , and so

$$p(x \cdot 1, N) \leq p(dN^{-1}, N) = \left[ 1 - \left( 1 - \frac{1}{N} \right)^N \right]^d.$$

In particular,

$$\mathbb{P}_x (\sigma_\partial^N \leq 1) = \mathbb{P}_x (\tau_\partial^N \leq N) \leq \left[ 1 - \left( 1 - \frac{1}{N} \right)^N \right]^d,$$

and so for any  $K \subset \Delta^o$ ,

$$\sup_{N > 1} \sup_{x \in K} \mathbb{P}_x (\tau_\partial^N \leq N) \leq (1 - e^{-2})^d \doteq \alpha_0.$$

Thus, for  $K \subset \Delta^o$ ,

$$\mathbb{P}_x (\tilde{X}_1^N \in K | \sigma_\partial^N > 1) = 1 - \mathbb{P}_x (\tilde{X}_1^N \in K^c | \sigma_\partial^N > 1) = 1 - \frac{\mathbb{P}_x (\tilde{X}_1^N \in K^c; \sigma_\partial^N > 1)}{\mathbb{P}_x (\sigma_\partial^N > 1)},$$

and

$$\sup_{N > 1} \sup_{x \in K} \frac{\mathbb{P}_x (\tilde{X}_1^N \in K^c; \sigma_\partial^N > 1)}{\mathbb{P}_x (\sigma_\partial^N > 1)} \leq \frac{\sup_{N > 1} \sup_{x \in K} \mathbb{P}_x (\tilde{X}_1^N \in K^c; \sigma_\partial^N > 1)}{1 - \alpha_0}.$$

We will now argue that for some  $r_1 \in (0, \infty)$ ,

$$\sup_{r \geq r_1} \sup_{N > 1} \sup_{x \in K_r^N} \mathbb{P}_x (\tilde{X}_1^N \in (K_r^N)^c; \sigma_\partial^N > 1) < 1 - \alpha_0. \tag{33}$$

Fix  $r > 0$  and let  $x \in K_r^N$ . As before, let  $a = \max_i \|F_i\|_\infty$ . Fix  $k \in \mathbb{N}$  and define, for  $a_1 \in (0, \infty)$ ,

$$m = m(N, k, a_1) \doteq \max \left\{ 1 \leq j \leq k : X_j^N \cdot 1 \leq a_1 \right\}.$$

Let  $Y_k^N \doteq X_k^N \cdot 1$ . Then

$$Y_k^N = Y_m^N + \frac{1}{N} \sum_{j=m+1}^k \eta_j^N (X_k^N) \cdot 1 \leq a_1 + \frac{1}{N} \max_{\{1 \leq l \leq k\}} \sum_{j=l}^k (U_j - V_j),$$

where  $U_j$  are i.i.d.  $\text{Poi}(ad)$ ,  $V_j$  are i.i.d.  $\text{Bin}(Na_1, 1/N)$ , and  $\{U_j, V_j, j, j' \in \mathbb{N}\}$  are mutually independent. For  $a_2 > a_1$ ,

$$\begin{aligned} \mathbb{P}_x (Y_k^N \geq a_2) &\leq \mathbb{P} \left( \max_{\{1 \leq l \leq k\}} \sum_{j=l}^k (U_j - V_j) \geq N (a_2 - a_1) \right) \\ &\leq \sum_{l=1}^k \mathbb{P} \left( \sum_{j=l}^k (U_j - V_j) \geq N (a_2 - a_1) \right). \end{aligned}$$

Thus, for each  $\gamma > 0$ , by Markov's inequality,

$$\begin{aligned} \mathbb{P}_x(Y_k^N \geq a_2) &\leq e^{-\gamma N(a_2 - a_1)} \sum_{l=1}^k [\mathbb{E}e^{\gamma U_1}]^{(k-l+1)} [\mathbb{E}e^{-\gamma V_1}]^{(k-l+1)} \\ &= e^{-\gamma N(a_2 - a_1)} \frac{\mathbb{E}e^{\gamma U_1} \mathbb{E}e^{-\gamma V_1} \left(1 - (\mathbb{E}e^{\gamma U_1} \mathbb{E}e^{-\gamma V_1})^k\right)}{1 - \mathbb{E}e^{\gamma U_1} \mathbb{E}e^{-\gamma V_1}}. \end{aligned}$$

Note that for each  $N \geq 1$ ,

$$(\mathbb{E}e^{\gamma U_1} \mathbb{E}e^{-\gamma V_1}) = e^{ad(e^\gamma - 1)} \left(1 - \frac{1}{N} + \frac{1}{Ne^\gamma}\right)^{Na_1} \leq e^{ad(e^\gamma - 1)} e^{-a_1(1 - e^{-\gamma})}.$$

Let  $r_1$  be large enough so that  $e^{ad(e^\gamma - 1)} e^{r_1(e^{-\gamma} - 1)/2} < \frac{1}{2}$  and

$$r_1 > -2 \frac{\log(1 - \alpha_0)}{\gamma}.$$

If we fix  $r \geq r_1$  and let  $a_2 = r$  and  $a_1 = r/2$ , then

$$\mathbb{P}_x(Y_k^N \geq r) \leq e^{-\gamma N \frac{r}{2}} \leq e^{-\gamma N \frac{r_1}{2}} < (1 - \alpha_0)^N \leq (1 - \alpha_0).$$

This proves (33) and hence

$$\inf_{r \geq r_1} \inf_{N > 1} \inf_{x \in K_r^N} \mathbb{P}_x(\tilde{X}_1^N \in K_r^N \mid \sigma_\partial^N > 1) \doteq c_0 > 0.$$

Finally, for all  $N > 1$ ,  $r \geq r_1$ , and  $x \in K_r^N$ ,

$$\mathbb{P}_x(\tilde{X}_1^N \in K_r^N; \sigma_\partial^N > 1) = \mathbb{P}_x(\tilde{X}_1^N \in K_r^N \mid \sigma_\partial^N > 1) \mathbb{P}_x(\sigma_\partial^N > 1) \geq c_0 (1 - \alpha_0) > 0.$$

The result follows.

Denote by  $\mathcal{Q}^N$  the collection of all  $\mu \in \mathcal{P}(\Delta_N^o)$  such that for every  $c \in (0, \infty)$ , there exists an  $r \in (0, \infty)$  such that

$$\mathbb{E}_\mu \left( e^{c \hat{\sigma}_r^N} \right) < \infty.$$

The following result gives the existence of QSD for the chain  $X^N$  for each  $N$  and provides an important characterization of these QSD.

**Theorem 11.** *There is a probability measure  $\mu_N$  on  $\Delta_N^o$  such that for all  $x_N \in \Delta_N^o$ ,*

$$\frac{\delta_{x_N} P_n^N}{\delta_{x_N} P_n^N \left( 1_{\Delta_N^o} \right)} \rightarrow \mu_N$$

*in the total variation distance. For each  $N \in \mathbb{N}$ , the measure  $\mu_N$  is a QSD for  $\{X_n^N\}$ . It is the unique QSD for  $\{X_n^N\}$  that belongs to  $\mathcal{Q}^N$ .*

*Proof.* Fix  $N \in \mathbb{N}$  and let  $r_1 \in (0, \infty)$  and  $\theta_2 \in (0, 1]$  be as in the second statement in Lemma 27. Fix  $r_2 \geq r_1$ , let  $K = K_{r_2}^N$ , and define  $\varphi_2: \Delta_N^o \rightarrow \mathbb{R}_+$  by  $\varphi_2(x) \doteq 1_K(x)$ .

Fix an arbitrary  $\theta_1 \in (0, \theta_2)$ . From Lemma 25 there is an  $r_3 > r_2$  such that for any fixed  $r \geq r_3$ ,

$$\varphi_1(x) \doteq \mathbb{E}_x \left( \theta_1^{-\hat{\tau}_r^N} \right) < \infty \quad \text{for all } x \in \Delta_N^o. \tag{34}$$

We now verify the conditions of Theorem 10 with the above choice of  $K$ ,  $\varphi_1$ ,  $\varphi_2$ ,  $\theta_1$ , and  $\theta_2$ . It is clear that the condition (B1) is satisfied with  $n_2(x) = 1$ . Also, (B2)(b) is satisfied, since  $\varphi_2(x) = 1$  for each  $x \in K$ . Since  $\theta_1 \in (0, 1)$ ,

$$\inf_{x \in \Delta_N^o} \varphi_1(x) \geq 1.$$

Also, since  $K \subset K_r^N$ ,

$$\sup_{x \in K} \varphi_1(x) = 1,$$

and so (B2)(a) holds. Next, an application of Lemma 26 and the Markov property show that (B2)(c) holds with

$$c_2 \doteq \sup_{y \in K} \mathbb{E}_y \left( \varphi_1 \left( X_1^N \right) 1_{\{\sigma_\theta^N > 1\}} \right).$$

Finally, the validity of (B2)(d) follows from Lemma 27.

Also, since

$$\inf_{x, y \in K} \mathbb{P}_y \left( X_1^N = x \right) \doteq \kappa_1 > 0,$$

the inequality in (30) is satisfied with  $C = \kappa_1^{-1}$ . Thus, from Theorem 10 it follows that there exists a QSD  $\mu_N$  for  $\{X_n^N\}$  that satisfies

$$\mathbb{E}_{\mu_N} \left( \theta_1^{-\hat{\tau}_r^N} \right) < \infty, \quad \text{and } \mu_N = \lim_{n \rightarrow \infty} \frac{\delta_{x_N} P_n^N}{\delta_{x_N} P_n^N (1_{\Delta_N^o})}, \quad \text{for any } x_N \in K. \tag{35}$$

We now show that  $\mu_N \in \mathcal{Q}^N$ . Fix  $c \in (0, \infty)$ . Let  $\varphi_2$ ,  $\theta_2$ , and  $K$  be as above. Choose  $\theta_1^* \in (0, \theta_2 \wedge e^{-c})$ . From the second statement in Lemma 25, there exists an  $r_4 > r_3$  such that

$$\tilde{\varphi}_1(x) \doteq \mathbb{E}_x \left( (\theta_1^*)^{-\hat{\tau}_{r_4}^N} \right) < \infty \quad \text{for all } x \in \Delta_N^o.$$

Then from the previous argument, there is a QSD  $\tilde{\mu}_N$  for  $\{X_n^N\}$  such that

$$\mathbb{E}_{\tilde{\mu}_N} \left( e^{c\hat{\tau}_{r_4}^N} \right) \leq \mathbb{E}_{\tilde{\mu}_N} \left( (\theta_1^*)^{-\hat{\tau}_{r_4}^N} \right) < \infty$$

and

$$\tilde{\mu}_N = \lim_{n \rightarrow \infty} \frac{\delta_{x_N} P_n^N}{\delta_{x_N} P_n^N (1_{\Delta_N^o})} \quad \text{for any } x_N \in K.$$

From (35) we now see that  $\mu_N = \tilde{\mu}_N$  and that

$$\mathbb{E}_{\mu_N} \left( e^{c\hat{\tau}_{r_4}^N} \right) < \infty.$$

Since  $c > 0$  is arbitrary, it follows that  $\mu_N \in \mathcal{Q}^N$ . Also, since  $r_2 \geq r_1$  is arbitrary, we see (by choosing a larger  $K$  if needed) that the convergence in (35) holds for all  $x_N \in \Delta_N^o$ .

Finally we argue uniqueness. Let  $\tilde{\mu}_N \in \mathcal{Q}^N$  be a QSD for  $\{X_n^N\}$ . Choose  $r_5 \geq r_1$  such that  $\tilde{\mu}_N(K_{r_5}^N) > 0$ . Consider  $\tilde{K} = K_{r_5}^N$  and  $\tilde{\varphi}_2 = 1_{\tilde{K}}$ . Fix  $\theta_1 \in (0, \theta_2)$  and let  $r > r_5$  be such that

$$\mathbb{E}_{\tilde{\mu}_N} \left( (\theta_1)^{-\hat{\tau}_r^N} \right) < \infty, \quad \text{and} \quad \mathbb{E}_x \left( (\theta_1)^{-\hat{\tau}_r^N} \right) < \infty \quad \text{for all } x \in \Delta_N^o.$$

Then by the previous argument (and Theorem 10),

$$\mathbb{E}_{\mu_N} \left( (\theta_1)^{-\hat{\tau}_r^N} \right) < \infty \quad \text{and} \quad \mu_N(\tilde{K}) > 0.$$

But since the above two properties are also satisfied by  $\tilde{\mu}_N$ , from Theorem 10 we must have  $\mu_N = \tilde{\mu}_N$ . □

### 8.6. Tightness of quasi-stationary distributions

We now prove the tightness of the sequence of QSD  $\{\mu_N\}$  given in Theorem 11.

**Theorem 12.** *For  $N \in \mathbb{N}$ , let  $\mu_N$  be as given in Lemma 11. Then the sequence  $\{\mu_N\}$  is tight.*

*Proof.* Recall the definition of  $P_n^N$  from (7), and let  $\tilde{P}_n^N \doteq P_{nN}^N$ . From Lemma 11, for all  $x_N \in \Delta_N^o$ ,

$$\lim_{n \rightarrow \infty} \frac{\delta_{x_N} \tilde{P}_n^N}{\delta_{x_N} \tilde{P}_n^N(1_{\Delta^o})} = \mu_N.$$

Thus, to show that the sequence  $\{\mu_N\}$  is tight, it suffices to show that the collection

$$\left\{ \frac{\delta_{x_N} \tilde{P}_n^N}{\delta_{x_N} \tilde{P}_n^N(1_{\Delta^o})}, n, N \in \mathbb{N} \right\} \tag{36}$$

is tight for some sequence  $\{x_N\}$ , where  $x_N \in \Delta_N^o$  for each  $N$ . For this it suffices to show that for every  $\varepsilon > 0$ , there is an  $L_1 \in (0, \infty)$  such that

$$\sup_{N \in \mathbb{N}} \sup_{n \in \mathbb{N}} \mathbb{P}_{x_N} \left( \tilde{X}_n^N \cdot 1 \geq L_1 \mid \sigma_\partial^N > n \right) \leq \varepsilon.$$

From Lemma 27, for all  $r \geq r_1$ ,

$$\theta_2^r \doteq \inf_{N \geq 1} \inf_{x \in K_r} \mathbb{P}_x \left( \tilde{X}_1^N \in K_r; \sigma_\partial^N > N \right) \geq \theta_2 > 0,$$

so for every  $r \geq r_1$ , with  $\varphi_2^r(x) = \varphi_2(x) \doteq 1_{K_r}(x)$ , for each  $N \in \mathbb{N}$ ,

$$P_1^N \varphi_2^r(x) \geq \theta_2 \varphi_2^r(x) \quad \text{for all } x \in \Delta_N^o.$$

Recall that  $a = \max_{1 \leq i \leq d} \|F_i\|_\infty < \infty$  and  $\tilde{X}_k^N = X_{Nk}^N$  for  $k \in \mathbb{N}$ . We now consider a coupling between the sequence of  $d$ -dimensional random variables  $\{X_k^N\}$  and a sequence  $\{Z_k^N\}$  of  $\mathbb{N}/N$ -valued random variables that preserves certain monotonicity properties. Note that  $\{X_k^N\}$  can be constructed as follows. Consider a collection of i.i.d. random fields  $\{(U_k^N(x), V_k^N(x)), x \in \Delta_N^o\}_{k \in \mathbb{N}}$  where  $U_k^N(x)$  is a  $d$ -dimensional random variable with mutually independent coordinates distributed as Poisson random variables with means  $F_i^N(x)$ ,  $i \in \{1, \dots, d\}$ , and  $V_k^N(x)$  is



a  $d$ -dimensional random variable, independent of  $U_k^N(x)$ , of mutually independent binomial random variables with parameters  $(Nx_i, 1/N)$ ,  $i \in \{1, \dots, d\}$ . Then

$$\begin{aligned} X_{k+1}^N &= X_k^N + \frac{1}{N} (U_{k+1}^N(X_k^N) - V_k^N(X_k^N)), & k \in \mathbb{N}_0, \\ X_0^N &= x_N, \end{aligned} \tag{37}$$

gives a construction for the Markov chain  $\{X_k^n\}$ . We can then construct, along with the above i.i.d. random fields, i.i.d. fields  $\left\{ \left( A_k^N(z), B_k^N(z) \right); z \in \mathbb{N}/N \right\}_{k \in \mathbb{N}}$  such that

$$A_k^N(x \cdot 1) \sim \text{Poi}(ad - F(x) \cdot 1), \quad \text{and } D_k^N(x) \doteq A_k^N(x \cdot 1) + U_k^N(x) \cdot 1 \sim \text{Poi}(ad), \quad \text{for all } x \in \Delta_N^o,$$

and

$$\begin{aligned} B_k^N(z) &\sim \text{Bin}(Nz, 1/N), \\ \text{and whenever } z &\geq x \cdot 1, & (B_k^N(z) - V_k^N(x) \cdot 1) \leq z - x \cdot 1, & \text{for } x \in \Delta_N^o \text{ and } z \in \mathbb{N}/N. \end{aligned}$$

For  $z_N \in \mathbb{N}/N$  with  $z_N \geq x_N \cdot 1$ , define

$$Z_{k+1}^N = Z_k^N + \frac{1}{N} [D_k^N(X_k^N) - B_k^N(Z_k^N)], \quad Z_0^N = z_N.$$

The sequence  $Z_k^N$  describes the evolution of the (scaled) population size of a single-type population in which at each time step any particle can die with probability  $1/N$  independently of other particles, and  $\text{Poi}(ad)$  new particles are born. Let  $Y_k^N \doteq X_k^N \cdot 1$ . Then, by construction,  $Z_k^N \geq Y_k^N$  for all  $k, N$ .

Fix  $r \geq r_1$  and let  $x_N$  be in  $K_r \cap \Delta_N$  for each  $N$ . Also, let  $z_N = x_N \cdot 1$ . To prove the tightness of the collection in (36) it suffices to show that for every  $\varepsilon > 0$ , there is an  $L_1 \in (0, \infty)$  such that

$$\sup_{N \in \mathbb{N}} \sup_{n \in \mathbb{N}} \mathbb{P}_{x_N} (\tilde{X}_n^N \cdot 1 \geq L_1 \mid \sigma_\partial^N > n) \leq \varepsilon.$$

Let  $\tilde{Z}_n^N \doteq Z_{nN}^N$  for  $n \in \mathbb{N}_0, N \in \mathbb{N}$ , and define

$$\sigma_r^{N,Z} \doteq \inf \{ n \in \mathbb{N}_0; \tilde{Z}_n^N \leq r \}, \quad \sigma_\partial^{Z,N} \doteq \inf \{ n; \tilde{Z}_n^N = 0 \}.$$

Using similar arguments as in the proofs of Lemmas 25 and 26, we can assume without loss of generality that  $r$  is large enough so that there is a  $\theta_1 \in (0, \theta_2)$  such that for

$$\varphi_1^N(z) \doteq \mathbb{E}_z \left( \theta_1^{-\left( \sigma_r^{N,Z} \wedge \sigma_\partial^{Z,N} \right)} \right), \quad z \in \mathbb{N}/N,$$

and

$$C \doteq \sup_{N \geq 1} \sup_{y \in \mathbb{N}/N, y \leq r} \mathbb{E}_y \left( \mathbb{E}_{\tilde{Z}_1^N} \left( \theta_1^{-\left( \sigma_r^{N,Z} \wedge \sigma_\partial^{Z,N} \right)} \mathbf{1}_{1 < \sigma_\partial^{Z,N}} \right) \right),$$

we have  $C < \infty$  and

$$\mathbb{E}_z \left( \varphi_1^N(\tilde{Z}_1^N) \mathbf{1}_{\sigma_\partial^{Z,N} > 1} \right) \leq \theta_1 \varphi_1^N(z) + C \mathbf{1}_B(z), \quad z \in \mathbb{N}/N, N \in \mathbb{N}.$$

For fixed  $L < \infty$ , there is an  $L_1 \in (r_0, \infty)$  such that for all  $z \geq L_1$ , we have  $\varphi_1^N(z) \geq L$  for all  $N \in \mathbb{N}$ . Then, with  $\varphi_2 = \varphi_2^{r_0}$ ,

$$\begin{aligned} \mathbb{P}_{x_N}(\tilde{X}_n^N \cdot 1 \geq L_1 \mid \sigma_\partial^N > n) &\leq \mathbb{P}_{z^N}(\tilde{Z}_n^N \geq L_1 \mid \sigma_\partial^N > n) \leq \mathbb{P}(\varphi_1^N(\tilde{Z}_n^N) \geq L \mid \sigma_\partial^N > n) \\ &\leq L^{-1} \mathbb{E}(\varphi_1^N(\tilde{Z}_n^N) \mid \sigma_\partial^N > n) = L^{-1} \frac{\mathbb{E}(\varphi_1^N(\tilde{Z}_n^N) 1_{\sigma_\partial^N > n})}{\mathbb{P}(\sigma_\partial^N > n)} \\ &\leq L^{-1} \frac{\mathbb{E}(\varphi_1^N(\tilde{Z}_n^N) 1_{\sigma_\partial^N > n})}{\mathbb{E}(\varphi_2(\tilde{X}_n^N) 1_{\sigma_\partial^N > n})}, \end{aligned}$$

where the last inequality uses the property  $\varphi_2 \leq 1$ . Also,

$$\mathbb{E}_{x_N}(\varphi_2(\tilde{X}_n^N) 1_{\sigma_\partial^N > n}) \geq \theta_2 \mathbb{E}_{x_N}(\varphi_2(\tilde{X}_{n-1}^N) 1_{\sigma_\partial^N > n-1}) = \theta_2 \mathbb{E}_{x_N}(1_{[0,r]}(\tilde{X}_{n-1}^N \cdot 1) 1_{\sigma_\partial^N > n-1}),$$

and, with  $\mathcal{F}_n = \sigma\{\tilde{X}_k, \tilde{Z}_k, k \leq n\}$ ,

$$\begin{aligned} \mathbb{E}_{x_N}(\varphi_1^N(\tilde{Z}_n^N) 1_{\sigma_\partial^N > n}) &= \mathbb{E}_{x_N}(\mathbb{E}(\varphi_1^N(\tilde{Z}_n^N) 1_{\sigma_\partial^N > n} 1_{\sigma_\partial^N > n-1} \mid \mathcal{F}_{n-1})) \\ &= \mathbb{E}_{x_N}(\mathbb{E}_{x_N}(\varphi_1^N(\tilde{Z}_n^N) 1_{\sigma_\partial^N > n} \mid \mathcal{F}_{n-1}) 1_{\sigma_\partial^N > n-1}) \\ &\leq \mathbb{E}_{x_N}(\mathbb{E}_{x_N}(\varphi_1^N(\tilde{Z}_n^N) 1_{\sigma_\partial^N > n} \mid \mathcal{F}_{n-1}) 1_{\sigma_\partial^N > n-1}) \\ &\leq \theta_1 \mathbb{E}_{x_N}(\varphi_1^N(\tilde{Z}_{n-1}^N) 1_{\sigma_\partial^N > n-1}) + C \mathbb{E}_{x_N}(1_{[0,r]}(\tilde{Z}_{n-1}^N) 1_{\sigma_\partial^N > n-1}) \\ &\leq \theta_1 \mathbb{E}_{x_N}(\varphi_1^N(\tilde{Z}_{n-1}^N) 1_{\sigma_\partial^N > n-1}) + C \mathbb{E}_{x_N}(1_{[0,r]}(\tilde{X}_{n-1}^N \cdot 1) 1_{\sigma_\partial^N > n-1}). \end{aligned}$$

Thus,

$$\frac{\mathbb{E}_{x_N}(\varphi_1^N(Z_n^N) 1_{\sigma_\partial^N > n})}{\mathbb{E}_{x_N}(\varphi_2(\tilde{X}_n^N) 1_{\sigma_\partial^N > n})} \leq \frac{\theta_1 \mathbb{E}_{x_N}(\varphi_1^N(Z_{n-1}^N) 1_{\sigma_\partial^N > n-1})}{\theta_2 \mathbb{E}_{x_N}(\varphi_2(\tilde{X}_{n-1}^N) 1_{\sigma_\partial^N > n-1})} + \frac{C}{\theta_2}.$$

Iterating this inequality, we have

$$\frac{\mathbb{E}_{x_N}(\varphi_1^N(Z_n^N) 1_{\sigma_\partial^N > n})}{\mathbb{E}_{x_N}(\varphi_2(\tilde{X}_n^N) 1_{\sigma_\partial^N > n})} \leq \left(\frac{\theta_1}{\theta_2}\right)^n \frac{\varphi_1^N(z_N)}{\varphi_2(x_N)} + \frac{C}{\theta_2} \frac{1}{1 - (\theta_1/\theta_2)}.$$

Since  $x_N \in K_r$  for each  $N$ ,

$$\mathbb{P}_{x_N}(\tilde{X}_n^N \cdot 1 \geq L_1 \mid \sigma_\partial^N > n) \leq L^{-1} \left[1 + \frac{C}{\theta_2 - \theta_1}\right].$$

Tightness follows. □

### 8.7. Completing the proof of Theorem 2

We can now complete the proof of Theorem 2. We will apply Theorem 1. From Sections 8.1, 8.2, 8.3, and 8.4 it follows that Assumptions 1, 2, 3, and 4 are satisfied. From Section 8.5 it follows that there is a  $\mu_N \in \mathcal{P}(\Delta_N^0)$  such that for every  $N \in \mathbb{N}$  and  $x_N \in \Delta_N^0$ ,

$$\frac{\delta_{x_N} P_n^N}{\delta_{x_N} P_n^N(1_{\Delta_N^0})}$$

converges to  $\mu_N$  in the total variation distance as  $n \rightarrow \infty$ . Furthermore, the measure  $\mu_N$  is a QSD for  $\{X^N\}$ . From Section 8.5 the sequence  $\{\mu_N\}_{N \in \mathbb{N}}$  is relatively compact as a sequence of probability measures on  $\Delta$ . Theorem 2 is now immediate from Theorem 1.

### Funding

The research of A. B. was supported in part by the NSF (DMS-1814894, DMS-1853968).

### Competing Interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

### References

- [1] BUCKLEY, F. AND POLLET, P. (2010). Limit theorems for discrete-time metapopulation models. *Prob. Surv.* **7**, 53–83.
- [2] CHAMPAGNAT, N. AND VILLEMONAIS, D. (2017). *General criteria for the study of quasi-stationarity*. Preprint. Available at <https://arxiv.org/abs/1712.08092>.
- [3] DUPUIS, P. AND ELLIS, R. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. John Wiley, New York.
- [4] FAURE, M. AND SCHREIBER, S. (2014). Quasi-stationary distributions for randomly perturbed dynamical systems. *Ann. Appl. Prob.* **24**, 553–598.
- [5] FERRARI, P., KESTEN, A., MARTÍNEZ, H. AND PICCO, P. (1995). Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Prob.* **23**, 501–521.
- [6] FREIDLIN, M. AND WENTZELL, A. (1984). *Random Perturbations of Dynamical Systems*. Springer, New York.
- [7] GOSSELIN, F. (1996). The expected extinction time of a population within a system of interacting biological populations. *Bull. Math. Biol.* **58**, 555–568.
- [8] GOSSELIN, F. (2001). Asymptotic behavior of absorbing Markov chains conditional on non-absorption for applications in conservation biology. *Ann. Appl. Prob.* **11**, 261–284.
- [9] GYLLENBERG, M. AND SILVESTROV, D. (1994). Quasi-stationary distributions of a stochastic metapopulation model. *J. Math. Biol.* **33**, 35–70.
- [10] HASSELBLATT, B. AND KATOK, A. (1995). *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press.
- [11] HÖGNÄS, G. (1997). On the quasi-stationary distribution of a stochastic Ricker model. *Stoch. Process. Appl.* **70**, 243–263.
- [12] KIFER, Y. (1988). *Random Perturbations of Dynamical Systems*. Birkhäuser, Basel.
- [13] MARMET, B. (2013). Quasi-stationary distributions for stochastic approximation algorithms with constant step size. Preprint. Available at <https://arxiv.org/abs/1303.7081>.
- [14] MÉLÉARD, S. AND VILLEMONAIS, D. (2012). Quasi-stationary distributions and population processes. *Prob. Surveys* **9**, 340–410.
- [15] POLLETT, P. (1996). Modelling the long-term behaviour of evanescent ecological systems. *Ecol. Modelling* **86**, 135–139.
- [16] POLLETT, P. (2001). Quasi-stationarity in populations that are subject to large-scale mortality or emigration. *Environm. Internat.* **27**, 231–236.
- [17] SALINS, M., BUDHIRAJA, A. AND DUPUIS, P. (2017). Uniform large deviation principles for Banach space valued stochastic differential equations. To appear in *Trans. Amer. Math. Soc.*
- [18] VAN DOORN, E. A. (1991). Quasi-stationary distributions and convergence to quasi-stationarity of birth–death processes. *Adv. Appl. Prob.* **23**, 683–700.
- [19] VAN DOORN, E. A. AND POLLETT, P. K. (2009). Quasi-stationary distributions for reducible absorbing Markov chains in discrete time. *Markov Process. Relat. Fields* **15**, 191–204.