Limit of a Poissonian SDE depending upon a parameter, and asymptotic analysis of the invariant measure

Elma Nassar* and Etienne Pardoux**

May 31, 2017

* Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
elma_nassar@hotmail.com
** Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
CMI, 39 rue F. Joliot Curie 13453 Marseille cedex 13
etienne.pardoux@univ-amu.fr, corresponding author.

Abstract

We consider a Poissonian SDE for the lack of fitness of a population subject to a continuous change of its environment, and an accumulation of advantageous mutations. We neglect the time of fixation of new mutations, so that the population is monomorphic at all times. We consider the asymptotic of small and frequent mutations. In that limit, we establish a law of large numbers and a central limit theorem. For small enough mutations, the original process is Harris recurrent and ergodic. We show in which sense the limits as $t \to \infty$ of the law of large and number and central limit theorem give a good approximation of the invariant probability measure of the original process.

Keys words: Poissonian SDE, Law of large numbers, Central limit theorem, Approximation of invariant measure, canonical equation of adaptive dynamics, moving optimum model.

Introduction

The present work is motivated by the moving optimum model in theoretical biology, which aims at evaluating the possibility for mutations to rescue a given population undergoing a linear change in its environment which deteriorates its survival conditions. We refer the reader to [6] and [8] for the presentation of this model. The authors have set in [10] a rigorous mathematical study of the moving optimum model by introducing a stochastic differential equation driven by a Poisson point process describing the evolution of a quantitative one–dimensional phenotypic trait in accordance to the biological description of this evolutionary rescue model. They studied the large time behavior of its solution, which is Harris recurrent when the speed of the environment v is smaller than the mean effect of the beneficial mutations m per unit time, transient if $v > m$. In the case of equality between the two parameters, the solution of the stochastic differential equation can either be transient or Harris recurrent depending upon additional technical conditions.

One is mainly interested in the positive recurrent case. However, the limitation of the ability to draw biological conclusions from these result is due to the difficulty to compute explicitly any quantity related to the invariant probability measure. This led us to study the small jumps limit, which is obtained by multiplying the jumps' sizes by ε, dividing the rates by ε^2, and then letting $\varepsilon \to 0$. In this paper, we study the limit as $\varepsilon \to 0$ in the such rescaled multidimensional version of the SDE from [10]. The solution X_t^ε then tends to $\bar{X}_t$, the solution of of an ODE. This is a law of large number type of result. The limiting ODE can be interpreted as the canonical equation of adaptive dynamics in the context of a changing environment, see [2] and [1]. The next step is to establish a Central Limit Theorem. Indeed, we define $U_t^\varepsilon = \varepsilon^{-1/2}(X_t^\varepsilon - \bar{X}_t)$, and show that U_t^ε converges in law to an Ornstein–Uhlenbeck process U_t. In our framework, it is not hard to show that as $t \to \infty$, both $\bar{X}_t \to \bar{X}_\infty$ and the law of U_t converges to a Gaussian law $\mathcal{N}(0, \bar{S}^2)$.

Both $\bar{X}_\infty$ and $\bar{S}^2$ can be easily computed with high accuracy, given the parameters of the model. On the other hand, we show that there exists $\varepsilon_0 > 0$ such that, for any $0 < \varepsilon \leq \varepsilon_0$, the process X_t^ε is Harris recurrent and possesses a unique invariant probability measure μ^ε. Since our motivation for studying the small jumps limit is to get informations about μ^ε, it is desirable to show that the pair $(\bar{X}_\infty, \bar{S}^2)$ gives a precise approximation of the invariant measure μ^ε, for small ε. This is a delicate question, since it amounts in a sense to invert the two limits $\varepsilon \to 0$ and $t \to \infty$.

We first show that the collection of probability measures $\{\mu^\varepsilon,\ \varepsilon \leq \varepsilon_0\}$ is tight. It is then not too difficult to deduce that $\mu^\varepsilon \Rightarrow \delta_{\bar{X}_\infty}$, as $\varepsilon \to 0$. We want to prove more, namely that μ^ε is close to ν^ε, which is the law of $\bar{X}_\infty + \sqrt{\varepsilon}\xi$, if $\xi \simeq \mathcal{N}(0, \bar{S}^2)$. This is done by analysing the probability measure

$$\mu_t^\varepsilon(A) = \frac{1}{t}\int_0^t \mathbf{1}_A(X_s^\varepsilon)ds$$

for large t and small ε.

We believe that those results are original, and have an interest, not only for the specific model which we study, but could also be useful in different frameworks, where a process converges in law to a limiting process, and one wants to compare large time behaviors. Note also that numerical simulations tend to indicate that the approximation is valid even for not very small values of ε. This will be reported upon elsewhere.

The paper is organized as follows: After giving some useful notations, section 1 gives a detailed presentation of the model, from the biological literature. Furthermore we present the stochastic differential equation that describes the evolution of the vector phenotypic lag between the population and its environment, explaining the fixation mechanism of mutations.

In section 2, we conduct a small jumps limit based on the above mentioned equation, prove the law of large numbers and the central limit theorem.

Section 3 is dedicated to the study of the large time behavior of X_t^ε which will turn out to be positive Harris recurrent for ε sufficiently small, admitting thus a unique invariant measure. Then we proceed to prove that the sequence of invariant measures is tight and converges in law. We finally give a precise statement which describes in which sense the invariant leasure is well approximated by a combination of the limit as $t \to \infty$ of the LLN and the CLT limits.

Notation

We remind that the quadratic variation of a scalar discontinuous bounded variation local martingale M_t is the sum of the squares of its jumps and is denoted by :

$$[M_t] = \sum_{s \le t} |\Delta M_s|^2.$$

Its predictable quadratic variation $\langle M \rangle_t$ is the unique increasing predictable process such that $[M]_t - \langle M \rangle_t$, and hence also $M_t^2 - \langle M \rangle_t$ is a martingale.

In the d-dimensional case, we define the quadratic variation of a discontinuous bounded variation local martingale M_t as :

$$[[M]]_t = \sum_{s \le t} \Delta M_s \otimes \Delta M_s.$$

Its predictable quadratic variation $\langle\langle M \rangle\rangle_t$ is the unique S^d-valued predictable increasing process such that both $[[M]]_t - \langle\langle M \rangle\rangle_t$ and $M_t \otimes M_t - \langle\langle M \rangle\rangle_t$ are S^d-valued martingales. Here S^d denotes the set of symmetric positive semi–definite $d \times d$ matrices. Note that $[[M]]_t$ (resp. $\langle\langle M \rangle\rangle_t$) is the matrix whose i, j element is $[M^i, M^j]_t$ (resp. $\langle M^i, M^j \rangle_t$).

We shall use the notation

$$\langle M \rangle_t = \mathrm{Tr} \langle\langle M \rangle\rangle_t,$$

so that $|M_t|^2 - \langle M \rangle_t$ is an $\mathbb{R}$-valued martingale, and the notations in the scalar and vector case are coherent. See [13] for more details.

1 The model

The model from Matuszewski *et al.* [8] is set up as follows: a population of constant size N is subject to Gaussian stabilizing selection with a moving optimum that increases linearly with speed vector $v \in \mathbb{R}^d$. That is, at time t, the phenotypic lag between an individual with trait value $z \in \mathbb{R}^d$ and the optimum equals $x = z - vt \in \mathbb{R}^d$, and the corresponding fitness is

$$\mathcal{W}(x) = \exp\left(-x' \Sigma^{-1} x\right), \tag{1}$$

where Σ describes the shape of the fitness landscape. For the adaptive-walk approximation, the population is assumed to be monomorphic at all times (i.e., its state is completely characterized by x). Mutations arise at rate $\Theta/2 = N\mu$ (where μ is the *per-capita* mutation rate and $\Theta = 2N\mu$ is a standard population-genetic parameter), and their phenotypic effects α are drawn from a distribution $p(\alpha)$. We neglect the possibility of fixation for deleterious mutations. Yet even beneficial mutations have a significant probability of being lost due to the effects of genetic drift while they are rare. A mutation with effect α that arises in a population with phenotypic lag x has a probability of fixation

$$g(x, \alpha) = \begin{cases} 1 - \exp(-2s(x, \alpha)) & \text{if } s(x, \alpha) > 0, \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

where

$$s(x, \alpha) = \frac{\mathcal{W}(x + \alpha)}{\mathcal{W}(x)} - 1 \approx -(2x + \alpha)' \Sigma^{-1} \alpha \tag{3}$$

is the selection coefficient. Formula (2) is a good approximation of the fixation probability derived under a diffusion approximation [5, 7], as long as the population size N is not too small. Note that Matuszewski *et al.* [8] used the even simpler approximation $g(x, \alpha) \approx 2s(x, \alpha)$ ([4]; for more exact approximations for the fixation probability in changing environments, see [14, 12]). Once a mutation gets fixed, it is assumed to do so instantaneously (which is of course a simplification which is not realistic), and the phenotypic lag x of the population is updated accordingly.

We make the following assumptions (see [8]):

1. v is a horizontal vector,
2. Σ is isotropic, i.e. $\Sigma = \sigma^2 \mathbf{I}_{\mathbb{R}^d}$.

It is always possible to reduce the situation to our assumptions, via a change of variables.

The evolution of the phenotypic lag of the population can be described by the following equation:

$$X_t = x_0 - vt + \int_{[0,t]\times\mathbb{R}^d\times[0,1]} \alpha\Gamma(X_{s^-}, \alpha, \xi) N(ds, d\alpha, d\xi). \tag{4}$$

Here, N is a Poisson point process over $\mathbb{R}_+ \times \mathbb{R}^d \times [0,1]$ with intensity $ds\ \nu(d\alpha)\ d\xi$ where $\nu(d\alpha)$ is the measure of new mutations and

$$\Gamma(x, \alpha, \xi) = \mathbf{1}_{\{\xi \leq g(x,\alpha)\}},$$

where the fixation probability $g(x, \alpha)$ of a mutation of size α that hits the population when the lag is x, as defined by (2) and (3), can be expressed as

$$g(x, \alpha) = \Big(1 - \exp\Big(2\sigma^{-2}\left(2x + \alpha \mid \alpha\right)\Big)\Big) \times \mathbf{1}_{\{(2x+\alpha|\alpha)\leq 0\}}.$$

Following the model by [8], we consider that

$$\nu(d\alpha) = \frac{\Theta}{2} p(\alpha) d\alpha, \tag{5}$$

where p is the density of a centered multidimensional Gaussian distribution $\mathcal{N}(0, M)$, M being a positive definite symmetric matrix. Under the above assumptions around the speed vector v and the fitness matrix Σ, M is generally not an isotropic matrix.

Mechanism of Fixation of Mutations : The points of this Poisson Point Process (T_i, A_i, Ξ_i) are such that the (T_i, A_i) form a Poisson Point Process over $\mathbb{R}_+ \times \mathbb{R}^d$ of the mutations that hit the population with intensity $ds\nu(d\alpha)$, and the Ξ_i are i.i.d. $\mathcal{U}[0, 1]$, globally independent of the Poisson Point Process of the (T_i, A_i). T_i's are the times when mutations are proposed and A_i's are the effect sizes of those mutations. The Ξ_i are auxiliary variables determining fixation: a mutation gets instantaneously fixed if $\Xi_i \leq g(X_{T_i}, A_i)$, and is lost otherwise.

Note that the limit of the probability of fixation as $|x| \to \infty$, while $\frac{x}{|x|}$ remains a constant unit vector, is $\mathbf{1}_{\{(x|\alpha)\leq 0\}}$. This means that when the process is sufficiently far away from 0, the fixation mechanism tends to accept all mutations inside the half space $\left(\frac{x}{|x|} \mid \alpha\right) \leq 0$.

Define the covariance matrix of fixed mutations :

$$\bar{V}(x) = \int_{(x|\alpha)\leq 0} \alpha \otimes \alpha \ \nu(d\alpha). \tag{6}$$

Proposition 1. *Under the definition of ν given by* (5)*, $\bar{V}(x)$ is independent of the direction of x and*

$$\bar{V} = \frac{\Theta}{4} M.$$

Proof. We assume without loss of generality that we have chosen as orthonormal basis of $\mathbb{R}^d$ a basis made of eigenvectors of M, the covariance matrix of p. In that case,

$$M = P'DP,$$

where D is a diagonal matrix and P is the matrix representing the change of basis such that $P' = P^{-1} = P^*$. First, we will show that

$$\frac{1}{(2\pi)^{\frac{d}{2}} (\det D)^{\frac{1}{2}}} \int_{(x|\alpha)\leq 0} \alpha \otimes \alpha \ e^{-\frac{1}{2}\alpha' D^{-1}\alpha} d\alpha = \frac{D}{2}. \tag{7}$$

This is equivalent to showing that for $X_1, \ldots, X_d$ being mutually independent zero mean Gaussian random variables, and for any vector $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$, with the notation $(X|a) = \sum_{i=1}^d a_i X_i$,

$$\mathbb{E}\left[X_j^2; (X|a) \leq 0\right] = \frac{1}{2}\mathbb{E}\left[X_j^2\right], \quad \text{for any } j \in \{1, \ldots, d\}$$

and

$$\mathbb{E}\left[X_j X_\ell; (X|a) \leq 0\right] = 0, \quad \text{for any } j \neq \ell \in \{1, \ldots, d\}.$$

The first of these two identities follows from the fact that if X and Y are two mutually independent zero mean Gaussian random variables, then

$$\mathbb{E}[X^2; Y < X] = \frac{1}{2}\mathbb{E}[X^2].$$

Indeed, if F_Y denotes the distribution function of the zero mean Gaussian r.v. Y, and σ^2 is the variance of X, since $F_Y(x)+F_Y(-x)=1$ for all $x\in\mathbb{R}$,

$$\begin{aligned}\mathbb{E}[X^2;Y\le X] &= \mathbb{E}[X^2F_Y(X)]\\ &= \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty x^2[F_Y(x)+F_Y(-x)]e^{-x^2/2\sigma^2}dx\\ &= \frac{1}{2}\mathbb{E}[X^2].\end{aligned}$$

We now establish the second formula. All we have to compute is the following quantity, where X,Y,Z are mutually independent zero mean Gaussian random variables, and a,b are arbitary real numbers,

$$\begin{aligned}\mathbb{E}[XY;Z\le aX+bY] &= \mathbb{E}[XYF_Z(aX+bY)]\\ &= \frac{1}{\sigma\tau 2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} xyF_Z(ax+by)e^{-\frac{x^2}{2\sigma^2}-\frac{y^2}{2\tau^2}}dxdy\\ &= \frac{1}{\sigma\tau 2\pi}\int_0^\infty\int_0^\infty xy[F_Z(ax+by)+F_Z(-ax-by)\\ &\qquad -F_Z(-ax+by)-F_Z(ax-by)]dxdy\\ &= 0,\end{aligned}$$

since clearly $F_Z(ax+by+F_Z(-ax-by)=F_Z(-ax+by)+F_Z(ax-by)=1$.

Now that (7) is established, the change of variables $\alpha=P'\tilde{\alpha}$ in the integral formula for $\bar{V}$ yields

$$\bar{V}=\frac{\Theta}{2}\left(P'\frac{D}{2}P\right)=\frac{\Theta}{4}M.$$

□

2 Small Jumps Limit

In the following, we introduce the rescaling

$$\tilde{\alpha}=\varepsilon\alpha \quad\text{and}\quad \tilde{s}=\frac{s}{\varepsilon^2}\text{ with }\varepsilon>0$$

of the jumps and the time, respectively. In other words, we rewrite our process (4) as

$$X_t^\varepsilon = x_0-vt+\int_0^t\int_{\mathbb{R}^d}\int_0^1 \varepsilon\alpha\varphi(X_{s^-}^\varepsilon,\varepsilon\alpha,\xi)M_\varepsilon(ds,d\alpha,d\xi),$$

where the intensity measure of the Poisson Point Process M_ε is

$$\left(\frac{1}{\varepsilon^2}ds\right) \times \nu(d\alpha) \times d\xi.$$

The above SDE can be rewritten as

$$X_t^\varepsilon = x_0 - vt + \int_0^t \frac{1}{\varepsilon^2} m_\varepsilon(X_s^\varepsilon) ds + \mathcal{M}_t^\varepsilon, \tag{8}$$

with

$$\begin{aligned} m_\varepsilon(x) &= \int_{\mathbb{R}^d} \varepsilon\alpha g(x, \varepsilon\alpha)\nu(d\alpha), \\ \text{where} \quad g(x, \varepsilon\alpha) &\leq 2\sigma^{-2} \left|(2x + \varepsilon\alpha \mid \varepsilon\alpha)\right| \mathbf{1}_{\{(2x+\varepsilon\alpha|\varepsilon\alpha)\leq 0\}} \\ &\leq 4\sigma^{-2}\varepsilon \left|(x \mid \alpha)\right| \mathbf{1}_{\{(x|\alpha)\leq 0\}}, \end{aligned}$$

and

$$\mathcal{M}_t^\varepsilon = \int_0^t \int_{\mathbb{R}^d} \int_0^1 \varepsilon\alpha\varphi(X_{s-}^\varepsilon, \varepsilon\alpha, \xi) \bar{M}_\varepsilon(ds, d\alpha, d\xi)$$

is a martingale.

Lemma 1. *The following two properties hold:*

1. *the collection* $\{X^\varepsilon, \ 0 < \varepsilon \leq 1\}$ *is tight in* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$*;*
2. *we have*

$$\langle \mathcal{M}^\varepsilon \rangle_t \underset{\varepsilon\to 0}{\longrightarrow} 0 \text{ in probability, locally uniformly in } t.$$

Proof. It is plain that

$$\frac{1}{\varepsilon^2} \left|m_\varepsilon(X_t^\varepsilon)\right| \leq 4|X_t^\varepsilon|\sigma^{-2} \int_{\mathbb{R}^d} |\alpha|^2 \nu(d\alpha), \tag{9}$$

and

$$\begin{aligned} \langle \mathcal{M}^\varepsilon \rangle_t &= \frac{1}{\varepsilon^2} \int_0^t \int_{\mathbb{R}^d} |\varepsilon\alpha|^2 g(X_s^\varepsilon, \varepsilon\alpha)\nu(d\alpha)ds \\ &\leq 4\sigma^{-2}\varepsilon \int_0^t \int_{(\alpha|X_s^\varepsilon)\leq 0} |\alpha|^2 \left|(\alpha \mid X_s^\varepsilon)\right| \nu(d\alpha)ds \\ &\leq 4\sigma^{-2}\varepsilon \left(\int_{\mathbb{R}^d} |\alpha|^3 \nu(d\alpha)\right) \int_0^t \left|X_s^\varepsilon\right| ds. \end{aligned} \tag{10}$$

Moreover for all x, we have

$$\frac{1}{\varepsilon^2} \left(m_\varepsilon(x) \mid x\right) \leq 0.$$

On the other hand,

$$|X_t^\varepsilon|^2 = |X_0^\varepsilon|^2 - 2\int_0^t (v \mid X_s^\varepsilon)\, ds + \frac{2}{\varepsilon^2}\int_0^t (m_\varepsilon(X_s^\varepsilon) \mid X_s^\varepsilon)\, ds + 2\int_0^t X_{s^-} dM_s^\varepsilon + \langle \mathcal{M}^\varepsilon \rangle_t$$

Hence, for fixed T, we have for all $0 \le t \le T$

$$\begin{aligned} \mathbb{E}|X_t^\varepsilon|^2 &\le \mathbb{E}|X_0^\varepsilon|^2 + 2v_1 \int_0^t \mathbb{E}|X_s^\varepsilon| ds + \mathbb{E}\langle \mathcal{M}^\varepsilon \rangle_t \\ &\le \mathbb{E}|X_0^\varepsilon|^2 + v_1 \left(t + \int_0^t \mathbb{E}(X_s^\varepsilon)^2 ds \right) + 2\varepsilon \int_{\mathbb{R}^d} |\alpha|^3 \nu(d\alpha) \left(t + \int_0^t \mathbb{E}(X_s^\varepsilon)^2 ds \right) \\ &\le \mathbb{E}|X_0^\varepsilon|^2 + \left(v_1 + 2\varepsilon \int_{\mathbb{R}^d} |\alpha|^3 \nu(d\alpha) \right) t + \left(v_1 + 2\varepsilon \int_{\mathbb{R}^d} |\alpha|^3 \nu(d\alpha) \right) \int_0^t \mathbb{E}(X_s^\varepsilon)^2 ds \\ &\le \left(\mathbb{E}|X_0^\varepsilon|^2 + \left(v_1 + 2\varepsilon \int_{\mathbb{R}^d} |\alpha|^3 \nu(d\alpha) \right) T \right) e^{\left[v_1 + 2\varepsilon \left(\int_{\mathbb{R}^d} |\alpha|^3 \nu(d\alpha) \right) \right] T}, \end{aligned} \tag{11}$$

since $2\mathbb{E}|X| \le 1 + \mathbb{E}X^2$. We deduce from (9), (10) and (11) that the process X^ε is tight in $\mathbb{D}\left(\mathbb{R}_+, \mathbb{R}^d\right)$ due to Remark 14, part 2 following Proposition 37 in [11] and from (10),

$$\langle \mathcal{M}^\varepsilon \rangle_t \xrightarrow[\varepsilon \to 0]{} 0 \text{ in probability.}$$

This convergence is uniformly in t since $t \mapsto \langle \mathcal{M}^\varepsilon \rangle_t$ is increasing. □

Lemma 2. *For all $x \in \mathbb{R}^d$, we have that*

1. $m_\varepsilon(x) \xrightarrow[\varepsilon \to 0]{} Lx$, *where*

$$Lx = 4\sigma^{-2} \int_{(x|\alpha) \le 0} \alpha \, |(x \mid \alpha)| \, \nu(d\alpha) = -4\sigma^{-2} \bar{V} x = -\Theta \sigma^{-2} M x. \tag{12}$$

2. $\int \alpha \otimes \alpha \frac{g(x, \varepsilon\alpha)}{\varepsilon} \nu(d\alpha) \xrightarrow[\varepsilon \to 0]{} \Lambda(x)$, *where*

$$\Lambda(x) = 4\sigma^{-2} \int_{(x|\alpha) \le 0} |\,(x \mid \alpha)\,| \alpha \otimes \alpha \nu(d\alpha), \tag{13}$$

Remark 1. *Note that the two above Lemmas remain true if the measure ν is not Gaussian, but satisfies the following moment condition :*

$$\int_{\mathbb{R}^d} |\alpha|^4 \nu(d\alpha) < \infty. \tag{14}$$

In this case the limit $L(x)$ is given by the expression

$$L(x) = 4\sigma^{-2} \int_{(x|\alpha) \le 0} \alpha \, |(x \mid \alpha)| \, \nu(d\alpha) = -4\sigma^{-2} \bar{V}(x) x,$$

where $\bar{V}$ depends upon the direction of x. The advantage of taking a Gaussian measure (up to a multiplicative constant) of new mutations is the resulting linear behavior of the limit Lx, due to the fact that $\bar{V}(x)$ is a constant matrix. More generally, $\bar{V}(x)$ would depend upon $\frac{x}{|x|}$.

Proof of Lemma 2. Let for all $\varepsilon > 0$ and $x, \alpha \in \mathbb{R}^d$

$$y_\varepsilon = -2\sigma^{-2}\left(2x + \varepsilon\alpha \mid \alpha\right) \text{ and } y = -4\sigma^{-2}\left(x \mid \alpha\right).$$

We have that

$$\frac{1 - e^{-\varepsilon y_\varepsilon}}{\varepsilon}\mathbf{1}_{y_\varepsilon \geq 0} - y\mathbf{1}_{y\geq 0} = \left(\frac{1 - e^{-\varepsilon y_\varepsilon}}{\varepsilon}\mathbf{1}_{y_\varepsilon \geq 0} - y_\varepsilon\mathbf{1}_{y_\varepsilon \geq 0}\right) + \left(y_\varepsilon\mathbf{1}_{y_\varepsilon\geq 0} - y\mathbf{1}_{y\geq 0}\right). \quad (15)$$

In addition,

$$-\varepsilon\frac{y_\varepsilon^2}{2}\mathbf{1}_{y_\varepsilon\geq 0} \leq \frac{1 - e^{-\varepsilon y_\varepsilon}}{\varepsilon}\mathbf{1}_{y_\varepsilon \geq 0} - y_\varepsilon\mathbf{1}_{y_\varepsilon \geq 0} \leq 0, \quad (16)$$

since for all $z > 0$

$$z - \frac{z^2}{2} \leq 1 - e^{-z} \leq z.$$

Combining (16) with the fact that

$$y_\varepsilon^2\mathbf{1}_{y_\varepsilon\geq 0} \leq y^2\mathbf{1}_{y\geq 0} \leq 16\sigma^{-4}|x|^2|\alpha|^2,$$

we obtain

$$-8\sigma^{-4}\varepsilon|x|^2|\alpha|^2 \leq \frac{1 - e^{-\varepsilon y_\varepsilon}}{\varepsilon}\mathbf{1}_{y_\varepsilon \geq 0} - y_\varepsilon\mathbf{1}_{y_\varepsilon \geq 0} \leq 0. \quad (17)$$

Moreover, $\{y_\varepsilon \geq 0\} \subset \{y \geq 0\}$. It follows that

$$\begin{aligned} y_\varepsilon\mathbf{1}_{y_\varepsilon\geq 0} - y\mathbf{1}_{y\geq 0} &= (y_\varepsilon - y)\mathbf{1}_{y\geq 0} - y_\varepsilon\mathbf{1}_{y\geq 0\backslash y_\varepsilon\geq 0} \\ &= -2\sigma^{-2}\varepsilon|\alpha|^2\mathbf{1}_{y\geq 0} - y_\varepsilon\mathbf{1}_{y\geq 0\backslash y_\varepsilon\geq 0}. \end{aligned} \quad (18)$$

Furthermore, $\{y \geq 0\backslash y_\varepsilon \geq 0\} = \{-\varepsilon|\alpha|^2 < (2x \mid \alpha) < 0\}$. Thus,

$$0 < -y_\varepsilon\mathbf{1}_{y\geq 0\backslash y_\varepsilon\geq 0} < 2\sigma^{-2}\varepsilon|\alpha|^2\mathbf{1}_{y\geq 0}. \quad (19)$$

Combining (18) and (19), we get

$$-2\sigma^{-2}\varepsilon|\alpha|^2 \leq y_\varepsilon\mathbf{1}_{y_\varepsilon\geq 0} - y\mathbf{1}_{y\geq 0} \leq 0. \quad (20)$$

We deduce from (15), (17) and (20) that

$$-\varepsilon\left(8\sigma^{-4}|x|^2 + 2\sigma^{-2}\right)|\alpha|^2 \leq \frac{1 - e^{-\varepsilon y_\varepsilon}}{\varepsilon}\mathbf{1}_{y_\varepsilon \geq 0} - y\mathbf{1}_{y\geq 0} \leq 0.$$

Hence,

$$\left|\alpha\left(\frac{1-e^{-\varepsilon y_\varepsilon}}{\varepsilon}\mathbf{1}_{y_\varepsilon\geq 0}-y\mathbf{1}_{y\geq 0}\right)\right|\leq \varepsilon\left(8\sigma^{-4}|x|^2+2\sigma^{-2}\right)|\alpha|^3. \tag{21}$$

By integrating (21) with respect to ν, we obtain

$$\left|\frac{1}{\varepsilon^2}m_\varepsilon(x)-L(x)\right|\leq \varepsilon\left(8\sigma^{-4}|x|^2+2\sigma^{-2}\right)\int_{\mathbb{R}^d}|\alpha|^3\nu(d\alpha). \tag{22}$$

Hence,

$$m_\varepsilon(x)\xrightarrow[\varepsilon\to 0]{} Lx.$$

By a similar argument, we have that

$$\left|\int \alpha\otimes\alpha\frac{g(x,\varepsilon\alpha)}{\varepsilon}\nu(d\alpha)-\Lambda(x)\right|\leq \varepsilon\left(8\sigma^{-4}|x|^2+2\sigma^{-2}\right)\int_{\mathbb{R}^d}|\alpha|^4\nu(d\alpha), \tag{23}$$

and we deduce the second result of the Lemma. □

It follows from Lemma 1 that we can extract a subsequence which we still denote X^ε by an abuse of notation such that $X^\varepsilon\Rightarrow\bar{X}$, and we have the following result:

Proposition 2. *We have that*

$$\frac{1}{\varepsilon^2}m_\varepsilon(X^\varepsilon_.)\Rightarrow L\bar{X}. \quad \text{in } D(\mathbb{R}_+,\mathbb{R}^d),$$

and

$$\frac{1}{\varepsilon}\langle\langle\mathcal{M}^\varepsilon\rangle\rangle_.\Rightarrow\Lambda(\bar{X}_.) \quad \text{in } D(\mathbb{R}_+,\mathbb{R}^d),$$

Proof. It follows from (22), that for all $\delta, C>0$, there exists $\varepsilon_{\delta,C}$ such that if $\varepsilon<\varepsilon_{\delta,C}$ and for all $|x|\leq C$,

$$\left|\frac{1}{\varepsilon^2}m_\varepsilon(x)-Lx\right|\leq\delta,$$

thus, for an arbitrary $T>0$ and for $\varepsilon<\varepsilon_{\delta,C}$,

$$\mathbb{P}\left(\sup_{t\leq T}\left|\frac{1}{\varepsilon^2}m_\varepsilon(X^\varepsilon_t)-L(X^\varepsilon_t)\right|>\delta\right)\leq\mathbb{P}\left(\sup_{t\leq T}|X^\varepsilon_t|>C\right).$$

It follows that for all $\delta, C>0$,

$$\limsup_{\varepsilon\to 0}\mathbb{P}\left(\sup_{t\leq T}\left|\frac{1}{\varepsilon^2}m_\varepsilon(X^\varepsilon_t)-L(X^\varepsilon_t)\right|>\delta\right)\leq\sup_\varepsilon\mathbb{P}\left(\sup_{t\leq T}|X^\varepsilon_t|>C\right).$$

From the tightness of X^ε, for all $\eta > 0$, we can choose $C > 0$ such that

$$\sup_{\varepsilon} \mathbb{P}\left(\sup_{t\leq T} |X_t^\varepsilon| > C\right) \leq \eta.$$

Hence for all $\delta, \eta > 0$

$$\limsup_{\varepsilon\to 0} \mathbb{P}\left(\sup_{t} \left|\frac{1}{\varepsilon^2} m_\varepsilon(X_t^\varepsilon) - L(X_t^\varepsilon)\right| > \delta\right) \leq \eta.$$

Moreover $L(X^\varepsilon) \Rightarrow L(\bar{X})$ in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ since L is continuous function. Consequently,

$$\frac{1}{\varepsilon^2} m_\varepsilon(X^\varepsilon) = \frac{1}{\varepsilon^2} m_\varepsilon(X^\varepsilon) - L(X^\varepsilon) + L(X^\varepsilon) \Rightarrow L(\bar{X}).$$

By a similar argument using (23), we can prove the second part of this Lemma since

$$\int |\alpha|^4 \nu(d\alpha) < \infty.$$

□

It follows

Theorem 1. $X_t^\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} \bar{X}_t$ *in probability, locally uniformly in* t*, where*

$$\frac{d\bar{X}_t}{dt} = -v + L\bar{X}_t, \quad X_0^\varepsilon = \bar{X}_0 = x_0. \tag{24}$$

Proof. The result for a subsequence is an immediate consequence of (8), Lemma 1 and Proposition 2. Since the limit $\bar{X}_t$ is uniquely determined, the whole process X_t^ε converges in probability towards $\bar{X}_t$.

□

The differential equation (24) represents a deterministic approximation for the stochastic process $\bar{X}^\varepsilon$ in the limit of small jumps. We note that

$$\bar{X}_t \underset{t\to\infty}{\longrightarrow} \bar{X}_\infty = -\frac{M^{-1}v}{\Theta\sigma^{-2}}. \tag{25}$$

To estimate the fluctuations of the process in the small-jumps limit, we now consider the following process

$$U_t^\varepsilon = \frac{X_t^\varepsilon - \bar{X}_t}{\sqrt{\varepsilon}}.$$

Theorem 2. *$U^\varepsilon \Rightarrow U$, where U is an Ornstein-Uhlenbeck process:*

$$\begin{aligned} &dU_t = LU_t dt + \Lambda^{\frac{1}{2}}(\bar{X}_t)dB_t,\\ &U_0 = 0, \end{aligned} \tag{26}$$

with B being a d-dimensional standard Brownian motion. In other words,

$$U_t = \int_0^t e^{L(t-s)}\Lambda^{\frac{1}{2}}(\bar{X}_s)dB_s.$$

Proof. We have that

$$\begin{aligned} U_t^\varepsilon &= \int_0^t \frac{\varepsilon^{-2}m_\varepsilon(X_s^\varepsilon) - L\bar{X}_s}{\sqrt{\varepsilon}}ds + \frac{1}{\sqrt{\varepsilon}}\mathcal{M}_t^\varepsilon,\\ &= L\int_0^t \frac{X_s^\varepsilon - \bar{X}_s}{\sqrt{\varepsilon}}ds + \frac{1}{\sqrt{\varepsilon}}\int_0^t \left(\frac{1}{\varepsilon^2}m_\varepsilon(X_s^\varepsilon) - LX_s^\varepsilon\right)ds + \frac{1}{\sqrt{\varepsilon}}\mathcal{M}_t^\varepsilon\\ &= L\int_0^t U_s^\varepsilon ds + \frac{1}{\sqrt{\varepsilon}}\int_0^t \left(\frac{1}{\varepsilon^2}m_\varepsilon(X_s^\varepsilon) - LX_s^\varepsilon\right)ds + \frac{1}{\sqrt{\varepsilon}}\mathcal{M}_t^\varepsilon\\ &= L\int_0^t U_s^\varepsilon ds + G_\varepsilon(t), \end{aligned}$$

where

$$G_\varepsilon(t) = \frac{1}{\sqrt{\varepsilon}}\int_0^t \left(\frac{1}{\varepsilon^2}m_\varepsilon(X_s^\varepsilon) - LX_s^\varepsilon\right)ds + \frac{1}{\sqrt{\varepsilon}}\mathcal{M}_t^\varepsilon.$$

Thus

$$U_t^\varepsilon = \int_0^t e^{L(t-s)}dG_\varepsilon(s).$$

We deduce from (22) that

$$\frac{1}{\sqrt{\varepsilon}}\int_0^t e^{L(t-s)}\left(\frac{1}{\varepsilon^2}m_\varepsilon(X_s^\varepsilon) - LX_s^\varepsilon\right)ds \underset{\varepsilon\to 0}{\longrightarrow} 0.$$

Furthermore, $\frac{1}{\sqrt{\varepsilon}}\mathcal{M}^\varepsilon$ are tight martingales by a similar argument as before since

$$\sup_{\{t>0,\ \varepsilon\}}\int_{\mathbb{R}^2}|\alpha|^2 g(X_t^\varepsilon, \varepsilon\alpha)\nu(d\alpha) \le \int_{\mathbb{R}^2}|\alpha|^2\nu(d\alpha) < \infty.$$

Moreover for any arbitrary $T > 0$,

$$\sup_{t\le T}\left|\frac{1}{\sqrt{\varepsilon}}\left(\mathcal{M}_t^\varepsilon - \mathcal{M}_{t-}^\varepsilon\right)\right| \underset{\varepsilon\to 0}{\longrightarrow} 0,$$

since the jumps are multiplied by ε. Hence every converging subsequence of $\frac{1}{\sqrt{\varepsilon}}\mathcal{M}^\varepsilon$ converges to a continuous martingale $\mathcal{M}$ as ε goes to 0, and using Lemma 2

$$\langle\langle \frac{1}{\sqrt{\varepsilon}}\mathcal{M}^\varepsilon\rangle\rangle_t = \int_0^t \alpha\otimes\alpha \frac{g(X_s^\varepsilon,\varepsilon\alpha)}{\varepsilon}\nu(d\alpha)ds \underset{\varepsilon\to 0}{\longrightarrow} \int_0^t \Lambda(\bar{X}_s)ds,$$

it follows that

$$\langle\langle\mathcal{M}\rangle\rangle_t = \int_0^t \Lambda(\bar{X}_s)ds.$$

We deduce thanks to the representation theorem of continuous martingales that there exists a d-dimensional Brownian motion (B_t) such that

$$\mathcal{M}_t = \int_0^t \Lambda^{\frac{1}{2}}(\bar{X}_s)dB_s,\ t\geq 0.$$

This being true for any subsequence of $\frac{1}{\sqrt{\varepsilon}}\mathcal{M}^\varepsilon$, the limit is unique (in law). Finally

$$U_t^\varepsilon \Rightarrow \int_0^t e^{L(t-s)}\Lambda^{\frac{1}{2}}(\bar{X}_s)dB_s.$$

□

It follows that

$$\mathbb{E}(U_t\otimes U_t) = \int_0^t e^{-\Theta\sigma^{-2}M(t-s)}\Lambda(\bar{X}_s)e^{-\Theta\sigma^{-2}M'(t-s)}ds.$$

Consequently

$$\mathbb{E}(U_t\otimes U_t) \underset{t\to\infty}{\longrightarrow} \bar{S}^2 = \int_0^\infty e^{-\Theta\sigma^{-2}M(t-s)}\Lambda(\bar{X}_\infty)e^{-\Theta\sigma^{-2}M'(t-s)}ds. \qquad (27)$$

3 Large time behavior of X_t^ε for small $\varepsilon > 0$

3.1 Large time behavior of the process X_t^ε

Theorem 3. *There exists $\varepsilon_0 > 0$ such that for any $0<\varepsilon\leq\varepsilon_0$, the process X_t^ε is positive Harris recurrent, with a unique invariant probability measure.*

Let us first establish:

Lemma 3. *There exists three constants $c, C, a > 0$ such that for all $0<\varepsilon<1/4$,*

$$|X_t^\varepsilon| \leq |X_0^\varepsilon| + C\times t - c\int_0^t \left(|X_s^\varepsilon|\wedge\frac{a}{\varepsilon}\right)ds + \mathcal{M}_\varepsilon(t),$$

where $\mathcal{M}_\varepsilon(t)$ is a martingale, for each $\varepsilon>0$.

Proof. We have

$$|X_t^\varepsilon| = |X_0^\varepsilon| - \int_0^t \left(v \mid \frac{X_s^\varepsilon}{|X_s^\varepsilon|} \right) ds + \sum_{s \le t} \left(|X_{s^-}^\varepsilon + \Delta X_s^\varepsilon| - |X_{s^-}^\varepsilon| \right).$$

Recall that

$$|X_{s^-}^\varepsilon + \Delta X_s^\varepsilon| - |X_{s^-}^\varepsilon| \le 0,$$

and moreover we shall restrict ourselves to the jumps such that

$$|X_{s^-}^\varepsilon + \Delta X_s^\varepsilon| - |X_{s^-}^\varepsilon| \le -\frac{1}{2}|\Delta X_s^\varepsilon|$$

Let $C(x)$ denote the set of vectors α such that

$$|x+\alpha| - |x| \le -\frac{1}{2}|\alpha|.$$

Let us first define

$$\gamma = \text{angle between the two directions } -x \text{ and } \alpha.$$

From the identity

$$|x+\alpha| = \sqrt{|x|^2 + |\alpha|^2 - 2|x|\,|\alpha|\cos\gamma},$$

we deduce that $\alpha \in C(x)$ iff both

$$|\gamma| < \frac{\pi}{3}, \text{ and } |\alpha| \le \frac{4}{3}(2\cos\gamma - 1)|x|.$$

We note in particular that whenever $\alpha \in C(x)$ and $\varepsilon \le 1$, then $\varepsilon\alpha \in C(x)$. Now

$$\begin{aligned}
\sum_{s\le t} \left(|X_{s-}^\varepsilon + \Delta X_s^\varepsilon| - |X_{s-}^\varepsilon| \right) &= \int_{[0,t]\times\mathbb{R}^d\times[0,1]} \left(|X_{s-}^\varepsilon + \varepsilon\alpha| - |X_{s-}^\varepsilon| \right) \mathbf{1}_{u \le g(X_{s-}^\varepsilon, \varepsilon\alpha)} M_\varepsilon(ds, d\alpha, du) \\
&= \int_{[0,t]\times\mathbb{R}^d} \frac{|X_{s-}^\varepsilon + \varepsilon\alpha| - |X_{s-}^\varepsilon|}{\varepsilon} \times \frac{g(X_{s-}^\varepsilon, \varepsilon\alpha)}{\varepsilon} \nu(d\alpha)ds + \mathcal{M}_\varepsilon(t) \\
&\le -\frac{1}{2}\int_{[0,t]}\int_{C(X_{s-}^\varepsilon)/\varepsilon} |\alpha| \frac{g(X_{s-}^\varepsilon, \varepsilon\alpha)}{\varepsilon} \nu(d\alpha)ds + \mathcal{M}_\varepsilon(t) \\
&\le -\frac{b}{2}\int_{[0,t]}\int_{C(X_{s-}^\varepsilon)/\varepsilon} |\alpha| \left(\frac{(2X_{s-}^\varepsilon + \varepsilon\alpha, \alpha)_-}{\sigma^2} \wedge \frac{1}{\varepsilon} \right) \nu(d\alpha)ds + \mathcal{M}_\varepsilon(t),
\end{aligned}$$

where $\mathcal{M}_\varepsilon(t)$ is a martingale, $b = 1 - e^{-1}$, and we have exploited the elementary inequality

$$1 - e^{-u} \ge (1 - e^{-1})(u \wedge 1), \text{ for all } u \ge 0.$$

We now need to lower bound the factor of $-b/2$ in the last right hand side. For that sake, we consider the expression

$$\int_{\alpha \in C(x)/\varepsilon} |\alpha| \left(\frac{(2x + \varepsilon\alpha, \alpha)_-}{\sigma^2} \wedge \frac{1}{\varepsilon} \right) \nu(d\alpha).$$

For $|x| \le 2$, we lower bound this integral by 0. We now consider the case $|x| > 2$. We lower bound the integral by reducing the integration to the set

$$A_\varepsilon(x) = \frac{C(x)}{\varepsilon} \cap \{1 \le |\alpha| \le 2\}.$$

It is not hard to see that whenever $\alpha \in A_\varepsilon(x)$, $-2 \le \left(\alpha, \frac{x}{|x|}\right) < -1/2$. Moreover, since $|x| > 2$, if $\alpha \in A_\varepsilon(x)$,

$$-2|x| < (x, \alpha) < -\frac{|x|}{2} < -1, \quad \text{while } \varepsilon|\alpha|^2 \le 1,$$

provided $\varepsilon \le 1/4$. Consequently $(x, \alpha) + \varepsilon|\alpha|^2 \le 0$ and

$$(2x + \varepsilon\alpha, \alpha)_- \ge |(x, \alpha)| \ge \frac{|x|}{2},$$

so that

$$\begin{aligned} \int_{A_\varepsilon(x)} |\alpha| \left(\frac{(2x + \varepsilon\alpha, \alpha)_-}{\sigma^2} \wedge \frac{1}{\varepsilon} \right) \nu(d\alpha) &\ge \frac{1}{2\sigma^2} \left(|x| \wedge \frac{2\sigma^2}{\varepsilon} \right) \int_{A_\varepsilon(x)} |\alpha| \nu(d\alpha) \\ &\ge \frac{\beta}{2\sigma^2} \left(|x| \wedge \frac{2\sigma^2}{\varepsilon} \right), \end{aligned}$$

where $\beta = \inf_{|x|>2, \varepsilon \le 1} \int_{A_\varepsilon(x)} |\alpha| \nu(d\alpha) > 0$. We have proved that, with $a = 2\sigma^2$ and $c = b\beta/(4\sigma^2)$,

$$|X_t^\varepsilon| \le |X_0^\varepsilon| + |v| \times t - c \int_0^t \mathbf{1}_{|X_s^\varepsilon|>2} \left(|X_s^\varepsilon| \wedge \frac{a}{\varepsilon} \right) ds + \mathcal{M}_\varepsilon(t).$$

The result follows with $C = |v| + 2c$. □

Corollary 1. *There exists $\varepsilon_0 > 0$, $b, d > 0$ and B a compact subset of $\mathbb{R}^d$ such that for any $0 < \varepsilon \le \varepsilon_0$,*

$$|X_t^\varepsilon| \le |X_0^\varepsilon| - bt + d \int_0^t \mathbf{1}_B(X_s^\varepsilon) ds + \mathcal{M}_\varepsilon(t). \tag{28}$$

Proof. We choose $B = \{x;\ |x| \le 2C/c\}$ and $\varepsilon_0 = \frac{ca}{2C}$. (28) with $b = C$, $d = 2C$ now follows from Lemma 3. □

We finally turn to the

Proof of Theorem 3. The process X_t^ε satisfies (28) which is exactly condition (CD2) from [9]. In order to deduce positive Harris recurrence from Theorem 4.2 in [9], it remains to show that $B = \{x;\ |x| \le K\}$ is a closed petite set, with $K = 2C/c$.

Note that $K > 2$. We choose $0 < \eta < 1$ and define the event

$$A_{K,\eta}^\varepsilon = \left\{ \begin{array}{l} X_t^\varepsilon \text{ does not jump on the time interval } \left[0, \frac{2K-\eta}{v}\right], \text{ and} \\ X_t^\varepsilon \text{ jumps exactly once on the time interval } \left(\frac{2K-\eta}{v}, \frac{2K}{v}\right]. \end{array} \right\}$$

It is plain that $\mathbb{P}(A_{K,\eta}^\varepsilon) > 0$. Let T^ε denote the first jump time of X_t^ε. We note that, provided that $X_0^\varepsilon = x$ satisfies $|x| \le K$, on the event $A_{K,\eta}^\varepsilon$, $-3K \le (X_{T^\varepsilon-}^\varepsilon)_1 \le -K + \eta$, while $-K \le (X_{T^\varepsilon-}^\varepsilon)_2 \le K$. Let Λ_ε denote a random vector which is such that the law of $\varepsilon^{-1}\Lambda_\varepsilon$ has the density p. Denote by $f_\varepsilon(y) = \varepsilon p(\varepsilon y)$ the density of the law of Λ_ε. We define

$$\Sigma_\varepsilon = \Lambda_\varepsilon \mathbf{1}_{\{|X_{T-}^\varepsilon + \Lambda_\varepsilon| \le 1+\eta\}}.$$

Since $K > 1 + \eta$,

$$c_{\varepsilon,K,\eta} := \inf_{\Sigma_\varepsilon \ne 0} g(X_{T-}^\varepsilon, \Sigma_\varepsilon) > 0.$$

We denote by ξ the random variable with the uniform distribution on the interval $[0,1]$, which is such that whenever $\xi \le g(X_{T-}^\varepsilon, \Sigma_\varepsilon)$, the "proposed" jump Σ_ε happens at time T^ε. Recall that $0 \le \frac{2K}{v} - T^\varepsilon < \frac{\eta}{v}$. On the event $A_{K,\eta}^\varepsilon \cap \{\xi \le c_{\varepsilon,K,\eta}\}$,

$$X_{\frac{2K}{v}} = X_{T^\varepsilon}^\varepsilon + \Sigma_\varepsilon + v\left(\frac{2K}{v} - T^\varepsilon\right),$$

and for any $h \in C(\mathbb{R}^d, \mathbb{R}^+)$,

$$\mathbb{E}h\left(X_{\frac{2K}{v}}^\varepsilon\right) \ge \mathbb{P}(A_{K,\eta}^\varepsilon) c_{\varepsilon,K,\eta} a_{\varepsilon,K,\eta} \int_{|y|\le 1} h(y)dy,$$

with

$$a_{\varepsilon,K,\eta} = \inf_{-3K < x_1 < -3K+\eta, -K < x_2 < K, |x+y| \le 1+\eta} f_\varepsilon(y).$$

Hence the law of $X_{2K/v}^\varepsilon$ is lower-bounded by

$$\mathbb{P}(A_{K,\eta}^\varepsilon) c_{\varepsilon,K,\eta} a_{\varepsilon,K,\eta} \mathbf{1}_{|y|\le 1} dy,$$

for any $X_0^\varepsilon = x$ belonging to B, hence B is a petite set. □

3.2 Tightness of the invariant probability measure of X_t^ε

It follows from Theorem 3 that for any $\varepsilon \le \varepsilon_0$, X_t^ε possesses a unique invariant probability measure μ_ε. The aim of this subsection is to prove

Theorem 4. *For any sequence $\varepsilon_n \downarrow 0$, the sequence of invariant measures $\{\mu_{\varepsilon_n},\ n \ge 1\}$ is tight.*

The result will follow from the following statement

Proposition 3. *There exist two constants $c, C > 0$ such that for any $M > 0$, if $\varepsilon \le c/M$,*

$$\limsup_{t\to\infty} \mathbb{E}\left(|X_t^\varepsilon| \wedge M\right] \le C.$$

Let us first show how Theorem 4 follows from Proposition 3.

PROOF OF THEOREM 4. We deduce from the above Proposition that provided $\varepsilon \le c/M$,

$$\int_{\mathbb{R}^d} (|x| \wedge M)\mu_\varepsilon(dx) \le C,$$

and then also

$$\mu_\varepsilon(|x| > M) \le \frac{C}{M}.$$

Fix $\delta > 0$ arbitrarily small. Let us from now on fix $M \ge C/\delta$. Let n_0 be such that $\varepsilon_{n_0} \le c/M$. It follows from the above that for any $n \ge n_0$,

$$\mu_{\varepsilon_n}(|x| > M) \le \delta.$$

It is finally easy to find $M' \ge M$ such that

$$\mu_{\varepsilon_n}(|x| > M') \le \delta$$

fro any $1 \le n \le n_0$, hence the result. □

Now return to the

PROOF OF PROPOSITION 3. It follows readily from Lemma 3 that whenever $M < a/\varepsilon$,

$$\begin{aligned}
\mathbb{E}\left[|X_t^\varepsilon| \wedge M\right] &\le \mathbb{E}(|X_t^\varepsilon|) \\
&\le |X_0^\varepsilon| + Ct - c\int_0^t \mathbb{E}\left[|X_s^\varepsilon| \wedge (a/\varepsilon)\right] ds \\
&\le |X_0^\varepsilon| + Ct - c\int_0^t \mathbb{E}\left[|X_s^\varepsilon| \wedge M\right] ds.
\end{aligned}$$

Thanks to a classical comparison theorem for ODEs, this implies that

$$\mathbb{E}\left[|X_t^\varepsilon| \wedge M\right] \leq |X_0^\varepsilon| e^{-ct} + C\int_0^t e^{-c(t-s)} ds$$
$$= \left(|X_0^\varepsilon| - \frac{C}{c}\right) e^{-ct} + \frac{C}{c},$$

which implies the result.

□

3.3 Asymptotic analysis of the large time behavior of X_t^ε

We now want to analyze the large time behavior of X_t^ε, for small ε. We first show

Theorem 5. *As $\varepsilon \to 0$, $\mu^\varepsilon \Rightarrow \delta_{\bar{X}_\infty}$.*

Proof We consider the X_t^ε equation started with $X_0^\varepsilon \simeq \mu^\varepsilon$. Since the collection $\{\mu^{\varepsilon_n}\}_{n\geq 1}$ is tight, along a subsequence still denoted the same way, $\mu^{\varepsilon_n} \Rightarrow \mu^0$. It follows from Theorem 1 that $X_t^\varepsilon \to \overline{X}_t$ in probability in $D([0,T])$, where $\overline{X}_t$ solves the ODE

$$\frac{d\overline{X}_t}{dt} = L(\overline{X}_t) - v, \ \overline{X}_0 \simeq \mu^0.$$

But for any $f \in C_b(\mathbb{R}^d)$, $t \to \mathbb{E}f(X_t^\varepsilon)$ is a constant, so this is true in the limit, which implies that $\mu^0 = \delta_{\overline{X}_\infty}$. This shows that the whole collection $\mu^\varepsilon \Rightarrow \delta_{\overline{X}_\infty}$ as $\varepsilon \to 0$. □

Remark 2. *We expect from the above results that for small enough ε, the invariant measure μ^ε is close to ν^ε, which is the law of $\bar{X}_\infty + \sqrt{\varepsilon}\xi$, where $\xi \sim \mathcal{N}(0, \bar{S}^2)$. If we interpret μ^ε as the mass of X_t^ε for large t, this is not really correct, since for some large t, X_t^ε will make a large deviation from $\bar{X}_t$, see [3].*

Therefore, we prefer to interpret $\mu^\varepsilon(A)$ as

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \mathbf{1}_A(X_s^\varepsilon) ds.$$

We now want to give a precise description of μ^ε for small $\varepsilon > 0$. For that sake, let us introduce some notation. For any Borel set A and for all $\varepsilon > 0$, let $A_\varepsilon := \bar{X}_\infty + \sqrt{\varepsilon}A$ and denote by $\lambda_\infty = \mathcal{N}(0, \bar{S}^2)$ the invariant Gaussian distribution of the Ornstein-Uhlenbeck process U_t.

Theorem 6. *Consider the process X_t^ε, starting at time $t=0$ from $X_0^\varepsilon = \bar{X}_\infty$. For any $\delta > 0$ there exist $t_\delta > 0$ large enough such that for any $t \geq t_\delta$, there exists $\varepsilon_{t,\delta} > 0$ such that for all $\varepsilon \leq \varepsilon_{t,\delta}$, with a probability larger than $1-\delta$, the fraction of the time in the interval $[0,t]$ which X_s^ε spends in the set $\bar{X}_\infty + \sqrt{\varepsilon}A$ belongs to the interval $[\lambda_\infty(A)-\delta, \lambda_\infty(A)+\delta]$.*

Proof. We assume w.l.o.g. that ∂A has zero d–dimensional Lebesgue measure. Then, as $\varepsilon \to 0$, for any fixed $t > 0$,

$$\frac{1}{t}\int_0^t \mathbf{1}_{A_\varepsilon}(X_s^\varepsilon)ds = \frac{1}{t}\int_0^t \mathbf{1}_A(U_s^\varepsilon)ds \Rightarrow \frac{1}{t}\int_0^t \mathbf{1}_A(U_s)ds,$$

and

$$\frac{1}{t}\int_0^t \mathbf{1}_A(U_s)ds \to \lambda_\infty(A) \text{ in probability, as } t \to \infty.$$

Hence for all $\delta > 0$, there exists $t_\delta > 0$ such that for any $t \geq t_\delta$,

$$\mathbb{P}\left(\left|\frac{1}{t}\int_0^t \mathbf{1}_A(U_s)ds - \lambda_\infty(A)\right| > \delta\right) \leq \delta.$$

Define a test function $\varphi_\delta \in C_b(\mathbb{R}^d;[0,1])$ such that

$$\varphi_\delta(x) = \begin{cases} 1 & \text{if } |x-\lambda_\infty(A)| \leq \frac{\delta}{2}, \\ 0 & \text{if } |x-\lambda_\infty(A)| > \delta. \end{cases}$$

Let us now fix an arbitrary $t \geq t_{\delta/2}$. We have

$$\mathbb{E}\left[\varphi_\delta\left(\frac{1}{t}\int_0^t \mathbf{1}_A(U_s)ds\right)\right] \geq 1 - \frac{\delta}{2}.$$

Now there exists $\varepsilon_{t,\delta}$ such that if $\varepsilon < \varepsilon_{t,\delta}$,

$$\mathbb{E}\left[\varphi_\delta\left(\frac{1}{t}\int_0^t \mathbf{1}_A(U_s^\varepsilon)ds\right)\right] \geq 1 - \delta,$$

yielding

$$\mathbb{P}\left(\left|\frac{1}{t}\int_0^t \mathbf{1}_A(U_s^\varepsilon)ds - \lambda_\infty(A)\right| > \delta\right) \leq \delta,$$

where we remind that

$$\mathbf{1}_A(U_s^\varepsilon) = \begin{cases} 1, & \text{if } X_s^\varepsilon \in \bar{X}_\infty + \sqrt{\varepsilon}A, \\ 0, & \text{otherwise,} \end{cases}$$

implying the result. □

Acknowledgements The authors want to express their gratitude to Micheal Kopp, who introduced us to the model, and suggested the study of this small jumps limit, and to Nicolas Champagnat, whose clever question during the thesis defense of the first author, led us to the developments of the above section 3.

References

[1] N. Champagnat and S. Méléard, Polymorphic evolution sequence and evolutionary branching. *Probability Theory and Related Fields* **151**, 45–94, 2011.

[2] U. Dieckmann and R. Law, The dynamical theorey of coevolution: a derivation from stochastic ecological processes, *J. Math. Biol.* **34**, 579–612, 1996.

[3] M. Freidlin and A. Wentzell, *Random Perturbations of Dynamical Systems.* Springer, Berlin, Heidelberg, 2012.

[4] J. B. S. Haldane, A mathematical theory of natural and artificial selection. Part V: selection and mutation, *Proc. Camb. Philos. Soc.* **23**, 838âĂŞ 844, 1927.

[5] M. Kimura, On the probability of fixation of mutant genes in a population, *Genetics* **47** p. 713, 1962.

[6] M. Kopp and J. Hermisson The genetic basis of phenotypic adaptation II: The distribution of adaptive substitutions in the moving optimum model. *Genetics* **183**, 1453–1476, 2009.

[7] G. Malécot, Les processus stochastiques et la méthode des fonctions génératrices ou caractéristiques. *Publications de l'Institut de Statistiques de l'Université de Paris* **1**, 1–16, 1952.

[8] S. Matuszewski, J. Hermisson, and M. Kopp, Fisher's geometric model with a moving-optimum, *Evolution*, **68**, 2571–2588, 2014.

[9] S. P. Meyn and R. L. Tweedie, Stability of markovian processes III: Foster- Lyapunov criteria for continuous-time processes, *Advances in Applied Probability*, **25**, no. 3, pp. 518–548, 1993.

[10] E. Nassar and E. Pardoux, On the large behavior of an sde driven by a Poisson point process, *Journal of Applied Probability* (To appear), 2017.

[11] E. Pardoux, *Probabilistic Models of Population Evolution.* Springer, 2016.

[12] S. Peischl and M. Kirkpatrick, Establishment of new mutations in changing environments, *Genetics*, 2012.

[13] P. E. Protter, *Stochastic Integration and Differential Equations*. Heidelberg: Springer-Verlag, 2005.

[14] H. Uecker and J. Hermisson, On the fixation process of a beneficial mutation in a variable environment, *Genetics* **188**, 915–930, 2011.