

## METASTABLE BEHAVIOUR OF SMALL NOISE LÉVY-DRIVEN DIFFUSIONS

PETER IMKELLER<sup>1</sup> AND ILYA PAVLYUKEVICH<sup>1</sup>

**Abstract.** We consider a dynamical system in  $\mathbb{R}$  driven by a vector field  $-U'$ , where  $U$  is a multi-well potential satisfying some regularity conditions. We perturb this dynamical system by a Lévy noise of small intensity and such that the heaviest tail of its Lévy measure is regularly varying. We show that the perturbed dynamical system exhibits metastable behaviour *i.e.* on a proper time scale it reminds of a Markov jump process taking values in the local minima of the potential  $U$ . Due to the heavy-tail nature of the random perturbation, the results differ strongly from the well studied purely Gaussian case.

**Mathematics Subject Classification.** 60E07, 60F10.

Received September 12, 2007.

### INTRODUCTION

This paper addresses the rigorous mathematical description of the phenomenon of metastability in systems with big jumps. The picture we shall study may be outlined as follows. Let us consider a one-dimensional deterministic dynamical system driven by a vector field  $-U'(\cdot)$ , where  $U(\cdot)$  is a multi-well potential with some smoothness conditions and a certain increase rate at infinity. According to the initial conditions the deterministic trajectories of the dynamical system converge to the local minima of the potential  $U$  or stay in its local maxima. Obviously, no transition between different domains of attraction is possible.

The situation becomes different if the dynamical system is perturbed by (small) random noise whose presence allows transitions between the potential wells. However depending on the system's initial conditions and noise's properties, certain potential wells may be reached only on appropriately long time scales or stay unvisited. The phenomenon of metastability means, roughly speaking, that for different time scales and initial conditions the system may reach different local statistical equilibria.

The system's behaviour is determined by the type of random perturbation. Unquestionably, dynamical systems subject to small *Gaussian perturbations* have been studied most extensively. The main reference on this subject is the book [10] where the large deviations theory for the perturbed trajectories is established. The large deviations estimates allow to solve the first exit problem from the domain of attraction of a stable point. It turns out that the mean exit time is exponentially large in the small noise parameter, and its logarithmic rate is proportional to the height of the potential barrier the trajectories have to overcome. Thus for a multi-well dynamical system we obtain a series of exponentially non-equivalent time scales given by the wells' mean exit

---

*Keywords and phrases.* Lévy process, jump diffusion, heavy tail, regular variation, metastability, extreme events, first exit time, large deviations.

<sup>1</sup> Institut für Mathematik, Humboldt-Universität zu Berlin, Rudower Chaussee 25, 12489 Berlin, Germany;  
[imkeller@mathematik.hu-berlin.de](mailto:imkeller@mathematik.hu-berlin.de); [pavljuke@mathematik.hu-berlin.de](mailto:pavljuke@mathematik.hu-berlin.de)

times. Moreover, one can prove that the normalised exit times are exponentially distributed (see [2,7,21]), and thus have the property of lack of memory which is referred to as unpredictability in the physics literature.

In the simplest situation in which the potential  $U$  has only two wells of different depths, one can observe two statistically different regimes. First, if the time horizon is shorter than the exit time from the shallow well, the system cannot leave the well where it started, and therefore stays in the neighbourhood of the well's local minimum. Second, if the time horizon is longer than the exit time from the shallow well, the system has enough time to reach the deepest well from any starting point, and stays in the vicinity of the global minimum. In [15] the following metastability result is established. Namely, there is a time scale on which the dynamical system converges to a Markov two-state process with one absorbing state corresponding to the deep well. It is easy to notice that this particular time scale is given by the mean exit time from the shallow well. More general results for multidimensional diffusions can be found in [17] and [11].

There is a very close connection between the metastability of a Markov process with small noise and the spectral properties of its infinitesimal generator. It can be shown that exponentially small eigenvalues of the infinitesimal generator are expressed in terms of mean life times in the domains of attraction, and the corresponding eigenfunctions are close to constants on these domains [16]. On the other hand, the generator's eigenvalues can be calculated with the help of variational principles [3,4].

Recently, systems subject to *non-Gaussian perturbations* with big jumps attract more attention. Instantaneous transitions between remote states are referred to as extreme events and are observed in the dynamics of asset prices, climate and telecommunication systems etc. In the physics literature, non-Gaussian symmetric stable Lévy processes are used especially often, under the name of *Lévy flights*.

Our particular interest in this type of systems arose from the acquaintance with the papers [8,9] by Ditlevsen devoted to a statistical analysis of real paleo-climatic ice-core data describing the evolution of the temperature of the Northern hemisphere during the past 100 000 years, *i.e.* the last ice age. Numerous reconstructions indicate about 25 abrupt catastrophic climate changes, the so-called Dansgaard-Oeschger events. The simplest mathematical model describing these events is a one-dimensional Langevin equation for an overdamped particle in a double-well potential, driven by some type of noise. The stable minima of the potential correspond to a cold (stadial) and a warmer (excited) climate regimes. On the basis of statistical data analysis performed in [8], Ditlevsen discovered a strong Lévy stable component of a stability index  $\alpha \approx 1.75$  in the driving noise. The corresponding Langevin equation was studied analytically on a physical level of rigour in [9].

With the aim of starting a rigorous mathematical underpinning of the heuristics these papers provide on the level of their pathwise behaviour, in [13] we have tackled the study of random dynamical systems driven by small heavy-tailed Lévy noise, in the simplest case of symmetric stable processes. We obtain results on the law and the mean value of the first exit time of the resulting jump diffusions from a *single potential well* the boundary of which does not contain saddle points (the boundary is *non-characteristic*) by purely probabilistic methods. More precisely, we show that the mean value of the exit time increases like a power of the noise amplitude and does not depend on the depth of the potential well but rather on the distance between the local minimum and the domain's boundary. Normalised by its mean value, the exit time has a standard exponential law in the small noise limit. We refer the reader to related work by Godovanchuk [12] and Wentzell [20] on large deviations of Markov processes with big jumps.

In the present paper we extend our research in two directions. Firstly, we take some effort to generalise the methods of [13] to investigate in detail the behaviour of the jump-diffusions at saddle points of the potential not accessible in [13]. This opens the way to also give a detailed description of the diffusions' metastable dynamics in *multi-well* potentials. Secondly, we consider the most general case of heavy-tailed Lévy processes, namely Lévy processes with an arbitrary Gaussian variance and drift and a jump measure having regularly varying tails. We note, that our research in progress shows that for lighter, *e.g.* subexponential but not polynomially heavy tails, the pattern of metastable behaviour we discover here does not hold any longer.

The results of this paper may also be of interest for physicists. Indeed, in [6], a numerical analysis treatment of transitions of Lévy flights between the domains of attraction of a double-well potential was given for stability

indices  $\alpha \in [1, 2)$ . The conclusions are not fully consistent with our results. They could be improved using the rigorous estimates of the present paper for calibration of physical simulations.

1. OBJECT OF STUDY AND MAIN RESULT

Let  $(\Omega, \mathcal{F}, (\mathcal{F})_{t \geq 0}, \mathbf{P})$  be a filtered probability space. We assume that the filtration satisfies the usual hypotheses in the sense of [18], *i.e.*  $\mathcal{F}_0$  contains all the  $\mathbf{P}$ -null sets of  $\mathcal{F}$ , and is right continuous.

We consider solutions  $X^\varepsilon = (X_t^\varepsilon)_{t \geq 0}$  of the one-dimensional stochastic differential equation

$$X_t^\varepsilon(x) = x - \int_0^t U'(X_{s-}^\varepsilon(x)) \, ds + \varepsilon L_t, \quad x \in \mathbb{R}, \tag{1.1}$$

where  $L$  is a Lévy process and  $U$  is a potential function satisfying the following assumptions.

Assumptions on  $L$ :

**L1**  $L$  has a generating triplet  $(d, \nu, \mu)$  with a Gaussian variance  $d \geq 0$ , an arbitrary drift  $\mu \in \mathbb{R}$  and a Lévy measure  $\nu$  satisfying the usual condition  $\int_{\mathbb{R} \setminus \{0\}} \max\{y^2, 1\} \nu(dy) < \infty$ . For  $u \geq 1$  denote the tails of the Lévy measure  $\nu$

$$H_-(-u) = \int_{(-\infty, -u)} \nu(dy), \quad H_+(u) = \int_{(u, +\infty)} \nu(dy), \tag{1.2}$$

and  $H(u) = H_-(-u) + H_+(u)$ .

**L2** Assume,  $H_+(\cdot)$  is regularly varying at infinity, *i.e.*

$$H_+(u) = u^{-r} l(u), \quad u \rightarrow +\infty, \tag{1.3}$$

for some  $r > 0$  and a slowly varying function  $l$ . Recall that a positive Lebesgue measurable function  $l$  is slowly varying at infinity if  $\lim_{u \rightarrow +\infty} l(\lambda u)/l(u) = 1$  for any  $\lambda > 0$ . For example, positive constants, logarithms and iterated logarithms are slowly varying functions.

**L3** Assume that

$$\lim_{u \rightarrow +\infty} \frac{H_-(-u)}{H_+(u)} = \kappa \in (0, +\infty) \quad \text{or} \quad \limsup_{u \rightarrow +\infty} \frac{H_-(-u)}{H_+(u)} = \kappa = 0. \tag{1.4}$$

Assumptions on  $U$ :

**U1**  $U \in \mathcal{C}^1(\mathbb{R}) \cap \mathcal{C}^3([-K, K])$  for some  $K > 0$  large enough.

**U2**  $U$  has exactly  $n$  local minima  $m_i$ ,  $1 \leq i \leq n$ , and  $n - 1$  local maxima  $s_i$ ,  $1 \leq i \leq n - 1$ , enumerated in increasing order

$$-\infty = s_0 < m_1 < s_1 < m_2 < \dots < s_{n-1} < m_n < s_n = +\infty. \tag{1.5}$$

All extrema of  $U$  are non-degenerate, *i.e.*  $U''(m_i) > 0$ ,  $1 \leq i \leq n$ , and  $U''(s_i) < 0$ ,  $1 \leq i \leq n - 1$ .

**U3**  $|U'(x)| > c_1|x|^{1+c_2}$  as  $x \rightarrow \pm\infty$  for some  $c_1, c_2 > 0$ .

The class of Lévy processes  $L$  under consideration covers for example compound Poisson processes with heavy-tail jumps or stable Lévy processes with Lévy measure

$$\nu(dy) = (c_1 \mathbb{I}\{y < 0\} + c_2 \mathbb{I}\{y > 0\}) \frac{dy}{|y|^{1+\alpha}}, \quad \alpha \in (0, 2), \quad c_1 \geq 0, c_2 > 0. \tag{1.6}$$

We consider  $X^\varepsilon$  for small values of  $\varepsilon$ ,  $\varepsilon \downarrow 0$ .

Since the Lévy process  $L$  is a semimartingale, the stochastic differential equation (1.1) is well defined, see also [18] for the general theory. However, since the drift term  $U'$  is not globally Lipschitz we need to show the existence and uniqueness of the strong solution of (1.1) which can be done by a slight modification of the argument in Section 2 of [19]. Essentially one shows that solutions of (1.1) do not explode since the drift  $U'(\cdot)$  drives them to stay in some bounded set. Similar results for Gaussian diffusions can be found in [5].

Under assumptions on  $U$ , the underlying deterministic ( $\varepsilon = 0$ ) equation

$$x_t^0(x) = x - \int_0^t U'(x_s^0(x)) ds \tag{1.7}$$

has a unique solution for any initial value  $x \in \mathbb{R}$  and all  $t \geq 0$ . The local minima of  $U$  are stable attractors for the dynamical system  $x^0$ , i.e. if  $x \in (s_{i-1}, s_i)$ ,  $1 \leq i \leq n$ , then  $x_t^0(x) \rightarrow m_i$  as  $t \rightarrow \infty$ . It is clear that the deterministic solution  $x^0$  does not leave the domain of attraction where it started.

Our goal is to describe the phenomenon of *metastability* which roughly speaking consists in the existence of a time scale for which the system reminds of a jump process taking values in the set of stable attractors. We prove the following main theorem.

**Theorem 1.1.** *Let  $X^\varepsilon(x) = (X_t^\varepsilon(x))_{t \geq 0}$  be a solution of (1.1). If  $x \in (s_{i-1}, s_i)$ , for some  $i = 1, \dots, n$ , then for  $t > 0$*

$$X_{t/H(1/\varepsilon)}^\varepsilon(x) \rightarrow Y_t(m_i), \quad \varepsilon \downarrow 0, \tag{1.8}$$

*in the sense of finite-dimensional distributions, where  $Y = (Y_t)_{t \geq 0}$  is a Markov process on a state space  $\{m_1, \dots, m_n\}$  with the infinitesimal generator  $Q = (q_{ij})_{i,j=1}^n$ ,*

$$\begin{aligned} q_{ij} &= \frac{\kappa \mathbb{I}\{j < i\} + \mathbb{I}\{j > i\}}{1 + \kappa} \left| |s_{j-1} - m_i|^{-r} - |s_j - m_i|^{-r} \right|, \quad i \neq j, \\ -q_{ii} = q_i &= \sum_{j \neq i} q_{ij} = \frac{\kappa}{1 + \kappa} |s_{i-1} - m_i|^{-r} + \frac{1}{1 + \kappa} |s_i - m_i|^{-r}. \end{aligned} \tag{1.9}$$

Let us consider a particular example of equation (1.1), namely a symmetric  $\alpha$ -stable process  $L$  (Lévy flight) in a double-well potential. Let  $U$  satisfy the assumptions formulated above and let  $s_1 = 0$ . The process  $L$  has a generating triplet  $(0, \nu, 0)$  with a Lévy measure  $\nu(dy) = |y|^{-1-\alpha} dy$ ,  $y \neq 0$ ,  $\alpha \in (0, 2)$ .

Due to Theorem 1.1, the main features of the process  $X^\varepsilon$  in the small noise limit are retained by a Markov jump process, and on the time scale  $\alpha\varepsilon^{-\alpha}$  we obtain the following convergence in the sense of finite dimensional distributions:

$$X_{\alpha t/\varepsilon^\alpha}^\varepsilon(x) \rightarrow Y_t, \quad t > 0, \quad \varepsilon \downarrow 0, \tag{1.10}$$

where  $Y$  is a Markov process on the state space  $\{m_1, m_2\}$  with the following matrix as infinitesimal generator

$$Q = \begin{pmatrix} -|m_1|^{-\alpha} & |m_1|^{-\alpha} \\ |m_2|^{-\alpha} & -|m_2|^{-\alpha} \end{pmatrix} \quad \text{and} \quad Y_0 = \begin{cases} m_1, & \text{if } x < 0, \\ m_2, & \text{if } x > 0. \end{cases} \tag{1.11}$$

To compare the result obtained with its Gaussian counterpart we refer to [15], where this problem was first studied. Let us consider a Gaussian diffusion  $\hat{X}^\varepsilon$  which solves the equation

$$\hat{X}_t^\varepsilon(x) = x - \int_0^t U'(\hat{X}_s^\varepsilon(x)) ds + \varepsilon W_t, \tag{1.12}$$

where  $W$  is a standard Brownian motion. Since it is well known that in the Gaussian case the height of the potential barriers plays a crucial role, we assume that the left well is deeper, i.e.  $U(0) - U(m_1) > U(0) - U(m_2)$ .

This leads to the following meta-stable behaviour of  $\hat{X}^\varepsilon$ , see Theorem 2.1 in [15]. There exists a time scale  $\lambda^\varepsilon$  such that

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \ln \lambda^\varepsilon = 2(U(0) - U(m_2)) \tag{1.13}$$

and

$$\hat{X}_{t\lambda^\varepsilon}^\varepsilon(x) \rightarrow \hat{Y}_t, \quad \varepsilon \downarrow 0, \tag{1.14}$$

in the sense of finite dimensional distributions, where  $\hat{Y}$  is a Markov process on  $\{m_1, m_2\}$  with the infinitesimal matrix

$$\begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad \hat{Y}_0 = \begin{cases} m_1, & \text{if } x < 0, \\ m_2, & \text{if } x > 0. \end{cases} \tag{1.15}$$

As we see, the main difference between Lévy and Gaussian dynamics consists not only in different intrinsic time scales – polynomial vs. exponential, – but also in a qualitatively different limiting behaviour. In the heavy-tail case, the states of the limiting process are recurrent, whereas in the Gaussian case, the minimum of the deepest well is absorbing.

In general case, we can summarise the differences as follows. First, we see that the characteristic time scale is algebraic in  $\varepsilon$ . Second, the properties of the limiting process  $Y$  depend on sizes of the potential wells and not on their depths. Further, if  $\kappa > 0$ , the all states of  $Y$  are recurrent. The process  $Y$  has a unique absorbing state  $m_n$  (the local minimum of the right peripheral well) if and only if  $\kappa = 0$ , *i.e.* when the positive tail of  $L$  dominates.

The material is organised as follows. In Section 2 we decompose the Lévy process  $L$  into a small jump part and a compound Poisson part and study the small-jump dynamics of the process  $X^\varepsilon$ . This section presents some technical challenges. It consists of a pathwise analysis of a jump-diffusion with relatively small jumps in  $\varepsilon$ -dependent neighbourhoods of the stable and unstable attractors. In Section 3 we determine the asymptotic law of the first exit time from a single well and its mean value, in particular the asymptotic probabilities to transit to different wells. In Section 4 we establish that the laws of transition times between neighbourhoods of the wells’ minima are asymptotically exponential, and provide the corresponding transition probabilities. The main Theorem 1.1 is proved in Section 5.

## 2. ONE-WELL DYNAMICS OF THE SMALL JUMP COMPONENT

### 2.1. Exponential estimate for the small-jump component

For  $\rho > 0$  and  $0 < \varepsilon \leq 1$  consider the decomposition  $L = \xi^\varepsilon + \eta^\varepsilon$ , where the Lévy processes  $\xi^\varepsilon$  and  $\eta^\varepsilon$  have generating triplets  $(d, \nu_\xi^\varepsilon, \mu)$  and  $(0, \nu_\eta^\varepsilon, 0)$  respectively with

$$\nu_\xi^\varepsilon(\cdot) = \nu \left( \cdot \cap \left[ -\frac{1}{\varepsilon^\rho}, \frac{1}{\varepsilon^\rho} \right] \setminus \{0\} \right), \quad \nu_\eta^\varepsilon(\cdot) = \nu \left( \cdot \cap \mathbb{R} \setminus \left[ -\frac{1}{\varepsilon^\rho}, \frac{1}{\varepsilon^\rho} \right] \right). \tag{2.1}$$

The absolute value of jumps of the process  $\varepsilon\xi^\varepsilon$  does not exceed  $\varepsilon^{1-\rho}$ .

Thus the process  $\xi^\varepsilon$  has a Lévy measure with compact support, and the Lévy measure  $\nu_\eta^\varepsilon(\cdot)$  of  $\eta^\varepsilon$  is finite. Denote

$$\beta_\varepsilon = \nu_\eta^\varepsilon(\mathbb{R}) = \int_{\mathbb{R} \setminus [-\frac{1}{\varepsilon^\rho}, \frac{1}{\varepsilon^\rho}]} \nu(dy) = H \left( \frac{1}{\varepsilon^\rho} \right). \tag{2.2}$$

Then,  $\eta^\varepsilon$  is a compound Poisson process with intensity  $\beta_\varepsilon$ , and jumps distributed according to the law  $\beta_\varepsilon^{-1} \nu_\eta^\varepsilon(\cdot)$ .

Denote  $\tau_k^\varepsilon, k \geq 0$ , the jump times of  $\eta^\varepsilon$  with  $\tau_0^\varepsilon = 0$ . Let  $T_k^\varepsilon = \tau_k^\varepsilon - \tau_{k-1}^\varepsilon$  denote successive inter-jump periods, and  $W_k^\varepsilon = \eta_{\tau_k^\varepsilon}^\varepsilon - \eta_{\tau_{k-1}^\varepsilon}^\varepsilon$  the jump heights of  $\eta^\varepsilon$ . Then, the three processes  $(T_k^\varepsilon)_{k \geq 1}$ ,  $(W_k^\varepsilon)_{k \geq 1}$ , and  $(\xi^\varepsilon)_{t \geq 0}$  are

independent. Moreover,

$$\mathbf{P}(T_k^\varepsilon \geq u) = \int_u^\infty \beta_\varepsilon e^{-\beta_\varepsilon s} ds = e^{-\beta_\varepsilon u}, \quad u \geq 0, \quad \text{and} \quad \mathbf{E}T_k^\varepsilon = \frac{1}{\beta_\varepsilon}, \tag{2.3}$$

$$\mathbf{P}(W_k^\varepsilon < w) = \frac{\nu_\eta^\varepsilon(-\infty, w)}{\nu_\eta^\varepsilon(\mathbb{R})} = \frac{1}{\beta_\varepsilon} \int_{(-\infty, w)} \mathbb{I}\left\{|y| > \frac{1}{\varepsilon^\rho}\right\} \nu(dy), \quad w \in \mathbb{R}. \tag{2.4}$$

Between the arrival times of  $\eta^\varepsilon$  the process  $X^\varepsilon$  is driven by  $\varepsilon\xi^\varepsilon$ . The next lemma shows that on long time intervals  $\varepsilon\xi^\varepsilon$  does not essentially deviate from zero. Hence the dynamics of the process  $X^\varepsilon$  on the intervals between arrival times of the process  $\eta^\varepsilon$  can be seen as a small random perturbation of the underlying deterministic trajectory.

**Lemma 2.1.** *For any  $\rho \in (0, 1)$ , any  $\gamma \in (0, 1 - \rho)$  and  $q \in (0, 1 - \rho - \gamma)$  there is  $p_0 > 0$  and  $\varepsilon_0 > 0$  such that the inequality*

$$\mathbf{P}\left(\sup_{t \in [0, 1/\varepsilon^q]} |\varepsilon\xi_t^\varepsilon| \geq \varepsilon^\gamma\right) \leq \exp(-1/\varepsilon^p) \tag{2.5}$$

holds for all  $0 < \varepsilon \leq \varepsilon_0$  and  $0 < p \leq p_0$ .

*Proof.* Let  $\rho, \gamma$  and  $q$  be as in the statement of lemma. The Lévy measure of  $\varepsilon\xi^\varepsilon$  has compact support, hence the process  $\varepsilon\xi^\varepsilon$  has exponential moments. Moreover,  $\varepsilon\xi_t^\varepsilon - \mathbf{E}(\varepsilon\xi_t^\varepsilon)$  is a zero-mean martingale, so that

$$|\mathbf{E}(\varepsilon\xi_t^\varepsilon)| \leq \varepsilon t \left| \mu + \int_{1 \leq |y| \leq \varepsilon^{-\rho}} y \nu(dy) \right| = \varepsilon t m(\varepsilon). \tag{2.6}$$

Then Doob’s inequality for exponential functions of martingales yields for  $u > 0$ :

$$\mathbf{P}\left(\sup_{t \in [0, 1/\varepsilon^q]} \varepsilon\xi_t^\varepsilon \geq \varepsilon^\gamma\right) \leq e^{-u(\varepsilon^\gamma - 2\varepsilon^{1-q}m(\varepsilon))} \sup_{t \in [0, 1/\varepsilon^q]} \mathbf{E}e^{u\varepsilon\xi_t^\varepsilon}, \tag{2.7}$$

where the latter exponent is known from the Lévy–Hinchin representation. Let  $u = u(\varepsilon) = 1/\varepsilon^c$  for  $c = (1 - \rho + \gamma)/2$ . Then with help of some algebra we obtain for all  $0 < p \leq p_0 = (c - \gamma)/2$  that

$$\sup_{t \in [0, 1/\varepsilon^q]} \varphi(u(\varepsilon), \varepsilon, t) = \sup_{t \in [0, 1/\varepsilon^q]} \left[ \ln \mathbf{E} \exp(u(\varepsilon)\varepsilon\xi_t^\varepsilon) + 2u(\varepsilon)m\varepsilon^{1-q} - \varepsilon^\gamma u(\varepsilon) \right] \leq -\frac{1}{\varepsilon^p}, \quad \varepsilon \downarrow 0. \tag{2.8}$$

The inequality for the infimum is proved analogously. □

### 2.2. Dynamics on a compact interval $a > -\infty$

Our goal is to study the one-well dynamics of the small-jump process  $x^\varepsilon$  and its unperturbed counterpart  $x^0$ ,

$$x_t^\varepsilon(x) = x - \int_0^t U'(x_{s-}^\varepsilon) ds + \varepsilon\xi_t^\varepsilon. \tag{2.9}$$

For definiteness we assume that the well’s minimum is located at the origin and thus the corresponding domain of attraction for  $x^0$  is  $(a, b)$ ,  $-\infty < a < 0 < b < +\infty$ , if the well is inner, and  $(-\infty, b)$  if it is peripheral. In the first case, we also assume that  $a$  and  $b$  are non-degenerate local maxima of  $U$ . In the second case,  $b$  is a non-degenerate local maximum and  $U'(x)$  increases to infinity faster than linearly as  $x \rightarrow -\infty$ . Denote the critical point curvatures as  $U''(0) = M_0 > 0$ ,  $U''(b) = M_b > 0$  and  $U''(a) = M_a < 0$  (when defined). For  $\gamma > 0$

and  $t \geq 0$  we introduce an event

$$\mathcal{E}_t = \left\{ \omega : \sup_{s \in [0, t]} |\varepsilon \xi_s^\varepsilon| \leq \varepsilon^{4\gamma} \right\}. \tag{2.10}$$

We prove the following estimate.

**Proposition 2.2.** *For any  $\gamma > 0$ , any  $d > 0$  there is  $\varepsilon_0 > 0$  such that for  $0 < \varepsilon \leq \varepsilon_0$  the inequality*

$$\sup_{s \in [0, t]} |x_s^\varepsilon(x) - x_s^0(x)| \leq d\varepsilon^{2\gamma} \tag{2.11}$$

holds a.s. on the event  $\mathcal{E}_t$  uniformly for  $t \geq 0$  and  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$ .

The proof will follow directly from the estimates of Sections 2.2.1 and 2.2.2 below. It is based on the study of the representation of the process  $x^\varepsilon$  in powers of  $\varepsilon$

$$x_t^\varepsilon(x) = x_t^0(x) + \varepsilon Z_t^\varepsilon(x) + R_t^\varepsilon(x), \quad t \geq 0, \tag{2.12}$$

where  $Z^\varepsilon$  is the first approximation of  $x^\varepsilon$  satisfying the stochastic differential equation

$$Z_t^\varepsilon(x) = - \int_0^t U''(x_s^0(x)) Z_{s-}^\varepsilon(x) ds + \xi_t^\varepsilon, \tag{2.13}$$

and the remainder  $R^\varepsilon(x)$  is the absolutely continuous function starting at 0 and satisfying the integral equation

$$R_t^\varepsilon(x) = \int_0^t [-U'(x_s^0(x) + \varepsilon Z_{s-}^\varepsilon(x) + R_s^\varepsilon(x)) + U'(x_s^0(x)) + U''(x_s^0(x))\varepsilon Z_{s-}^\varepsilon(x)] ds. \tag{2.14}$$

The dynamics if the summanden in (2.12) in the limit of small  $\varepsilon$  mainly depends on the geometry of the potential and is quite different in the neighbourhoods of the stable point 0 and unstable interval boundaries (it suffices to consider only one unstable point, say  $x = a$ ). The following properties of the potential  $U$  will be extensively used:

- (1) The deterministic trajectories  $x_t^0(x)$ ,  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$  converge to 0 as  $t \rightarrow \infty$  due to the property  $xU'(x) > 0$  for  $x \neq a, b, 0$ .
- (2) The curvature of the potential at  $x = a, b$  is negative. In a small neighbourhood of  $a$  we have  $U(x) = U(a) - M_a \frac{(x-a)^2}{2} + o((x-a)^2)$ . Consequently  $x^0$  behaves there like  $a + e^{M_a t}$ , and the dynamics of  $x^\varepsilon$  reminds of the dynamics of an inverted process of Ornstein-Uhlenbeck type.
- (3) The curvature of the potential at  $x = 0$  is positive. In small neighbourhoods of 0 we have  $U(x) = U(0) + M_0 \frac{x^2}{2} + o(x^2)$ . Consequently  $x^0$  decays there like  $e^{-M_0 t}$ , and the dynamics of  $x^\varepsilon$  reminds of the dynamics of a process of Ornstein-Uhlenbeck type.

From now on, let  $\gamma > 0$  be fixed. Using assumptions on  $U$ , for technical reasons we fix some small  $\delta$ ,  $0 < \delta < \min\{|a|, b\}$ , and consider  $\delta$ -neighbourhoods of the critical points  $a, 0$  and  $b$  with the following properties:

- (1) there are some  $0 < m_1^a \leq M_a \leq m_2^a$ ,  $\frac{m_2^a}{m_1^a} < \frac{3}{2}$ , such that if  $a \leq x \leq a + \delta$  then  $m_1^a(x-a) \leq -U'(x) \leq m_2^a(x-a)$  and  $-U'(\cdot)$  is monotone increasing in  $x \in [a, a + \delta]$  (analogous estimates hold in  $\delta$ -neighbourhood of  $b$ );
- (2) there are some  $0 < m_1^0 < m_2^0$  such that the inequality  $m_1^0 < \inf_{|x| < \delta} U''(x) \leq \sup_{|x| < \delta} U''(x) < m_2^0$  holds.

For  $\varepsilon$  such that  $0 < \varepsilon^\gamma < \delta$  and for  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$ , consider the first entrance time of  $x^0(x)$  into the interval  $[a + \delta, b - \delta]$

$$t_\varepsilon(x) = \begin{cases} \int_x^{a+\delta} \frac{dy}{-U'(y)}, & \text{if } x \in [a + \varepsilon^\gamma, a + \delta], \\ \int_{b-\delta}^x \frac{dy}{U'(y)}, & \text{if } x \in [b - \delta, b - \varepsilon^\gamma], \\ 0, & \text{if } x \in [a + \delta, b - \delta]. \end{cases} \tag{2.15}$$

Also define the time period

$$\hat{T} = \max \left\{ \int_{a+\delta}^{-\delta} \frac{dy}{-U'(y)}, \int_{\delta}^{b-\delta} \frac{dy}{U'(y)} \right\}. \tag{2.16}$$

$\hat{T}$  has the property that for all  $x \in [a + \delta, b - \delta]$  and  $t \geq \hat{T}$ ,  $|x_t^0(x)| \leq \delta$ , *i.e.* after  $\hat{T}$  the trajectory of  $x^0(x)$  is within a  $\delta$ -neighbourhood of the stable point 0.

2.2.1. *Estimates on  $Z^\varepsilon$*

The solution to equation (2.13) is explicitly given by

$$Z_t^\varepsilon(x) = \xi_t^\varepsilon - \int_0^t \xi_s^\varepsilon U''(x_s^0(x)) e^{-\int_s^t U''(x_u^0(x)) du} ds, \tag{2.17}$$

and hence for  $t \geq 0$  and  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$  we have

$$\sup_{s \in [0, t]} |Z_s^\varepsilon(x)| \leq \left( 1 + \int_0^t |U''(x_s^0(x))| e^{-\int_s^t U''(x_u^0(x)) du} ds \right) \sup_{s \in [0, t]} |\xi_s^\varepsilon|. \tag{2.18}$$

We distinguish three cases:  $x \in [a + \delta, b - \delta]$ ,  $x \in [a + \varepsilon^\gamma, a + \delta]$  and  $x \in [b - \delta, b - \varepsilon^\gamma]$ . The case  $x \in [a + \delta, b - \delta]$  was considered in [13], Lemma 5, where for some  $C_1 > 2$  we established the estimate  $\sup_{s \in [0, t]} |Z_s^\varepsilon(x)| \leq C_1 \sup_{s \in [0, t]} |\xi_s^\varepsilon|$ .

Let now  $x \in [a + \varepsilon^\gamma, a + \delta]$  and  $t \leq t_\varepsilon(x)$ . Then using the fact that  $U''(x_t^0(x)) < 0$  we obtain:

$$1 + \int_0^t |U''(x_s^0(x))| e^{-\int_s^t U''(x_u^0(x)) du} ds = \frac{U'(x_t^0(x))}{U'(x)}. \tag{2.19}$$

For any  $t \geq 0$  we use (2.18) and (2.19) to obtain

$$1 + \int_0^t |U''(x_s^0(x))| e^{-\int_s^t U''(x_u^0(x)) du} ds \leq 1 - \frac{U'(x_t^0(x))}{U'(x_{t_\varepsilon(x) \wedge t}^0(x))} + \frac{U'(x_t^0(x))}{U'(x)} + C_1. \tag{2.20}$$

Note that for  $0 < \varepsilon \leq \varepsilon_0$  for some  $\varepsilon_0$  small enough and depending on  $U, a, b, \gamma$  and  $\delta$  we have

$$1 - \frac{U'(x_t^0(x))}{U'(x_{t_\varepsilon(x) \wedge t}^0(x))} + \frac{U'(x_t^0(x))}{U'(x)} \leq \begin{cases} \frac{U'(x_t^0(x))}{U'(x)} \leq \frac{m_2^a |a+\delta|}{m_1^a (a-x)} \leq \frac{C_2}{\varepsilon^\gamma}, & 0 \leq t \leq t_\varepsilon(x), \\ 1 + \max_{y \in [a, b]} |U'(y)| \left( \frac{1}{|U'(a+\delta)|} + \frac{1}{|U'(x)|} \right) \leq \frac{C_3}{\varepsilon^\gamma}, & t \geq t_\varepsilon(x), \end{cases} \tag{2.21}$$

and hence for some  $C_4 > 0$  we get the final estimate:

$$\sup_{s \in [0, t]} |Z_s^\varepsilon(x)| \leq \frac{C_4}{\varepsilon^\gamma} \sup_{s \in [0, t]} |\xi_s^\varepsilon|. \tag{2.22}$$

The case  $x \in [b - \delta, b - \varepsilon^\gamma]$  is treated analogously, and thus for some  $C_Z > 0$  we obtain that  $\sup_{s \in [0, t]} |Z_s^\varepsilon(x)| \leq C_Z \varepsilon^{3\gamma}$  on  $\mathcal{E}_t$ . □



2.2.2. Estimates on  $R^\varepsilon$

To estimate the remainder term  $R^\varepsilon$  we need finer smoothness properties of the potential  $U$ . However, it was shown in Lemma 6 in [13] that this restriction only has to hold locally. Namely, we assume in this section that there is  $C > 0$  such that the inequality  $\sup_{s \in [0, t]} |R_s^\varepsilon(x)| \leq C$  holds on the event  $\mathcal{E}_t$  uniformly for  $t \geq 0$  and  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$ . In other words, the small jump process  $x_s^\varepsilon(x)$ ,  $s \in [0, t]$ , with initial state confined to  $[a + \varepsilon^\gamma, b - \varepsilon^\gamma]$ , stays bounded by a deterministic constant  $K$  on the set  $\mathcal{E}_t$ ,  $t \geq 0$ . Therefore, in the small noise limit, only local properties of  $U$  are relevant to our analysis.

As in the case of estimates of  $Z^\varepsilon$ , the behaviour of  $R^\varepsilon(x)$  for  $x \in [a + \delta, b - \delta]$  was studied in detail in our paper [13]. Let us consider the asymptotics of  $R^\varepsilon(x)$  for  $x$  being close to the saddle points.

For definiteness, we consider the case  $x \in [a + \varepsilon^\gamma, a + \delta]$ . Assume also that the final estimate on  $Z^\varepsilon$  from the previous section holds for  $0 < \varepsilon \leq \varepsilon_1$  with constant  $C_Z$ . We show that on the event  $\mathcal{E}_t$  the inequality

$$\sup_{s \in [0, t \wedge t_\varepsilon(x)]} |R_s^\varepsilon(x)| \leq C_1 \varepsilon^{3\gamma} \tag{2.23}$$

holds uniformly over  $x$ .

Consider the integral equation (2.14), and let  $f$  denote the integrand

$$f(R, x^0, \varepsilon Z) = -U'(x^0 + \varepsilon Z + R) + U'(x^0) + U''(x^0)(\varepsilon Z). \tag{2.24}$$

Let the constant  $K$  from Assumption **U1** be bigger than  $C$ . We write the Taylor expansion for the integrand  $f$  with some  $|q| \leq K$ :

$$f(R, x^0, \varepsilon Z) = -U''(x^0)R - \frac{U^{(3)}(q)}{2}(R + \varepsilon Z)^2. \tag{2.25}$$

Since  $U \in \mathcal{C}^3$ ,  $|U^{(3)}|$  is bounded, say by  $L$ , on  $[-K, K]$ . Using the inequality  $(R + \varepsilon Z)^2 \leq 2(R^2 + \varepsilon^2 Z^2)$  we obtain that for  $t \geq 0$ ,

$$-U''(x_t^0(x))R_t^\varepsilon - L(R_t^\varepsilon)^2 - A^2 \varepsilon^{6\gamma} < f(R_t^\varepsilon(x), x_t^0(x), \varepsilon Z_{t-}^\varepsilon(x)) < -U''(x_t^0(x))R_t^\varepsilon + L(R_t^\varepsilon)^2 + A^2 \varepsilon^{6\gamma}, \tag{2.26}$$

on the event  $\mathcal{E}_t$ , with  $A^2 = 2C_Z^2 L$ .

Together with (2.14) consider the Riccati equation

$$p_t^\varepsilon(x) = \int_0^t (m_2^a p_s^\varepsilon + L(p_s^\varepsilon)^2 + A^2 \varepsilon^{6\gamma}) ds, \quad 0 \leq t \leq t_\varepsilon(x), \tag{2.27}$$

whose solution can be found explicitly:

$$p_t^\varepsilon = A^2 \varepsilon^{6\gamma} \frac{e^{t\lambda^\varepsilon} - e^{-t\lambda^\varepsilon}}{(m_2^a + \lambda^\varepsilon)e^{-t\lambda^\varepsilon} - (m_2^a - \lambda^\varepsilon)e^{t\lambda^\varepsilon}} \quad \text{with} \quad \lambda^\varepsilon = \sqrt{(m_2^a)^2 - 4LA^2 \varepsilon^{6\gamma}}. \tag{2.28}$$

The function  $p_t^\varepsilon$  is non-negative monotonically increasing function, which starts at 0 and has a singularity at

$$t^*(\varepsilon) = \frac{1}{2\lambda^\varepsilon} \ln \left( \frac{m_2^a + \lambda^\varepsilon}{m_2^a - \lambda^\varepsilon} \right). \tag{2.29}$$

However, one easily sees, that for  $\varepsilon \downarrow 0$

$$t_\varepsilon(x) \leq t_\varepsilon = \frac{1}{m_1^a} \int_{a+\varepsilon^\gamma}^{a+\delta} \frac{dy}{|a-y|} = \frac{\gamma |\ln \varepsilon|}{m_1^a} (1 + \mathcal{O}(|\ln \varepsilon|^{-1})) < \frac{3\gamma |\ln \varepsilon|}{m_2^a} (1 + \mathcal{O}(|\ln \varepsilon|^{-1})) \leq t^*(\varepsilon), \tag{2.30}$$

and  $p_t^\varepsilon$  is well defined on the time interval under consideration.

The estimate (2.23) will follow from two statements:

- a)  $R_t^\varepsilon(x) \leq p_t^\varepsilon$  for  $0 \leq t \leq t_\varepsilon(x)$ ,
- b)  $p_t^\varepsilon \leq C_1 \varepsilon^{3\gamma}$  for  $0 \leq t \leq t_\varepsilon(x)$ .

To show a) we note that at the starting point  $t = 0$ ,

$$D^+ R_t^\varepsilon(x) \Big|_{t=0} = \lim_{h \downarrow 0} \frac{R_h^\varepsilon(x) - 0}{h} = 0 < A^2 \varepsilon^{6\gamma} = \dot{p}_t^\varepsilon(x) \Big|_{t=0}, \quad (2.31)$$

consequently it follows from the continuity of  $R^\varepsilon$  and  $p^\varepsilon$  that  $p_t^\varepsilon > R_t^\varepsilon$  for at least positive and small  $t$ . Assume there exists  $\tau = \inf\{t > 0 : R_t^\varepsilon(x) = p_t^\varepsilon\}$  such that  $\tau \leq t_\varepsilon(x)$ . At the point  $\tau$  the left derivative of  $R^\varepsilon(x)$  is necessarily not less than the derivative of  $p^\varepsilon$  which leads to the following contradiction:

$$\begin{aligned} D^- R_t^\varepsilon(x) \Big|_{t=\tau} &= \lim_{h \downarrow 0} \frac{R_\tau^\varepsilon(x) - R_{\tau-h}^\varepsilon(x)}{h} = f(R_\tau^\varepsilon(x), x_\tau^0(x), Z_{\tau-}^\varepsilon(x)) \geq \dot{p}_t^\varepsilon \Big|_{t=\tau} = m_2^a p_\tau^\varepsilon + L(p_\tau^\varepsilon)^2 + A^2 \varepsilon^{6\gamma}, \\ f(R_\tau^\varepsilon(x), x_\tau^0(x), Z_{\tau-}^\varepsilon(x)) &= f(p_\tau^\varepsilon, x_\tau^0(x), Z_{\tau-}^\varepsilon(x)) < m_2^a p_\tau^\varepsilon + L(p_\tau^\varepsilon)^2 + A^2 \varepsilon^{6\gamma}. \end{aligned} \quad (2.32)$$

To prove b), we use the inequality  $\sup_{t \in [0, t_\varepsilon(x)]} p_t^\varepsilon \leq p_{t_\varepsilon}^\varepsilon \leq p_{t_\varepsilon}^\varepsilon$ , the formula (2.28), and the asymptotics (2.30) for  $t_\varepsilon$  to obtain the estimate

$$p_{t_\varepsilon}^\varepsilon \leq A^2 \varepsilon^{6\gamma} \frac{c_1 \varepsilon^{-\gamma \frac{m_2^a}{m_1^a}}}{c_2 \varepsilon^{\gamma \frac{m_2^a}{m_1^a}} - c_3 \varepsilon^{\gamma(6 - \frac{m_2^a}{m_1^a})}} \leq C_1 \varepsilon^{3\gamma}, \quad (2.33)$$

on the event  $\mathcal{E}_t$  for some positive  $c_1, c_2, c_3$  and  $C_1$ .

The proof of the lower bound in (2.23) is analogous. The estimate  $\sup_{s \in [0, t]} |R_s^\varepsilon(x)| \leq C_R \varepsilon^{3\gamma}$  for all  $t \geq 0$ ,  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$  and some  $C_R > 0$  follows from (2.23) and slightly modified estimates from Lemmas 7 and 8 in [13].

### 2.3. Dynamics on unbounded interval, $a = -\infty$ . Return from infinity

In this section we show that with high probability the process  $x^\varepsilon(x)$  reaches some fixed compact neighbourhood of the origin in finite time.

For  $M > 0$  large enough and  $x \leq -M$  define a stopping time  $\tau_x = \inf\{t \geq 0 : x_t^\varepsilon(x) \geq -M\}$  and a return time  $T_M = \int_{-\infty}^{-M+1} \frac{dv}{c_1 |v|^{1+c_2}}$ .

**Lemma 2.3.** *There is  $M > 0$  such that  $\tau_x \leq T_M$  a.s. on the event  $\mathcal{E}_{T_M}$  for all  $x \leq -M$ .*

*Proof.* Recall that due to Assumption **U3** there is  $N > 0$  such that  $-U'(x) > c_1 |x|^{1+c_2}$ , for some  $c_1, c_2 > 0$  and  $x \leq -N$ . Additionally, we assume that  $N$  is sufficiently large, so that for any  $x < -N$

$$-|x|^{1+c_2} + |x + \frac{1}{2}|^{1+c_2} + \frac{1}{4}(1+c_2)|x|^{c_2} = -\frac{1}{4}(1+c_2)|x|^{c_2} + o(|x|^{c_2}) < 0. \quad (2.34)$$

We compare  $x^\varepsilon(x)$  with the solution of the SDE

$$v_t^\varepsilon(v) = v + c_1 \int_0^t |v_{s-}^\varepsilon|^{1+c_2} ds + \varepsilon \xi_t^\varepsilon, \quad t \geq 0. \quad (2.35)$$

Let  $M > N$  and  $x \leq -M$ . Then a comparison argument similar to those in Section 2.2.2 a) yields, that for  $v < x < -M$ ,  $v_t^\varepsilon(v) < x_t^\varepsilon(x)$  a.s. on  $t \in [0, \tau_x)$ , i.e. the random trajectory  $x^\varepsilon$  does not deviate to  $-\infty$  faster than  $v^\varepsilon$ .

Let now  $T_M = \int_{-\infty}^{-M+1} \frac{dv}{c_1 |v|^{1+c_2}}$  be such that  $v_{T_M}^0(-M) \leq -N$ . The statement of the lemma will follow if we show that on the event  $\mathcal{E}_{T_M}$  the inequality  $\sup_{t \in [0, T_M]} |v_t^\varepsilon(v) - v_t^0(v)| \leq 1$  holds a.s. uniformly over  $v \leq -M$ .

To prove the latter inequality, similarly to Section 2.2 consider the representation  $v_t^\varepsilon(v) = v_t^0(v) + \varepsilon w_t^\varepsilon(v) + r_t^\varepsilon(v)$  with

$$\begin{aligned} v_t^0(v) &= v + c_1 \int_0^t |v_s^0|^{1+c_2} ds, \\ w_t^\varepsilon(v) &= \xi_t^\varepsilon - c_1(1+c_2)|v_t^0(v)|^{1+c_2} \int_0^t \xi_{s-}^\varepsilon \frac{ds}{|v_s^0(v)|}, \\ r_t^\varepsilon(v) &= c_1 \int_0^t (|v_s^0(v) + \varepsilon w_{s-}^\varepsilon(v) + r_{s-}^\varepsilon(v)|^{1+c_2} - |v_s^0(v)|^{1+c_2} + (1+c_2)|v_s^0(v)|^{c_2} \varepsilon w_{s-}^\varepsilon(v)) ds. \end{aligned} \tag{2.36}$$

To estimate  $w^\varepsilon$  we recall equations (2.18) and (2.19) and immediately get that  $\sup_{[0, T_M]} |w_t^\varepsilon(v)| \leq 2 \sup_{[0, T_M]} |\xi_t^\varepsilon| \leq 2\varepsilon^{4\gamma}$  for  $v \leq -M$ .

The estimate  $\sup_{t \in [0, T_M]} |r_t^\varepsilon| < 3/4$  is obtained by a comparison argument used in Section 2.2.2 a). □

### 2.4. Final estimate for $|x_t^\varepsilon - x_t^0|$

**Lemma 2.4.** *Let  $x^\varepsilon(x)$  and  $x^0(x)$  satisfy (2.9), and let  $T(\varepsilon)$  be a non-negative random variable. Then for any  $\gamma > 0, d > 0, q > 0$  and any integer  $k \geq 1$  the following inclusion holds for  $\varepsilon$  small enough:*

$$\left. \begin{aligned} &\left\{ \sup_{t \in [0, T(\varepsilon)]} |x_t^\varepsilon(x) - x_t^0(x)| \geq d\varepsilon^{2\gamma} \right\}, \quad \text{if } a > -\infty, \\ &\left\{ \sup_{t \in [0, \tau_x \wedge T(\varepsilon)]} x_t^\varepsilon(x) \geq -M + 1 \right\} \\ &\cup \left\{ \sup_{t \in [\tau_x \wedge T(\varepsilon), T(\varepsilon)]} |x_t^\varepsilon(x) - x_{t-\tau_x \wedge T(\varepsilon)}^0(x_{\tau_x}^\varepsilon(x))| \geq d\varepsilon^{2\gamma} \right\}, \\ &\text{if } a = -\infty. \end{aligned} \right\} \subseteq \{T(\varepsilon) \geq k/\varepsilon^q\} \cup \bigcup_{j=0}^{k-1} \left\{ \sup_{t \in [\frac{j}{\varepsilon^q}, \frac{j+1}{\varepsilon^q}]} |\varepsilon \xi_t^\varepsilon - \varepsilon \xi_{\frac{j}{\varepsilon^q}}^\varepsilon| \geq \varepsilon^{4\gamma} \right\} \tag{2.37}$$

*Proof.* Let  $a > -\infty$ . For any  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$  we have the obvious inclusion

$$\left\{ \sup_{t \in [0, T(\varepsilon)]} |x_t^\varepsilon(x) - x_t^0(x)| \geq d\varepsilon^{2\gamma} \right\} \subseteq \left\{ T(\varepsilon) \geq k/\varepsilon^q \right\} \cup \left\{ \sup_{t \in [0, k/\varepsilon^q]} |x_t^\varepsilon(x) - x_t^0(x)| \geq d\varepsilon^{2\gamma} \right\}. \tag{2.38}$$

Further we note that for  $\varepsilon$  small enough and  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$ , the estimate  $t_\varepsilon(x) + \hat{T} \ll 1/\varepsilon^q$  holds, and thus due to the inequality  $|x_t^0(x)| \leq |x|e^{-m_1 t}$  which holds in a  $\delta$ -neighbourhood of 0, we obtain that  $|x_{1/\varepsilon^q}^0(x)| \leq \frac{d}{3}\varepsilon^{2\gamma}$  for all  $x \in [a + \varepsilon^\gamma, b - \varepsilon^\gamma]$ . Consequently,  $|x_{1/\varepsilon^q}^\varepsilon(x)| \leq \frac{2d}{3}\varepsilon^{2\gamma}$  under the condition that  $|x_t^\varepsilon(x) - x_t^0(x)| \leq \frac{d}{3}\varepsilon^{2\gamma}$  for all  $t \in [0, 1/\varepsilon^q]$ . Then with help of Proposition 2.2 we obtain the following chain of inclusions and prove the

statement of the lemma in case of finite interval:

$$\begin{aligned}
 & \left\{ \sup_{t \in [0, k/\varepsilon^q]} |x_t^\varepsilon(x) - x_t^0(x)| \geq d\varepsilon^{2\gamma} \right\} \\
 & \subseteq \bigcup_{j=0}^{k-1} \left[ \bigcap_{i=1}^{j-1} \left\{ \sup_{t \in [\frac{i-1}{\varepsilon^q}, \frac{i}{\varepsilon^q}] } |x_t^\varepsilon(x) - x_{t-\frac{i-1}{\varepsilon^q}}^0(x_{\frac{i-1}{\varepsilon^q}}^\varepsilon)| < \frac{d}{3}\varepsilon^{2\gamma} \right\} \cap \left\{ \sup_{t \in [\frac{j}{\varepsilon^q}, \frac{j+1}{\varepsilon^q}] } |x_t^\varepsilon(x) - x_{t-\frac{j}{\varepsilon^q}}^0(x_{\frac{j}{\varepsilon^q}}^\varepsilon)| \geq d\varepsilon^{2\gamma} \right\} \right] \\
 & \subseteq \left\{ \sup_{t \in [0, \frac{1}{\varepsilon^q}]} |x_t^\varepsilon(x) - x_t^0(x)| \geq d\varepsilon^{2\gamma} \right\} \cup \bigcup_{j=1}^{k-1} \left\{ \sup_{t \in [\frac{j}{\varepsilon^q}, \frac{j+1}{\varepsilon^q}]} |x_t^\varepsilon(y) - x_{t-\frac{j}{\varepsilon^q}}^0(y)| \geq d\varepsilon^{2\gamma} \text{ for all } |y| \leq \frac{2d}{3}\varepsilon^{2\gamma} \right\} \\
 & \subseteq \bigcup_{j=0}^{k-1} \left\{ \sup_{t \in [\frac{j}{\varepsilon^q}, \frac{j+1}{\varepsilon^q}]} |\varepsilon \xi_t^\varepsilon - \varepsilon \xi_{j/\varepsilon^q}^\varepsilon| \geq \varepsilon^{4\gamma} \right\}.
 \end{aligned} \tag{2.39}$$

Let  $a = -\infty$ . Since  $\tau_x = 0$  for  $x \in [-M, b - \varepsilon^\gamma]$ , we only have to consider the case  $x \leq -M$ . Then using the estimate  $T_M < 1/\varepsilon^q$  and Lemma 2.3 we obtain

$$\begin{aligned}
 & \left\{ \sup_{t \in [0, \tau_x \wedge T(\varepsilon)]} x_t^\varepsilon(x) \geq -M + 1 \right\} \cup \left\{ \sup_{t \in [\tau_x \wedge T(\varepsilon), T(\varepsilon)]} |x_t^\varepsilon(x) - x_{t-\tau_x}^0(x_{\tau_x}^\varepsilon(x))| \geq d\varepsilon^{2\gamma} \right\} \\
 & \subseteq \mathcal{E}_{1/\varepsilon^q}^c \cup \{T(\varepsilon) \geq k/\varepsilon^q\} \cup \left\{ \sup_{t \in [0, \tau_x \wedge \frac{k}{\varepsilon^q}]} x_t^\varepsilon(x) \geq -M + 1, \mathcal{E}_{1/\varepsilon^q} \right\} (= \emptyset) \\
 & \cup \left\{ \sup_{t \in [0, \tau_x \wedge \frac{k}{\varepsilon^q}]} x_t^\varepsilon(x) < -M + 1, \sup_{t \in [\tau_x \wedge \frac{k}{\varepsilon^q}, \frac{k}{\varepsilon^q}]} |x_t^\varepsilon(x) - x_{t-\tau_x}^0(x_{\tau_x}^\varepsilon(x))| \geq d\varepsilon^{2\gamma}, \mathcal{E}_{1/\varepsilon^q} \right\} \\
 & \subseteq \mathcal{E}_{1/\varepsilon^q}^c \cup \{T(\varepsilon) \geq k/\varepsilon^q\} \cup \left\{ \sup_{t \in [\frac{1}{\varepsilon^q}, \frac{k}{\varepsilon^q}]} |x_t^\varepsilon(y) - x_{t-\frac{1}{\varepsilon^q}}^0(y)| \geq d\varepsilon^{2\gamma} \text{ for all } |y| \leq 2d\varepsilon^{2\gamma}/3 \right\}.
 \end{aligned} \tag{2.40}$$

We finish the proof applying the finite interval argument to the last event in the previous formula.  $\square$

### 3. EXIT FROM A SINGLE WELL

For  $i = 1, \dots, n$ , consider the wells of the potential  $U$  with local minima at  $m_i$ . For  $\varepsilon > 0$  and  $\gamma > 0$  consider the following  $\varepsilon$ -dependent inner neighbourhoods of the wells:

$$\begin{aligned}
 \Omega^i &= (s_{i-1}, s_i), \\
 \Omega_\varepsilon^i &= [s_{i-1} + 2\varepsilon^\gamma, s_i - 2\varepsilon^\gamma],
 \end{aligned} \tag{3.1}$$

where by convention  $\Omega^1 = (-\infty, s_1)$ ,  $\Omega_\varepsilon^1 = (-\infty, s_1 - 2\varepsilon^\gamma)$ ,  $\Omega^n = (s_{n-1}, +\infty)$ , and  $\Omega_\varepsilon^n = [s_{n-1} + 2\varepsilon^\gamma, +\infty)$ .

Consider the following life times of the process  $X^\varepsilon$  in the potential wells:

$$\sigma^i(\varepsilon) = \inf\{t \geq 0 : X_t^\varepsilon(\cdot) \notin [s_{i-1} + \varepsilon^\gamma, s_i - \varepsilon^\gamma]\}, \quad i = 1, \dots, n. \tag{3.2}$$

Let

$$\lambda^i(\varepsilon) := H_- \left( \frac{s_{i-1} - m_i}{\varepsilon} \right) + H_+ \left( \frac{s_i - m_i}{\varepsilon} \right), \quad i = 1, \dots, n. \tag{3.3}$$

**Proposition 3.1.** *There exists  $\gamma_0 > 0$  such that for any  $0 < \gamma \leq \gamma_0$ ,  $x \in \Omega_\varepsilon^i$ ,  $i = 1, \dots, n$ , any  $\theta > -1$  the following limit holds:*

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}_x e^{-\theta \lambda(\varepsilon) \sigma(\varepsilon)} = \frac{1}{\theta + 1}, \tag{3.4}$$

and hence  $\lambda(\varepsilon) \sigma(\varepsilon) \xrightarrow{d} \exp(1)$ , and

$$\lim_{\varepsilon \downarrow 0} \mathbf{E}_x [\lambda^i(\varepsilon) \sigma^i(\varepsilon)]^p = \int_0^\infty y^p e^{-y} dy, \quad p \geq 0. \tag{3.5}$$

Moreover, for  $j \neq i$ ,

$$\lim_{\varepsilon \downarrow 0} \mathbf{P}_x (X_{\sigma^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j) = \frac{q_{ij}}{q_i}. \tag{3.6}$$

All limits hold uniformly over  $x \in \Omega_\varepsilon^i$ .

Proposition 3.1 will easily follow from Lemmas 3.5 and 3.4 formulated below. The proof is rather technical and consists in applying the strong Markov property and accurate estimations of certain probabilities.

### 3.1. Useful technicalities

#### 3.1.1. Dynamics between big jumps

Due to the strong Markov property, for any stopping time  $\tau$  the process  $\xi_{t+\tau}^\varepsilon - \xi_\tau^\varepsilon$ ,  $t \geq 0$ , is also a Lévy process with the same law as  $\xi^\varepsilon$ . For  $k \geq 1$  consider processes

$$\begin{aligned} \xi_t^{k,\varepsilon} &= \xi_{t+\tau_{k-1}}^\varepsilon - \xi_{\tau_{k-1}}^\varepsilon, \\ x_t^{k,\varepsilon}(x) &= x - \int_0^t U'(x_{s-}^{k,\varepsilon}) ds + \varepsilon \xi_t^{k,\varepsilon}, \quad t \in [0, T_k^\varepsilon]. \end{aligned} \tag{3.7}$$

In our notation, for  $x \in \mathbb{R}$ ,

$$X_{t+\tau_{k-1}}^\varepsilon = x_t^{k,\varepsilon}(x_{\tau_{k-1}}^{k-1,\varepsilon} + \varepsilon W_{k-1}^\varepsilon) + \varepsilon W_k^\varepsilon \mathbb{I}\{t = T_k^\varepsilon\}, \quad t \in [0, T_k^\varepsilon], \quad k \geq 1. \tag{3.8}$$

Denote  $W_0^\varepsilon = T_0^\varepsilon = 0$ ,  $x^1(0) = x$ , and write  $\mathbb{I}\{A\}$  for the indicator function of a measurable set  $A$ .

#### 3.1.2. Constants $\rho$ , $\gamma$ , $p_0$ , and $c$

We assume that the threshold power  $\rho > 0$  and  $\gamma$  satisfy

$$\frac{1}{2} < \rho < 1, \quad 0 < \gamma \leq \gamma_0 := \frac{1-\rho}{5}. \tag{3.9}$$

Throughout this section we use a constant  $c$  such that the following holds for  $\varepsilon \in (0, \varepsilon_0]$  for some  $\varepsilon_0 > 0$ :

$$\begin{aligned} \sup_{y \in [s_{i-1} + \varepsilon^\gamma, s_i - \varepsilon^\gamma]} |X_t^0(y) - m_i| &\leq \frac{\varepsilon^{2\gamma}}{2} \text{ for } t \geq c|\ln \varepsilon|, \quad i = 1, \dots, n, \\ \sup_{|y-s_i| \geq \varepsilon^\gamma} |X_t^0(y) - s_i| &\geq \varepsilon^\gamma + 2\varepsilon^{2\gamma} \text{ for } t \geq c\varepsilon^\gamma, \quad i = 1, \dots, n-1. \end{aligned} \tag{3.10}$$

Let us show that these inequalities hold for some  $c > 0$ . Let  $T(x, y) = \inf\{t \geq 0 : X_t^0(x) = y\}$ . Then for any  $i = 1, \dots, n$ , and due to the properties of  $U$  we need to show that

$$\begin{aligned} T(s_{i-1} + \varepsilon^\gamma, m_i - \varepsilon^{2\gamma}/2), T(s_i - \varepsilon^\gamma, m_i + \varepsilon^{2\gamma}/2) &\leq c|\ln \varepsilon|, \quad i = 1, \dots, n, \\ T(s_i - \varepsilon^\gamma, s_i - \varepsilon^\gamma - 2\varepsilon^{2\gamma}), T(s_i + \varepsilon^\gamma, s_i + \varepsilon^\gamma + 2\varepsilon^{2\gamma}) &\leq c\varepsilon^\gamma, \quad i = 1, \dots, n-1, \end{aligned} \tag{3.11}$$

what easily follows from nondegeneracy properties of potential's extrema (Assumption **U2**).

3.1.3. *Technical lemmas*

For definiteness, we assume as in Section 2 that the well’s minimum is located at the origin, and denote well’s boundaries as  $-\infty \leq a < 0 < b < +\infty$ . Denote  $\lambda(\varepsilon) := H_- \left(\frac{a}{\varepsilon}\right) + H_+ \left(\frac{b}{\varepsilon}\right)$  and

$$\begin{aligned} I &:= [a, b], & I &:= (-\infty, b], \\ I_{\varepsilon,1} &:= [a + \varepsilon^\gamma, b - \varepsilon^\gamma], & \text{or} & & I_{\varepsilon,1} &:= (-\infty, b - \varepsilon^\gamma], \\ I_{\varepsilon,2} &:= [a + \varepsilon^\gamma + \varepsilon^{2\gamma}, b - \varepsilon^\gamma - \varepsilon^{2\gamma}] & & & I_{\varepsilon,2} &:= (-\infty, b - \varepsilon^\gamma - \varepsilon^{2\gamma}], \end{aligned} \tag{3.12}$$

if  $a > -\infty$  or  $a = -\infty$  respectively.

For  $y \in I_{\varepsilon,1}$ ,  $j \geq 1$ , we introduce the following events:

$$\begin{aligned} A_y^j &= A^j(y) = \{x_s^j(y) \in I_{\varepsilon,1}, s \in [0, T_j], x_{T_j}^j(y) + \varepsilon W_j \in I_{\varepsilon,1}\}, \\ A_y^{j,-} &= A^{j,-}(y) = \{x_s^j(y) \in I_{\varepsilon,1}, s \in [0, T_j], x_{T_j}^j(y) + \varepsilon W_j \in I_{\varepsilon,2}\}, \\ B_y^j &= B^j(y) = \{x_s^j(y) \in I_{\varepsilon,1}, s \in [0, T_j], x_{T_j}^j(y) + \varepsilon W_j \notin I_{\varepsilon,1}\}, \\ E_y^j &= \left\{ \omega : \sup_{t \in [0, T_j]} |x_t^j(y) - x_t^0(y)| \leq \frac{\varepsilon^{2\gamma}}{2}, \quad a > -\infty, \right. \\ E_y^j &= \left. \left\{ \omega : \sup_{t \in [0, \tau_y \wedge T_j]} x_t^j(y) \leq -M + 1 \text{ and } \sup_{t \in [\tau_y \wedge T_j, T_j]} |x_t^j(y) - x_{t-\tau_y \wedge T_j}^0(x_{\tau_y}^j(y))| \leq \frac{\varepsilon^{2\gamma}}{2} \right\}, \quad a = -\infty, \right. \end{aligned} \tag{3.13}$$

with  $M > 0$  and  $\tau_y$  defined in Section 2.3. Let also  $A_y = A_y^1$ ,  $A_y^- = A_y^{1,-}$ ,  $B_y = B_y^1$ ,  $E_y = E_y^1$ .

Let us recall Lemma 2.4 with  $d = 1/2$  and choose  $0 < q < (1 - \rho - \gamma) \wedge r\rho$  and  $k = k_\varepsilon := \lceil \varepsilon^{q/2} / \beta_\varepsilon \rceil$ , where  $\lceil x \rceil$  denotes the integer part of a real number  $x$ . Let

$$(E^j)^c := \{T_j \geq k_\varepsilon / \varepsilon^q\} \cup \bigcup_{m=0}^{k_\varepsilon-1} \left\{ \sup_{t \in [\frac{m}{\varepsilon^q}, \frac{m+1}{\varepsilon^q}]} |\varepsilon \xi_t^{j,\varepsilon} - \varepsilon \xi_{\frac{m}{\varepsilon^q}}^{j,\varepsilon}| \geq \varepsilon^{4\gamma} \right\}, \quad \text{and} \quad E := E^1. \tag{3.14}$$

Then Lemma 2.4 implies that  $(E_y^j)^c \subseteq (E^j)^c$ . Moreover, we obtain the following estimates.

**Lemma 3.2.** *For any  $\theta > -1$  there are  $p > 0$ ,  $C > 0$  and  $\varepsilon_0 > 0$  such that the following estimates hold for  $0 < \varepsilon \leq \varepsilon_0$ :*

$$\mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{E^c\}] \leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \exp(-\varepsilon^{-p}), \tag{3.15}$$

$$\mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{T_1 \leq c|\ln \varepsilon|\}] \leq \frac{C\beta_\varepsilon^2 |\ln \varepsilon|}{\theta\lambda(\varepsilon) + \beta_\varepsilon}, \tag{3.16}$$

$$\mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{T_1 \leq c\varepsilon^\gamma\}] \leq \frac{C\beta_\varepsilon^2 \varepsilon^\gamma}{\theta\lambda(\varepsilon) + \beta_\varepsilon}. \tag{3.17}$$

*Proof.* Let  $\theta > -1$  and  $\varepsilon_0$  be small enough such that  $\theta\lambda(\varepsilon) + \beta_\varepsilon > 0$  and Lemma 2.1 holds for  $0 < \varepsilon \leq \varepsilon_0$  with some  $p' > 0$ . Note also that  $k_\varepsilon$  increases to infinity slower than some power of  $1/\varepsilon$ . Then the Markov property and a straightforward calculation yield for some  $p > 0$ :

$$\begin{aligned} \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{E^c\}] &\leq \int_{k_\varepsilon/\varepsilon^q}^\infty e^{-(\theta\lambda(\varepsilon)+\beta_\varepsilon)t} \beta_\varepsilon dt + \sum_{m=0}^{k_\varepsilon-1} \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I} \left\{ \sup_{t \in [\frac{m}{\varepsilon^q}, \frac{m+1}{\varepsilon^q}]} |\varepsilon \xi_t^{j,\varepsilon} - \varepsilon \xi_{\frac{m}{\varepsilon^q}}^{j,\varepsilon}| \geq \varepsilon^{4\gamma} \right\} \right] \\ &\leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} (\exp(-(\theta\lambda(\varepsilon) + \beta_\varepsilon)k_\varepsilon \varepsilon^{-q}) + k_\varepsilon \exp(-\varepsilon^{-p'})) \leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \exp(-\varepsilon^{-p}). \end{aligned} \tag{3.18}$$

Inequalities (3.16) and (3.17) are obtained by a straightforward calculation taking into account that  $(\theta\lambda(\varepsilon) + \beta_\varepsilon)c|\ln \varepsilon| \leq c'\beta_\varepsilon|\ln \varepsilon|$  and  $(\theta\lambda(\varepsilon) + \beta_\varepsilon)c\varepsilon^\gamma \leq c'\beta_\varepsilon\varepsilon^\gamma$  in the limit of small  $\varepsilon$  for some  $c' > 0$ .  $\square$

We will also use an important result on the uniform convergence for slowly varying functions.

**Proposition 3.3** [1, Th. 1.2.1). *If  $l$  is slowly varying at  $+\infty$  then*

$$\lim_{u \rightarrow +\infty} \frac{l(\lambda u)}{l(u)} = 1, \tag{3.19}$$

uniformly for  $\lambda$  from a compact set in  $(0, +\infty)$ .

**3.2. Proof of Proposition 3.1. Lower estimate**

In this section we obtain an estimate for the Laplace transform of  $\lambda(\varepsilon)\sigma(\varepsilon)$  from below in the small noise limit  $\varepsilon \rightarrow 0$ .

**Lemma 3.4.** *For any  $\theta \geq -1$ , any  $C > 0$  there exists  $\varepsilon_0 > 0$  such that for all  $0 < \varepsilon \leq \varepsilon_0$*

$$\mathbf{E}_x e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} \geq \frac{1 - C}{1 + \theta + C} \tag{3.20}$$

uniformly over  $x \in I_{\varepsilon,2}$ .

*Proof.* For  $x \in I_{\varepsilon,2}$ , we use the total probability formula to obtain an estimate

$$\mathbf{E}_x e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} \geq \sum_{k=1}^{\infty} \mathbf{E}_x [e^{-\theta\lambda(\varepsilon)\tau_k} \mathbb{I}\{\sigma = \tau_k\}]. \tag{3.21}$$

Then applying the independence and law properties of the processes  $x^j$ ,  $j \in \mathbb{N}$ , the following chain of inequalities is deduced which results in a factorisation formula for the expectation under estimation:

$$\begin{aligned} \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\tau_k} \mathbb{I}\{\sigma = \tau_k\} \right] &= \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\tau_k} \mathbb{I}\{X_s^\varepsilon \in I_{\varepsilon,1}, s \in [0, \tau_k), X_{\tau_k}^\varepsilon \notin I_{\varepsilon,1}\} \right] \\ &= \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\tau_k} \prod_{j=1}^{k-1} \mathbb{I}\{A^j(X_{\tau_{j-1}}^\varepsilon)\} \cdot \mathbb{I}\{B^k(X_{\tau_{k-1}}^\varepsilon)\} \right] \\ &\geq \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\tau_k} \prod_{j=1}^{k-1} \mathbb{I}\{A^{j,-}(X_{\tau_{j-1}}^\varepsilon)\} \cdot \mathbb{I}\{B^k(X_{\tau_{k-1}}^\varepsilon)\} \right] \\ &= \mathbf{E} \left[ \prod_{j=1}^{k-1} e^{-\theta\lambda(\varepsilon)T_j} \inf_{y \in I_{\varepsilon,2}} \mathbb{I}\{A_y^{j,-}\} \cdot e^{-\theta\lambda T_k} \inf_{y \in I_{\varepsilon,2}} \mathbb{I}\{B_y^k\} \right] \\ &= \left( \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \inf_{y \in I_{\varepsilon,2}} \mathbb{I}\{A_y^-\} \right] \right)^{k-1} \cdot \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \inf_{y \in I_{\varepsilon,2}} \mathbb{I}\{B_y\} \right]. \end{aligned} \tag{3.22}$$

For  $y \in I_{\varepsilon,2}$ , we next specify separately in two steps the further estimation for the two different events appearing in the latter formula.

**Step B1-1.** Consider the event  $\mathbb{I}\{A_y^-\}$  for  $y \in I_{\varepsilon,2}$ . Then

$$\mathbb{I}\{A_y^-\} \geq \mathbb{I}\{A_y^-\} \mathbb{I}\{E_y\} \mathbb{I}\{|\varepsilon W_1| \leq \frac{\varepsilon^{2\gamma}}{2}\} \mathbb{I}\{T_1 \geq c\varepsilon^\gamma\} + \mathbb{I}\{A_y^-\} \mathbb{I}\{E_y\} \mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\} \mathbb{I}\{T_1 \geq c|\ln \varepsilon|\}. \tag{3.23}$$

It is easy to see that on the events  $E_y$  and  $\{T_1 \geq c|\ln \varepsilon|\}$  we have the inclusion  $A_y^- \supseteq \{\varepsilon W_1 \in [a + 3\varepsilon^\gamma, b - 3\varepsilon^\gamma]\}$ , *i.e.* since  $x^\varepsilon(y)$  jumps from a  $\varepsilon^\gamma$ -neighbourhood of zero, and the big jump  $\varepsilon W_1$  is bounded, the jump-diffusion does not leave the interval  $I_{\varepsilon,2}$ . Continuing the chain of inequalities we obtain with help of  $E_y^c \subseteq E^c$ :

$$\begin{aligned} \mathbb{I}\{A_y^-\} &\geq \mathbb{I}\{E_y\}\mathbb{I}\{|\varepsilon W_1| \leq \frac{\varepsilon^{2\gamma}}{2}\}\mathbb{I}\{T_1 \geq c\varepsilon^\gamma\} + \mathbb{I}\{E_y\}\mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\}\mathbb{I}\{T_1 \geq c|\ln \varepsilon|\}\mathbb{I}\{\varepsilon W_1 \in [a + 3\varepsilon^\gamma, b - 3\varepsilon^\gamma]\} \\ &\geq \mathbb{I}\{|\varepsilon W_1| \leq \frac{\varepsilon^{2\gamma}}{2}\}\mathbb{I}\{T_1 \geq c\varepsilon^\gamma\} + \mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\}\mathbb{I}\{T_1 \geq c|\ln \varepsilon|\}\mathbb{I}\{\varepsilon W_1 \in [a + 3\varepsilon^\gamma, b - 3\varepsilon^\gamma]\} - 2\mathbb{I}\{E_y^c\} \\ &\geq \mathbb{I}\{|\varepsilon W_1| \leq \frac{\varepsilon^{2\gamma}}{2}\} + \mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\}\mathbb{I}\{\varepsilon W_1 \in [a + 3\varepsilon^\gamma, b - 3\varepsilon^\gamma]\} \\ &\quad - \mathbb{I}\{T_1 < c\varepsilon^\gamma\} - \mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\}\mathbb{I}\{T_1 < c|\ln \varepsilon|\} - 2\mathbb{I}\{E^c\} \\ &= \mathbb{I}\{\varepsilon W_1 \in [a + 3\varepsilon^\gamma, b - 3\varepsilon^\gamma]\} - \mathbb{I}\{T_1 < c\varepsilon^\gamma\} - \mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\}\mathbb{I}\{T_1 < c|\ln \varepsilon|\} - 2\mathbb{I}\{E^c\}. \end{aligned} \tag{3.24}$$

**Step B1-2.** Similarly, the event  $\mathbb{I}\{B_y\}$ ,  $y \in I_{\varepsilon,2}$  may be estimated as follows:

$$\begin{aligned} \mathbb{I}\{B_y\} &\geq \mathbb{I}\{B_y\}\mathbb{I}\{E_y\}\mathbb{I}\{T_1 \geq c|\ln \varepsilon|\} \\ &\geq \mathbb{I}\{E_y\}\mathbb{I}\{T_1 \geq c|\ln \varepsilon|\}\mathbb{I}\{\varepsilon W_1 \notin [a - \varepsilon^\gamma - \varepsilon^{2\gamma}, b + \varepsilon^\gamma + \varepsilon^{2\gamma}]\} \\ &\geq \mathbb{I}\{\varepsilon W_1 \notin [a - 2\varepsilon^\gamma, b + 2\varepsilon^\gamma]\}(1 - \mathbb{I}\{T_1 < c|\ln \varepsilon|\} - \mathbb{I}\{E^c\}), \end{aligned} \tag{3.25}$$

where we used that if  $\varepsilon W_1 \notin [a - \varepsilon^\gamma - \varepsilon^{2\gamma}, b + \varepsilon^\gamma + \varepsilon^{2\gamma}]$  and the process  $x^\varepsilon(y)$  jumps from a  $\varepsilon^\gamma$ -neighbourhood of the origin, then it leaves the interval  $I_{\varepsilon,1}$  for sure.

Now we apply (3.24) and (3.25) to estimate the expectations appearing in the formula for  $\mathbf{E}_x[e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)}\mathbb{I}\{\sigma(\varepsilon) = \tau_k\}]$ . Let  $C > 0$  and let  $\varepsilon$  be small enough such that Lemma 3.2 holds and  $\lambda(\varepsilon)\beta_\varepsilon^{-1}(1 + C) < 1$ .

**Step B2-1.** Here we estimate  $\mathbf{E}[\inf_{y \in I_{\varepsilon,2}} \mathbb{I}\{A_y^-\}]$ . With help of (3.24), Lemma 3.2 and using the independence of  $T_1$  and  $W_1$  we obtain:

$$\begin{aligned} \mathbf{E} \left[ \inf_{y \in I_{\varepsilon,2}} \mathbb{I}\{A_y^-\} \right] &\geq \mathbf{E}e^{-\theta\lambda(\varepsilon)T_1} \mathbf{P}(\varepsilon W_1 \in [a + 3\varepsilon^\gamma, b - 3\varepsilon^\gamma]) - \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{T_1 < c\varepsilon^\gamma\}] \\ &\quad - \mathbf{P}(|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}) \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{T_1 < c|\ln \varepsilon|\}] - 2\mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{E^c\}] \\ &\geq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \left( 1 - \frac{H_-((a + 3\varepsilon^\gamma)/\varepsilon) + H_+((b - 3\varepsilon^\gamma)/\varepsilon)}{\beta_\varepsilon} - c_1\beta_\varepsilon\varepsilon^\gamma \right. \\ &\quad \left. - c_2H(1/(2\varepsilon^{1-2\gamma}))|\ln \varepsilon| - 2e^{-1/\varepsilon^p} \right) \\ &\geq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \left( 1 - \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 + C) \right). \end{aligned} \tag{3.26}$$

To obtain the latter inequality we used the uniform convergence from Proposition 3.3.

**Step B2-2.** Similarly, for  $\varepsilon$  small enough we obtain

$$\begin{aligned} \mathbf{E} \left[ \inf_{y \in I_{\varepsilon,2}} \mathbb{I}\{B_y\} \right] &\geq \mathbf{P}(\varepsilon W_1 \notin [a - 2\varepsilon^\gamma, b + 2\varepsilon^\gamma]) \left( \mathbf{E}e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} - \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{T_1 < c|\ln \varepsilon|\}] - \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{E^c\}] \right) \\ &\geq \frac{H_-(a - 2\varepsilon^\gamma) + H_+(b + \varepsilon^\gamma)}{\beta_\varepsilon} \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \left( 1 - c_1\beta_\varepsilon|\ln \varepsilon| - c_2e^{-1/\varepsilon^p} \right) \geq \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 - C). \end{aligned} \tag{3.27}$$



Consequently for  $\varepsilon$  small enough and  $x \in I_{\varepsilon,2}$  we have

$$\mathbf{E}_x[e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)}] \geq \sum_{k=1}^{\infty} \left[ \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \right]^k \left[ 1 - \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 + C) \right]^{k-1} \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 - C) \geq \frac{1 - C}{\theta + 1 + C}, \tag{3.28}$$

where an easy calculation justifies that the series converges for all  $\theta \geq -1$ . □

### 3.3. Proof of Proposition 3.1. Upper estimate

In this section we obtain an estimate for the Laplace transform of  $\lambda(\varepsilon)\sigma(\varepsilon)$  from above in the small noise limit  $\varepsilon \rightarrow 0$ . This leads to the following Lemma with a rather technical proof again.

**Lemma 3.5.** *For any  $\theta > -1$ , any  $C \in (0, \theta + 1)$  there exists  $\varepsilon_0 > 0$  such that for all  $0 < \varepsilon \leq \varepsilon_0$*

$$\mathbf{E}_x e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} \leq \frac{1 + C}{1 + \theta - C} \tag{3.29}$$

uniformly over  $x \in I_{\varepsilon,2}$ .

*Proof.* For  $x \in I_{\varepsilon,1}$ , we use the following obvious inequality

$$\mathbf{E}_x e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} = \sum_{k=1}^{\infty} \left( \mathbf{E}_x [e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} \mathbb{I}\{\sigma(\varepsilon) = \tau_k\}] + R_k(x, \varepsilon) \right), \tag{3.30}$$

where

$$R_k(x, \varepsilon) \leq \begin{cases} \mathbf{E}_x [e^{-\theta\lambda(\varepsilon)\tau_k} \mathbb{I}\{\sigma(\varepsilon) \in (\tau_{k-1}, \tau_k)\}], & -1 < \theta < 0, \\ \mathbf{E}_x [e^{-\theta\lambda(\varepsilon)\tau_{k-1}} \mathbb{I}\{\sigma(\varepsilon) \in (\tau_{k-1}, \tau_k)\}], & \theta \geq 0. \end{cases} \tag{3.31}$$

With arguments analogous to (3.22) we obtain the factorisation:

$$\begin{aligned} \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} \mathbb{I}\{\sigma(\varepsilon) = \tau_k\} \right] &= \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} \mathbb{I}\{X_s^\varepsilon \in I_{\varepsilon,1}, s \in [0, \tau_k), X_{\tau_k}^\varepsilon \notin I_{\varepsilon,1}\} \right] \\ &= \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\tau_k} \prod_{j=1}^{k-1} \mathbb{I}\{A^j(X_{\tau_{j-1}}^\varepsilon)\} \cdot \mathbb{I}\{B^k(X_{\tau_{k-1}}^\varepsilon)\} \right] \\ &\leq \left( \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \right] \right)^{k-1} \cdot \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{B_y\} \right]. \end{aligned} \tag{3.32}$$

Analogously we estimate the remainder term  $R_k$  which refers to the exit between the  $(k - 1)$ th and the  $k$ th arrival times of the compound Poisson process  $\eta^\varepsilon$ ,  $k \in \mathbb{N}$ . Here we distinguish two cases.

In the first case,  $k = 1$ ,  $x \in I_{\varepsilon,2}$ . Then

$$\begin{aligned} R_1(x, \varepsilon) &\leq \begin{cases} \mathbf{E}_x [e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{\exists s \in (0, T_1) : X_s^\varepsilon \notin I_{\varepsilon,1}\}], & -1 < \theta < 0, \\ \mathbf{E}_x [\mathbb{I}\{\exists s \in (0, T_1) : X_s^\varepsilon \notin I_{\varepsilon,1}\}], & \theta \geq 0, \end{cases} \\ &\leq \begin{cases} \mathbf{E}_x [e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,2}} \mathbb{I}\{\exists s \in (0, T_1) : X_s^\varepsilon \notin I_{\varepsilon,2}\}], & -1 < \theta < 0, \\ \mathbf{E}_x [\sup_{y \in I_{\varepsilon,2}} \mathbb{I}\{\exists s \in (0, T_1) : X_s^\varepsilon \notin I_{\varepsilon,2}\}], & \theta \geq 0. \end{cases} \end{aligned} \tag{3.33}$$

In the second case,  $k \geq 2$ ,  $x \in I_{\varepsilon,1}$ . Consider first non-negative values of  $\theta$ . Then

$$\begin{aligned}
 R_k(x, \varepsilon) &\leq \mathbf{E}_x [e^{-\theta\lambda(\varepsilon)\tau_{k-1}} \mathbb{I}\{X_s^\varepsilon \in I_{\varepsilon,1}, s \in [0, \tau_{k-1}]\} \mathbb{I}\{\exists s \in (\tau_{k-1}, \tau_k) : X_s^\varepsilon \notin I_{\varepsilon,1}\}] \\
 &= \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\tau_{k-1}} \prod_{j=1}^{k-1} \mathbb{I}\{A^j(X_{\tau_{j-1}}^\varepsilon)\} \cdot \mathbb{I}\{\exists s \in [0, T_k] : x_s^k(X_{\tau_{k-1}}^\varepsilon) \notin I_{\varepsilon,1}\} \right] \\
 &\leq \mathbf{E} \left[ \prod_{j=1}^{k-2} e^{-\theta\lambda(\varepsilon)T_j} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y^j\} \cdot e^{-\theta\lambda(\varepsilon)T_{k-1}} \right. \\
 &\quad \left. \times \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y^{k-1}\} \mathbb{I}\{\exists s \in [0, T_k] : x_s^k(x_{T_{k-1}}^{k-1}(y) + \varepsilon W_{k-1}) \notin I_{\varepsilon,1}\} \right] \\
 &= \left( \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \right] \right)^{k-2} \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(x_{T_1}^1(y) + \varepsilon W_1) \notin I_{\varepsilon,1}\} \right]. \tag{3.34}
 \end{aligned}$$

Analogously, for  $-1 < \theta < 0$  we get

$$\begin{aligned}
 R_k(x, \varepsilon) &\leq \left( \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \right] \right)^{k-2} \\
 &\quad \times \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)(T_1+T_2)} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(x_{T_1}^1(y) + \varepsilon W_1) \notin I_{\varepsilon,1}\} \right]. \tag{3.35}
 \end{aligned}$$

Next, we specify separately in four steps the further estimation for the four different events appearing in the formulae for expectations above. The argument is similar to Steps **B1-1** and **B1-2** and we omit intermediate calculations.

**Step A1-1.** Consider  $\mathbb{I}\{A_y\}$  with  $y \in I_{\varepsilon,1}$ . Similarly to Step **B1-1**, we may estimate this indicator as

$$\mathbb{I}\{A_y\} \leq \mathbb{I}\{\varepsilon W_1 \in I\} + \mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\} \mathbb{I}\{T_1 < c|\ln \varepsilon|\} + \mathbb{I}\{E^c\}. \tag{3.36}$$

**Step A1-2.** Consider  $\mathbb{I}\{B_y\}$  with  $y \in I_{\varepsilon,1}$ . As in Step **B1-2** we get

$$\mathbb{I}\{B_y\} \leq \mathbb{I}\{\varepsilon W_1 \notin [a + 2\varepsilon^\gamma, b - 2\varepsilon^\gamma]\} + \mathbb{I}\{|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}\} \mathbb{I}\{T_1 < c|\ln \varepsilon|\} + \mathbb{I}\{T_1 < c\varepsilon^\gamma\} + \mathbb{I}\{E^c\}. \tag{3.37}$$

**Step A1-3.** Consider  $\mathbb{I}\{\exists s \in [0, T_1] : x_s^1(y) \notin I_{\varepsilon,1}\}$ . For  $y \in I_{\varepsilon,2}$ , we may estimate

$$\mathbb{I}\{\exists s \in [0, T_1] : x_s^1(y) \notin I_{\varepsilon,1}\} \leq \mathbb{I}\{E_y^c\} + \mathbb{I}\{\exists s \in [0, T_1] : x_s^1(y) \notin I_{\varepsilon,1}\} \mathbb{I}\{E_y\} \leq \mathbb{I}\{E^c\}. \tag{3.38}$$

**Step A1-4.** Finally, for  $y \in I_{\varepsilon,1}$ , we may estimate

$$\begin{aligned}
& \mathbb{I}\{A_y\} \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(x_{T_1}^1(y) + \varepsilon W_1) \notin I_{\varepsilon,1}\} \\
&= \mathbb{I}\{x_s^1(y) \in I, s \in (0, T_1], x_{T_1}^1(y) + \varepsilon W_1 \in I_{\varepsilon,2}\} \cdot \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(x_{T_1}^1(y) + \varepsilon W_1) \in I_{\varepsilon,1}\} \\
&\quad + \mathbb{I}\{x_s^1(y) \in I, s \in (0, T_1], x_{T_1}^1(y) + \varepsilon W_1 \in I_{\varepsilon,1} \setminus I_{\varepsilon,2}\} \cdot \mathbb{I}\{s \in [0, T_2] : x_s^2(x_{T_1}^1(y) + \varepsilon W_1) \notin I_{\varepsilon,1}\} \\
&\leq \mathbb{I}\{x_s^1(y) \in I, s \in (0, T_1], x_{T_1}^1(y) + \varepsilon W_1 \in I_{\varepsilon,2}\} \cdot \sup_{y \in I_{\varepsilon,2}} \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(y) \notin I_{\varepsilon,1}\} \\
&\quad + \mathbb{I}\{x_s^1(y) \in I, s \in (0, T_1], x_{T_1}^1(y) + \varepsilon W_1 \in I_{\varepsilon,1} \setminus I_{\varepsilon,2}\} \\
&\leq \sup_{y \in I_{\varepsilon,2}} \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(y) \notin I_{\varepsilon,1}\} + \mathbb{I}\{x_s^1(y) \in I_{\varepsilon,1}, s \in (0, T_1], x_{T_1}^1(y) + \varepsilon W_1 \in I_{\varepsilon,1} \setminus I_{\varepsilon,2}\}. \tag{3.39}
\end{aligned}$$

The first term in the resulting expression in the **Step A1-4** is identical to the expression handled in **Step A1-3**, while the second term requires an inessential modification of the estimation in **Step A1-2**, namely we consider an event  $\{x_{T_1}^1(y) + \varepsilon W_1 \in I_{\varepsilon,1} \setminus I_{\varepsilon,2}\}$  instead of  $\{x_{T_1}^1(y) + \varepsilon W_1 \notin I_{\varepsilon,1}\}$ .

Now we apply (3.36), (3.37), (3.38) and (3.39) to estimate the expectations  $\mathbf{E}_x[e^{-\theta\lambda(\varepsilon)\tau_k} \mathbb{I}\{\sigma(\varepsilon) = \tau_k\}]$  and remainder terms  $R_k(x, \varepsilon)$ . Fix  $\theta$  and let  $C$  be an arbitrary small constant such that  $0 < C < \theta + 1$ .

**Step A2-1.** For  $\varepsilon$  small enough we get

$$\begin{aligned}
\mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \right] &\leq \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1}] \mathbf{P}(\varepsilon W_1 \in I) + \mathbf{P}(|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}) \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{T_1 < c|\ln \varepsilon|\}] \\
&\quad + \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{E^c\}] \leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \left( 1 - \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 - C) \right). \tag{3.40}
\end{aligned}$$

**Step A2-2.** Similarly to Step **B1-2** we estimate for  $\varepsilon$  small enough:

$$\mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{B_y\} \right] \leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \frac{\lambda(\varepsilon)}{\beta_\varepsilon} \left( 1 + \frac{C}{3} \right). \tag{3.41}$$

On these steps, we again used the uniform convergence of slowly varying functions, see Proposition 3.3.

**Step A2-3.** Estimate  $R_1(x, \varepsilon)$ . For  $\theta > -1$  and  $x \in I_{\varepsilon,2}$  we have

$$R_1(x, \varepsilon) \leq e^{-1/\varepsilon^p} \leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \frac{\lambda(\varepsilon)}{\beta_\varepsilon} \frac{C}{3}, \tag{3.42}$$

$p > 0$  being a constant from Lemma 3.2.

**Step A2-4.** Finally we estimate the last factor in the representation for  $R_k(x, \varepsilon)$  for  $\theta \geq 0$  as

$$\begin{aligned}
& \mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)T_1} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(x_{T_1}^1(y) + \varepsilon W_1) \notin I_{\varepsilon,1}\} \right] \\
&\leq 2\mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{E^c\}] + \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1}] (\mathbf{P}(|\varepsilon W_1 - a| \leq \varepsilon^\gamma) + \mathbf{P}(|\varepsilon W_1 - b| \leq \varepsilon^\gamma)) \\
&\quad + \mathbf{E}[e^{-\theta\lambda(\varepsilon)T_1} \mathbb{I}\{T_1 < c|\ln \varepsilon|\}] \mathbf{P}(|\varepsilon W_1| > \frac{\varepsilon^{2\gamma}}{2}) \\
&\leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \frac{\lambda(\varepsilon)}{\beta_\varepsilon} \frac{C}{3} \tag{3.43}
\end{aligned}$$

and analogously for  $-1 < \theta < 0$  as

$$\mathbf{E} \left[ e^{-\theta\lambda(\varepsilon)(T_1+T_2)} \sup_{y \in I_{\varepsilon,1}} \mathbb{I}\{A_y\} \mathbb{I}\{\exists s \in [0, T_2] : x_s^2(x_{T_1}^1(y) + \varepsilon W_1) \notin I_{\varepsilon,1}\} \right] \leq \left( \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \right)^2 \frac{\lambda(\varepsilon) C}{\beta_\varepsilon} \frac{1}{4} \tag{3.44}$$

$$\leq \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \frac{\lambda(\varepsilon) C}{\beta_\varepsilon} \frac{1}{3}.$$

Collecting together the above estimates yields the upper bound for the Laplace transform of the first exit time  $\sigma_x(\varepsilon)$  for  $x \in I_{\varepsilon,2}$ . Fix  $\theta > -1$  and let  $\varepsilon$  be small enough such that all estimates above hold and  $\frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} (1 - \frac{\lambda}{\beta}(1 - C)) \leq \frac{1}{2}$ . Then the following chain of inequalities finishes the proof:

$$\begin{aligned} \mathbf{E}_x \left[ e^{-\theta\lambda(\varepsilon)\sigma(\varepsilon)} \right] &\leq \sum_{k=1}^{\infty} \left[ \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \left( 1 - \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 - C) \right) \right]^{k-1} \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \frac{\lambda(\varepsilon)}{\beta_\varepsilon} \left( 1 + \frac{C}{3} \right) \\ &\quad + \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \frac{\lambda(\varepsilon) C}{\beta_\varepsilon} \frac{1}{3} + \sum_{k=2}^{\infty} \left[ \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \left( 1 - \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 - C) \right) \right]^{k-1} \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \frac{\lambda(\varepsilon) C}{\beta_\varepsilon} \frac{1}{3} \\ &\leq \frac{\lambda(\varepsilon)(1 + C)}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \sum_{k=1}^{\infty} \left[ \frac{\beta_\varepsilon}{\theta\lambda(\varepsilon) + \beta_\varepsilon} \left( 1 - \frac{\lambda(\varepsilon)}{\beta_\varepsilon} (1 - C) \right) \right]^{k-1} = \frac{1 + C}{\theta + 1 - C}. \quad \square \end{aligned} \tag{3.45}$$

*Proof of the Proposition 3.1.* The first statement of Proposition 3.1 follows directly from Lemmas 3.5 and 3.4 and immediately implies convergence in distribution. Since the Laplace transform is finite for at least one negative  $\theta$ , the random variables  $\lambda(\varepsilon)\sigma_x(\varepsilon)$  have finite moments of all orders. Consequently, for any  $p > 0$ ,  $(\lambda(\varepsilon)\sigma_x(\varepsilon))^p \xrightarrow{d} \xi^p$ ,  $\xi$  being a standard exponential random variable, and  $(\lambda(\varepsilon)\sigma_x(\varepsilon))^p, 0 < \varepsilon \leq \varepsilon_0$ , is a family of uniformly intergable non-negative random variables, which implies the convergence of moments (see [14], Lem. 4.11).

To obtain the third statement we repeat the steps of the argument of Lemmas 3.5 and 3.4 taking  $\theta = 0$  and redefining the event  $B_y^j$  in (3.13) and thereafter as

$$\{x_s^j(y) \in I_{\varepsilon,1}, s \in [0, T_j], x_{T_j}^j(y) + \varepsilon W_j \in \Omega_\varepsilon^j\}. \tag{3.46}$$

Then, it is easy to see that for  $x \in \Omega_\varepsilon^i$

$$\begin{aligned} \left[ \frac{H_-\left(\frac{s_{j-1}-m_i}{\varepsilon}\right) - H_-\left(\frac{s_{j-1}-m_i}{\varepsilon}\right)}{H_-\left(\frac{s_{i-1}-m_i}{\varepsilon}\right) + H_+\left(\frac{s_j-m_i}{\varepsilon}\right)} \right]^{-1} \mathbf{P}_x(X_{\sigma_{\varepsilon^i}^\varepsilon} \in \Omega_\varepsilon^j) &\rightarrow 1, \quad \text{if } j < i, \\ \left[ \frac{H_+\left(\frac{s_{j-1}-m_i}{\varepsilon}\right) - H_+\left(\frac{s_j-m_i}{\varepsilon}\right)}{H_+\left(\frac{s_{i-1}-m_i}{\varepsilon}\right) + H_+\left(\frac{s_j-m_i}{\varepsilon}\right)} \right]^{-1} \mathbf{P}_x(X_{\sigma_{\varepsilon^i}^\varepsilon} \in \Omega_\varepsilon^j) &\rightarrow 1, \quad \text{if } i < j, \end{aligned} \tag{3.47}$$

and the ratios in brackets converge to  $q_{ij}/q_i$  as defined in (1.9). □

4. TRANSITIONS BETWEEN THE WELLS

For  $0 < \Delta < \Delta_0 = \min_{1 \leq i \leq n} \{|m_i - s_{i-1}|, |m_i - s_i|\}$  and  $x \in \mathbb{R}$  denote  $B_\Delta(x) = \{y : |x - y| \leq \Delta\}$ . Consider the following stopping times:

$$T^i(\varepsilon) = \inf\{t \geq 0 : X_t^\varepsilon(\cdot) \in \cup_{k \neq i} \Omega_\varepsilon^k\}, \tag{4.1}$$

$$\tau^i(\varepsilon) = \inf\{t \geq 0 : X_t^\varepsilon(\cdot) \in \cup_{k \neq i} B_\Delta(m_k)\}, \tag{4.2}$$

$$S^i(\varepsilon) = \inf\{t \geq 0 : X_t^\varepsilon(\cdot) \notin B_{2\varepsilon^\gamma}(s_i)\}, \quad i = 1, \dots, n - 1. \tag{4.3}$$

For  $x \in \Omega_\varepsilon^i$ ,  $T^i$  is the transition time between the wells. For  $x \in B_\Delta(m_i)$ ,  $\tau^i$  is the transition time between  $\Delta$ -neighbourhoods of wells' minima, and for  $x \in B_{2\varepsilon^\gamma}(s_i)$ ,  $S_x^i$  is the exit time from a neighbourhood of the saddle point.

**Lemma 4.1.** *Let  $i = 1, \dots, n - 1$  and  $x \in B_{2\varepsilon^\gamma}(s_i)$ . Then*

$$\lim_{\varepsilon \downarrow 0} H(1/\varepsilon) \mathbf{E}_x S^i(\varepsilon) = 0. \tag{4.4}$$

*Proof.* To estimate  $\mathbf{E}_x S^i(\varepsilon)$  we notice that for  $x \in B_{2\varepsilon^\gamma}(s_i)$ ,

$$S_x^i(\varepsilon) \leq \inf\{t > 0 : |\varepsilon L_t - \varepsilon L_{t-}| > 4\varepsilon^\gamma\} = J(\varepsilon) \text{ a.s.}, \tag{4.5}$$

*i.e.* the first exit time of  $X^\varepsilon$  from the  $2\varepsilon^\gamma$ -neighbourhood of the saddle point  $s_i$  is a.s. bounded from above by the time of the first jump of  $\varepsilon L$  exceeding  $4\varepsilon^\gamma$ . Note that  $J(\varepsilon)$  is exponentially distributed with mean

$$\mathbf{E}J(\varepsilon) = \left( \int_{|y| > 4/\varepsilon^{1-\gamma}} \nu(dy) \right)^{-1} = \frac{1}{H(4/\varepsilon^{1-\gamma})}. \tag{4.6}$$

The statement of the lemma follows from the fact that  $H(1/\varepsilon)/H(4/\varepsilon^{1-\gamma}) \rightarrow 0$  as  $\varepsilon \downarrow 0$ . □

**Proposition 4.2.** *For  $x \in \Omega_\varepsilon^i$  and  $j \neq i$*

$$\lim_{\varepsilon \downarrow 0} \mathbf{P}_x(X_{T^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j) = \frac{q_{ij}}{q_i} \tag{4.7}$$

$$\lim_{\varepsilon \downarrow 0} \mathbf{P}_x(T^i(\varepsilon) > \sigma^i(\varepsilon)) = 0, \tag{4.8}$$

$$\lim_{\varepsilon \downarrow 0} \lambda^i(\varepsilon) \mathbf{E}_x T^i(\varepsilon) = 1. \tag{4.9}$$

*Proof.* It is obvious that for all  $x \in \Omega_\varepsilon^i$

$$\sigma^i(\varepsilon) \leq T^i(\varepsilon) \quad \mathbf{P}_x\text{-a.s.} \tag{4.10}$$

We have the inequality

$$\mathbf{P}_x(X_{T^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j) = \mathbf{P}_x(X_{\sigma^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j) + \mathbf{P}_x(X_{T^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j, T^i(\varepsilon) > \sigma^i(\varepsilon)) \geq \mathbf{P}_x(X_{\sigma^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j). \tag{4.11}$$

Recall (3.6) in Proposition 3.1 and note that  $\sum_{j \neq i} \frac{q_{ij}}{q_i} = 1$ . Then the limits (4.7) and (4.8) follow.

For any  $\delta > 0$  there exists  $\varepsilon_0 > 0$  such that for  $0 < \varepsilon \leq \varepsilon_0$  the following estimates hold

$$\begin{aligned} \sup_{x \in \Omega_\varepsilon^i} \mathbf{P}_x \left( X_{\sigma^i(\varepsilon)}^\varepsilon \in \cup_{j=1}^{n-1} B_{2\varepsilon^\gamma}(s_j) \right) &\leq \delta, \\ \sup_{x \in \Omega_\varepsilon^i} \lambda^i(\varepsilon) \mathbf{E}_x \sigma^i(\varepsilon) &\leq 1 + \delta, \\ \max_{1 \leq j \leq n-1} \sup_{x \in B_{2\varepsilon^\gamma}(s_j)} \lambda^i(\varepsilon) \mathbf{E}_x S^j(\varepsilon) &\leq \delta. \end{aligned} \tag{4.12}$$

Then is easy to see that

$$\lambda^i(\varepsilon) \mathbf{E}_x T^i(\varepsilon) \leq \lambda^i(\varepsilon) \mathbf{E}_x \sigma^i(\varepsilon) + \sum_{k=1}^{\infty} (k+1)(1+\delta) + k\delta \delta^k \leq 1 + \delta \cdot \text{Const.} \tag{4.13}$$

which proves (4.9). □

**Proposition 4.3.** *For any  $0 < \Delta < \Delta_0$  the following limits hold*

$$\lim_{\varepsilon \downarrow 0} \mathbf{P}_x \left( X_{\tau^i(\varepsilon)}^\varepsilon \in B_\Delta(m_j) \right) = \frac{q_{ij}}{q_i} \tag{4.14}$$

$$\lambda^i(\varepsilon) \tau^i(\varepsilon) \xrightarrow{d} \exp(1), \tag{4.15}$$

$$\lim_{\varepsilon \downarrow 0} \lambda^i(\varepsilon) \mathbf{E}_x \tau^i(\varepsilon) = 1, \tag{4.16}$$

uniformly for  $x \in B_\Delta(m_i)$  and  $i = 1, \dots, n, j \neq i$ .

*Proof.* It is obvious that for all  $x \in B_\Delta(m_i)$

$$\sigma^i(\varepsilon) \leq T^i(\varepsilon) \leq \tau^i(\varepsilon) \quad \mathbf{P}_x\text{-a.s.} \tag{4.17}$$

On the other hand, the main contribution to  $\tau(\varepsilon)$  is made by the switching time  $T(\varepsilon)$ , for if the trajectory overcomes the saddle point and is in  $\Omega_\varepsilon^j$  for some  $j \neq i$ , it follows the deterministic trajectory with high probability and reaches the set  $B_\Delta(m_j)$  in short (logarithmic) time.

First we show that

$$\lim_{\varepsilon \downarrow 0} \mathbf{P}_x \left( \tau^i(\varepsilon) \leq T^i(\varepsilon) + c|\ln \varepsilon| \right) = 1, \tag{4.18}$$

where  $c$  is defined in (3.10). Let  $X_{T^i(\varepsilon)}^\varepsilon(x) \in \Omega_\varepsilon^j$  for some  $j \neq i$ . On the event  $A_\varepsilon = \{\omega : \sup_{t \in [0, c|\ln \varepsilon|]} |\varepsilon L_{t+T^i(\varepsilon)} - \varepsilon L_{T^i(\varepsilon)}| \leq \varepsilon^{4\gamma}\}$  the trajectory  $X_t^\varepsilon(X_{T^i(\varepsilon)}^\varepsilon(x))$  follows the deterministic trajectory  $x_t^0(X_{T^i(\varepsilon)}^\varepsilon(x))$  which reaches the small neighbourhood of the local minimum  $m_j$  in time  $c|\ln \varepsilon|$ . The limit (4.18) holds since  $\mathbf{P}_x(A_\varepsilon) \rightarrow 1$ . Then

$$\begin{aligned} \mathbf{P}_x(X_{\tau^i(\varepsilon)}^\varepsilon \in B_\Delta(m_j)) &\geq \mathbf{P}_x(X_{\tau^i(\varepsilon)}^\varepsilon \in B_\Delta(m_j), X_{T^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j, A_\varepsilon) \\ &= \mathbf{P}_x(X_{T^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j, A_\varepsilon) \geq \mathbf{P}_x(X_{T^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^j) - \mathbf{P}_x(A_\varepsilon^c) \rightarrow \frac{q_{ij}}{q_i} \end{aligned} \tag{4.19}$$

and (4.14) is proved since  $\sum_{j \neq i} \frac{q_{ij}}{q_i} = 1$ .

Convergence (4.15) follows easily from inequality (4.17), limits (4.8) and (4.18) and the fact that  $\lambda^i(\varepsilon)|\ln \varepsilon| \rightarrow 0$ .

To prove (4.16) we repeat the argument of Proposition 4.2. Indeed, for any  $\delta > 0$  there is  $\varepsilon_0 > 0$  such that for  $0 < \varepsilon \leq \varepsilon_0$  the following inequalities hold:

$$\begin{aligned} \sup_{x \in \Omega_\varepsilon^i} \mathbf{P}_x \left( \sup_{t \in [0, c|\ln \varepsilon|]} |\varepsilon L_t| \leq \varepsilon^{4\gamma} \right) &\leq \delta, & \max_{1 \leq i \leq n} \lambda^i(\varepsilon) c |\ln \varepsilon| &\leq \delta, \\ \sup_{x \in \Omega_\varepsilon^i} \lambda^i(\varepsilon) \mathbf{E}_x T^i(\varepsilon) &\leq 1 + \delta, & \max_{1 \leq j \leq n-1} \sup_{x \in B_{2\varepsilon^\gamma}(s_j)} \lambda^i(\varepsilon) \mathbf{E}_x S^j(\varepsilon) &\leq \delta. \end{aligned} \tag{4.20}$$

Then it is easy to see that for  $0 < \varepsilon \leq \varepsilon_0$

$$\lambda^i(\varepsilon) \mathbf{E}_x \tau^i(\varepsilon) \leq \lambda^i(\varepsilon) (\mathbf{E}_x T^i(\varepsilon) + c|\ln \varepsilon|) + \sum_{k=1}^{\infty} [(1 + \delta + \lambda^i(\varepsilon) c |\ln \varepsilon|)(k + 1) + \delta k] \delta^k \leq 1 + \delta \cdot \text{Const.} \tag{4.21}$$

which finishes the proof. □

### 5. METASTABLE BEHAVIOUR. PROOF OF THEOREM 1.1

#### 5.1. Convergence on short time intervals

**Proposition 5.1.** *Let  $0 < \delta < r$ . Then if  $x \in \Omega^i$ ,  $i = 1, \dots, n$ , then for  $t > 0$*

$$X_{t/\varepsilon^\delta}^\varepsilon(x) \xrightarrow{\mathcal{D}} m_i, \quad \varepsilon \downarrow 0. \tag{5.1}$$

*Proof.* For some  $1 \leq i \leq n$ , let  $x \in \Omega^i$ . We shall prove a stronger result: for any  $A > 0$  and  $0 < \Delta < \Delta_0$

$$\mathbf{P}_x \left( \sup_{s \in [c\varepsilon^\delta |\ln \varepsilon|, A]} |X_{s/\varepsilon^\delta}^\varepsilon - m_i| \leq \Delta \right) = \mathbf{P}_x \left( \sup_{s \in [c|\ln \varepsilon|, A/\varepsilon^\delta]} |X_s^\varepsilon - m_i| \leq \Delta \right) \rightarrow 1, \quad \varepsilon \downarrow 0. \tag{5.2}$$

Indeed, recalling Section 2 we choose  $\gamma > 0$  and  $c > 0$  such that  $|X_{c|\ln \varepsilon|}^\varepsilon(x) - m_i| \leq \Delta/2$  a.s. on the event  $E = \mathcal{E}_{c|\ln \varepsilon|} \cap \{T_1 > c|\ln \varepsilon|\}$ , where  $\mathcal{E}_{c|\ln \varepsilon|} = \{\sup_{[0, c|\ln \varepsilon|]} |\varepsilon \xi_t^\varepsilon| \leq \varepsilon^{4\gamma}\}$ . This gives

$$\begin{aligned} \mathbf{P}_x \left( \sup_{s \in [c|\ln \varepsilon|, A/\varepsilon^\delta]} |X_s^\varepsilon - m_i| > \Delta \right) &\leq \sup_{|y - m_i| \leq \Delta/2} \mathbf{P}_y \left( \sup_{s \in [0, A/\varepsilon^\delta - c|\ln \varepsilon|]} |X_s^\varepsilon - m_i| > \Delta \right) + \mathbf{P}(E^c) \\ &\leq \sup_{|y - m_i| \leq \Delta/2} \mathbf{P}_y (\sigma_\Delta(\varepsilon) < A/\varepsilon^\delta - c|\ln \varepsilon|) + \mathbf{P}(E^c) \\ &\leq \sup_{|y - m_i| \leq \Delta/2} \mathbf{P}_y (\sigma_\Delta(\varepsilon) < A/\varepsilon^\delta) + \mathbf{P}(\mathcal{E}_{\mu|\ln \varepsilon|}^c) + \mathbf{P}(T_1 \leq c|\ln \varepsilon|), \end{aligned} \tag{5.3}$$

where  $\sigma_\Delta(\varepsilon) = \inf\{t > 0 : |X_t^\varepsilon - m_i| > \Delta\}$ . On the other hand we know that for  $\lambda_\Delta(\varepsilon) = H_-(-\Delta/\varepsilon) + H_+(\Delta/\varepsilon)$ ,

$$\lambda_\Delta(\varepsilon) \sigma_\Delta(\varepsilon) \xrightarrow{d} \exp(1). \tag{5.4}$$

Since  $\lambda_\Delta(\varepsilon)/\varepsilon^\delta \rightarrow 0$  as  $\varepsilon \downarrow 0$  we have  $\mathbf{P}_y (\sigma_\Delta(\varepsilon) < A/\varepsilon^\delta) \rightarrow 0$ , as well as  $\mathbf{P}(\mathcal{E}_{c|\ln \varepsilon|}^c) \rightarrow 0$  and  $\mathbf{P}(T_1 \leq c|\ln \varepsilon|) \rightarrow 0$  in the limit of small  $\varepsilon$ . This finishes the proof of (5.2). □

**Remark 5.2.** It is easy to notice in view of Section 2 that the convergence in Proposition 5.1 is uniform in  $x$  for  $x \in \Omega_\varepsilon^i$ .

5.2. Proof of Theorem 1.1

**Lemma 5.3.** For any  $t > 0$  and  $0 < \Delta < \Delta_0$ ,

$$\mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \in \cup_{j=1}^n B_\Delta(m_j) \right) \rightarrow 1, \quad \varepsilon \downarrow 0, \tag{5.5}$$

uniformly for  $x \in \mathbb{R}$ .

*Proof.* Choose  $\rho$  and  $\gamma$  such that Proposition 3.1 and Lemma 4.1 hold for small  $\varepsilon$ . 1. Let  $|x - s_i| \leq 2\varepsilon^\gamma$  for some  $i = 1, \dots, n - 1$ . We know (see Lem. 4.1) that  $S_x^i(\varepsilon) \leq J(\varepsilon) = \inf\{t > 0 : |\varepsilon L_t - \varepsilon L_{t-}| > 4\varepsilon^\gamma\}$  a.s. and  $J(\varepsilon) \sim \exp(\frac{1}{H(4/\varepsilon^{1-\gamma})})$  if  $1 - \gamma > \rho$ . We show that

$$\mathbf{P}_x \left( X_{2/\varepsilon^{r(1-\gamma/2)}}^\varepsilon \notin \cup_{j=1}^n B_\Delta(m_j) \right) \rightarrow 0. \tag{5.6}$$

Indeed, the strong Markov property implies

$$\begin{aligned} \mathbf{P}_x \left( X_{2/\varepsilon^{r(1-\gamma/2)}}^\varepsilon \notin \cup_{j=1}^n B_\Delta(m_j) \right) &\leq \mathbf{P}_x \left( X_{2/\varepsilon^{r(1-\gamma/2)}}^\varepsilon \notin \cup_{j=1}^n B_\Delta(m_j), S^i(\varepsilon) \leq 1/\varepsilon^{r(1-\gamma/2)} \right) \\ &\quad + \mathbf{P} \left( J(\varepsilon) > 1/\varepsilon^{r(1-\gamma/2)} \right) \\ &= \sum_{k=1}^n \mathbf{E}_x \left[ \mathbf{P}_{X_{S^i(\varepsilon)}^\varepsilon} \left( X_{2/\varepsilon^{r(1-\gamma/2)} - S^i(\varepsilon)}^\varepsilon \notin \cup_{j=1}^n B_\Delta(m_j) \right) \cdot \mathbb{I}\{S^i(\varepsilon) \leq 1/\varepsilon^{r(1-\gamma/2)}\} \right] \\ &\leq 1/\varepsilon^{r(1-\gamma/2)} \cdot \mathbb{I}\{X_{S^i(\varepsilon)}^\varepsilon \in \Omega_\varepsilon^k\} \\ &\quad + \mathbf{E}_x \left[ \mathbf{P}_{X_{S^i(\varepsilon)}^\varepsilon} \left( X_{2/\varepsilon^{r(1-\gamma/2)} - S^i(\varepsilon)}^\varepsilon \notin \cup_{j=1}^n B_\Delta(m_j) \right) \cdot \mathbb{I}\{S^i(\varepsilon) > 1/\varepsilon^{r(1-\gamma/2)}\} \right] \\ &\leq 1/\varepsilon^{r(1-\gamma/2)} \cdot \mathbb{I}\{X_{S^i(\varepsilon)}^\varepsilon \notin \cup_{j=1}^n \Omega_\varepsilon^j\} \\ &\quad + \mathbf{P} \left( J(\varepsilon) > 1/\varepsilon^{r(1-\gamma/2)} \right) \\ &\leq \sum_{k=1}^n \sup_{y \in \Omega_\varepsilon^k} \mathbf{P}_y \left( \sup_{s \in [c|\ln \varepsilon|, 2/\varepsilon^{r(1-\gamma/2)}]} |X_s^\varepsilon - m_k| > \Delta \right) \\ &\quad + \mathbf{P} \left( \sup_{t \in [0, 1/\varepsilon^{r(1-\gamma/2)}]} \varepsilon |L_t - L_{t-}| > a \right) \\ &\quad + \mathbf{P} \left( J(\varepsilon) > 1/\varepsilon^{r(1-\gamma/2)} \right), \end{aligned} \tag{5.7}$$

with  $a = \frac{1}{2} \min\{s_2 - s_1, \dots, s_{n-1} - s_{n-2}\}$ . The first summand in the latter formula tends to 0 due to Proposition 5.1. The second summand is estimated by  $1 - \exp(\varepsilon^{-r(1-\gamma/2)} H(a/\varepsilon)) \rightarrow 0$ , and the third summand also tends to 0 due to the definition of  $J(\varepsilon)$ .

2. It is clear from the proof that the limit (5.6) holds also for  $x \in \Omega_\varepsilon^i$ ,  $i = 1, \dots, n$ , and thus for all  $x \in \mathbb{R}$ . Then, for  $\varepsilon$  small enough such that  $t/H(1/\varepsilon) > 2/\varepsilon^{r(1-\gamma/2)}$  the application of the Markov property

$$\mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \notin \cup_{j=1}^n B_\Delta(m_j) \right) = \mathbf{E}_x \mathbf{P}_{X_{t/H(1/\varepsilon) - 2/\varepsilon^{r(1-\gamma/2)}}^\varepsilon} \left( X_{2/\varepsilon^{r(1-\gamma/2)}}^\varepsilon \notin \cup_{j=1}^n B_\Delta(m_j) \right) \tag{5.8}$$

finishes the proof. □



*Proof of Theorem 1.1.* It is clear from the Markov property that it is sufficient to show that for any  $t > 0$  and  $x \in \Omega_\varepsilon^i, i = 1, \dots, n,$

$$\mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \in B_\Delta(m_j) \right) \rightarrow \mathbf{P}_{m_i} (Y_t = m_j), \quad j \neq i. \tag{5.9}$$

Define a sequence of stopping times  $(\tau(k))_{k \geq 0}$  and states  $(m(k))_{k \geq 0}$  such that  $\tau(0) = 0, m(0) = m_i$  and for  $k \geq 1$

$$\begin{aligned} \tau(k) &= \inf\{t > \tau(k-1) : X_t^\varepsilon \in \bigcup_{i=1}^n B_\Delta(m_i) \setminus B_\Delta(m(k-1))\}, \\ m(k) &= \sum_{j=1}^n m_j \mathbb{I}\{X_{\tau(k)}^\varepsilon \in B_\Delta(m_j)\}. \end{aligned} \tag{5.10}$$

Define also a (non-Markovian) process  $\hat{X}^\varepsilon$  on a state space  $\{m_1, \dots, m_n\}$

$$\hat{X}_t^\varepsilon = \sum_{k=0}^\infty m(k) \cdot \mathbb{I}\{t \in [H(1/\varepsilon)\tau(k), H(1/\varepsilon)\tau(k+1))\}. \tag{5.11}$$

The strong Markov property of  $X^\varepsilon$  and Proposition 4.3 imply that

$$\begin{aligned} \text{Law} \left( H(1/\varepsilon)(\tau(k+1) - \tau(k)) \mid \hat{X}_{\tau(k)}^\varepsilon = m_i \right) &\rightarrow \exp(1/q_i), \\ \mathbf{P}_x(\hat{X}_{\tau(k+1)}^\varepsilon = m_j \mid \hat{X}_{\tau(k)}^\varepsilon = m_i) &\rightarrow \frac{q_{ij}}{q_i}, \end{aligned} \tag{5.12}$$

uniformly for  $k \geq 0$ .

The process  $Y$  defined in the statement of the Theorem is given by the sequence of its jump times and states,  $(\theta(k), Y_k)_{k \geq 0}$  with the property that the interjump times are conditionally independent and exponentially distributed and for  $k \geq 0, 0 \leq i, j \leq n, i \neq j,$

$$\begin{aligned} \text{Law}(\theta(k+1) - \theta(k) \mid Y_k = m_i) &= \exp(1/q_i), \\ \mathbf{P}(Y_{k+1} = m_j \mid Y_k = m_i) &= \frac{q_{ij}}{q_i}. \end{aligned} \tag{5.13}$$

Then

$$\begin{aligned} &\left| \mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \in B_\Delta(m_j) \right) - \mathbf{P}_{m_i} (Y_t = m_j) \right| \\ &\leq \left| \mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \in B_\Delta(m_j) \right) - \mathbf{P}_x \left( \hat{X}_t^\varepsilon = m_j \right) \right| + \left| \mathbf{P}_x \left( \hat{X}_t^\varepsilon = m_j \right) - \mathbf{P}_{m_i} (Y_t = m_j) \right|. \end{aligned} \tag{5.14}$$

The second summand in (5.14) vanishes in the limit of small  $\varepsilon$  due to the weak convergence of the jump process  $\hat{X}^\varepsilon$  to  $Y$ . Indeed, in this case the weak convergence is equivalent to the weak convergence of the sequences of jump times and jump sizes (see [22])  $(\tau(k), m(k))_{k \geq 0} \Rightarrow (\theta(k), Y_k)_{k \geq 0},$  which follows from (5.12) and (5.13).

To estimate the first summand in (5.14) we use Lemma 5.3. Indeed,

$$\begin{aligned} &\left| \mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \in B_\Delta(m_j) \right) - \mathbf{P}_x \left( \hat{X}_t^\varepsilon = m_j \right) \right| = \\ &\quad \left| \mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \in B_\Delta(m_j), \hat{X}_t^\varepsilon = m_j \right) + \mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \in B_\Delta(m_j), \hat{X}_t^\varepsilon \neq m_j \right) \right| (= 0) \\ &\quad - \mathbf{P}_x \left( \hat{X}_t^\varepsilon = m_j, X_{t/H(1/\varepsilon)}^\varepsilon \in B_\Delta(m_j) \right) - \mathbf{P}_x \left( \hat{X}_t^\varepsilon = m_j, X_{t/H(1/\varepsilon)}^\varepsilon \in \cup_{k \neq j} B_\Delta(m_k) \right) (= 0) \\ &\quad - \mathbf{P}_x \left( \hat{X}_t^\varepsilon = m_j, X_{t/H(1/\varepsilon)}^\varepsilon \notin \cup_{k=1}^n B_\Delta(m_k) \right) \leq \mathbf{P}_x \left( X_{t/H(1/\varepsilon)}^\varepsilon \notin \cup_{k=1}^n B_\Delta(m_k) \right) \rightarrow 0, \end{aligned} \tag{5.15}$$

which finishes the proof of the theorem. □

*Acknowledgements.* This work was supported by the DFG Research Project ‘Stochastic Dynamics of Climate States’, DFG Research Center MATHEON in Berlin and the Japan Society for the Promotion of Science. I. Pavlyukevich thanks Th. Mikosch, G. Samorodnitsky, Th. Simon and N. Yoshida for their valuable comments. The authors are grateful to anonymous referees for careful reading and helpful suggestions which lead to a significant improvement of the manuscript.

## REFERENCES

- [1] N.H. Bingham, C.M. Goldie and J.L. Teugels, *Regular variation, Encyclopedia of Mathematics and its applications* **27**. Cambridge University Press, Cambridge (1987).
- [2] A. Bovier, M. Eckhoff, V. Gaynard, and M. Klein, Metastability in reversible diffusion processes I: Sharp asymptotics for capacities and exit times. *Eur. Math. Soc.* **6** (2004) 399–424.
- [3] A. Bovier, V. Gaynard and M. Klein, Metastability in reversible diffusion processes II: Precise asymptotics for small eigenvalues. *Eur. Math. Soc.* **7** (2005) 69–99.
- [4] V.A. Buslov and K.A. Makarov, Life times and lower eigenvalues of an operator of small diffusion. *Matematicheskie Zametki* **51** (1992) 20–31.
- [5] S. Cerrai, *Second order PDE’s in finite and infinite dimension. A probabilistic approach. Lect. Notes Math.* Springer, Berlin Heidelberg (2001).
- [6] A.V. Chechkin, V.Yu Gonchar, J. Klafter and R. Metzler, Barrier crossings of a Lévy flight. *EPL* **72** (2005) 348–354.
- [7] M.V. Day, On the exponential exit law in the small parameter exit problem. *Stochastics* **8** (1983) 297–323.
- [8] P.D. Ditlevsen, Anomalous jumping in a double-well potential. *Phys. Rev. E* **60** (1999) 172–179.
- [9] P.D. Ditlevsen, Observation of  $\alpha$ -stable noise induced millennial climate changes from an ice record. *Geophysical Research Letters* **26** (1999) 1441–1444.
- [10] M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems, Grundlehren der Mathematischen Wissenschaften* **260**. Springer, New York, NY, second edition (1998).
- [11] A. Galves, E. Olivieri and M.E. Vares, Metastability for a class of dynamical systems subject to small random perturbations. *Ann. Probab.* **15** (1987) 1288–1305.
- [12] V.V. Godovanchuk, Asymptotic probabilities of large deviations due to large jumps of a Markov process. *Theory Probab. Appl.* **26** (1982) 314–327.
- [13] P. Imkeller and I. Pavlyukevich, First exit times of SDEs driven by stable Lévy processes. *Stochastic Process. Appl.* **116** (2006) 611–642.
- [14] O. Kallenberg, *Foundations of modern probability. Probability and Its Applications.* Springer, second edition (2002).
- [15] C. Kipnis and C.M. Newman, The metastable behavior of infrequently observed, weakly random, one-dimensional diffusion processes. *SIAM J. Appl. Math.* **45** (1985) 972–982.
- [16] V.N. Kolokol’tsov and K.A. Makarov, Asymptotic spectral analysis of a small diffusion operator and the life times of the corresponding diffusion process. *Russian J. Math. Phys.* **4** (1996) 341–360.
- [17] P. Mathieu, Spectra, exit times and long time asymptotics in the zero-white-noise limit. *Stoch. Stoch. Rep.* **55** 1–20 (1995).
- [18] Ph.E. Protter, *Stochastic integration and differential equations, Applications of Mathematics* **21**. Springer, Berlin, second edition (2004).
- [19] G. Samorodnitsky and M. Grigoriu, Tails of solutions of certain nonlinear stochastic differential equations driven by heavy tailed Lévy motions. *Stoch. Process. Appl.* **105** (2003) 69–97.
- [20] A.D. Wentzell, *Limit theorems on large deviations for Markov stochastic processes, Mathematics and Its Applications (Soviet Series)* **38**. Kluwer Academic Publishers, Dordrecht (1990).
- [21] M. Williams, Asymptotic exit time distributions. *SIAM J. Appl. Math.* **42** (1982) 149–154.
- [22] Ai H. Xia, Weak convergence of jump processes, in *Séminaire de Probabilités, XXVI, Lect. Notes Math.* **1526** Springer, Berlin (1992) 32–46.