Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 118 (2008) 1071–1105

stochastic processes and their applications

www.elsevier.com/locate/spa

Simulated annealing for Lévy-driven jump-diffusions☆

Ilya Pavlyukevich*

Institut für Mathematik, Humboldt-Universität zu Berlin, Rudower Chaussee 25, 12489 Berlin, Germany

Received 5 October 2006; received in revised form 30 June 2007; accepted 22 July 2007
Available online 26 July 2007

Abstract

We consider a one-dimensional dynamical system driven by a vector field $-U'$, where U is a multi-well potential satisfying some regularity conditions. We perturb this dynamical system by a stable symmetric non-Gaussian Lévy process whose scale decreases as a power function of time. It turns out that the limiting behaviour of the perturbed dynamical system is different for slow and fast decrease rates of the noise intensity. As opposed to the well-studied Gaussian case, the support of the limiting law is not located in the set of global minima of U.

MSC: 60E07; 60F10

Keywords: Non-Gaussian stable Lévy process; Jump-diffusion; Heavy tail; Metastability; Extreme events; First exit time; Large deviations; Simulated annealing; Cooling rate

1. Introduction

Classical simulated annealing is a stochastic algorithm for determining the global minimum of an unknown function U. Let U be a real-valued function on $\mathbb{R}^d$ with a unique global minimum. The idea of the method consists in running a time-nonhomogeneous Gaussian-diffusion

$$d\hat{Z}(t) = -\nabla U(\hat{Z}(t))dt + \sigma(t)dW(t) \tag{1.1}$$

with some dispersion matrix that satisfies $\sigma(t) \to 0$ as $t \to +\infty$. Physically, the process $\hat{Z}$ describes an evolution of an overdamped Brownian particle in a potential energy landscape,

☆ This work was supported by the DFG research project *Stochastic Dynamics of Climate States*.
* Tel.: +49 30 2093 5845; fax: +49 30 2093 5848.
E-mail address: pavljuke@mathematik.hu-berlin.de.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2007.07.012

$\sigma(\cdot)$ being the temperature. For small values of $\sigma(\cdot)$ the process spends most of the time in the neighbourhoods of the potential's local minima, occasionally making transitions between adjacent wells. The goal of the simulated annealing consists in an appropriate choice of the cooling rate $\sigma(\cdot)$ so that for large values of t the diffusion $\hat{Z}(t)$ settles down in a neighbourhood of the global minimum of U.

More precisely, the main result concerning the simulated annealing of the process $\hat{Z}$ is as follows. To guarantee convergence, one should choose $\sigma^2(t) \approx \frac{\theta}{\ln(\lambda+t)}$ with a positive cooling rate θ, and some $\lambda > 1$ to parameterise the initial temperature $\sigma(0)$. Then, there is a critical value $\hat{\theta} > 0$ such that if $\theta > \hat{\theta}$ then $\hat{Z}(t)$ converges (in probability) to the coordinate of the global minimum of U, and the convergence fails otherwise. Moreover, the critical constant $\hat{\theta}$ is the logarithmic rate $-\lim_{\sigma\to 0}\sigma^2 \ln|\lambda^1_\sigma|$ of the principal nonzero eigenvalue λ^1_σ of the generator of the time-homogeneous small-noise diffusion

$$\mathrm{d}\hat{X}(t) = -\nabla U(\hat{X}(t))\mathrm{d}t + \sigma \mathrm{d}W(t). \tag{1.2}$$

Heuristic justification of this convergence is based on the observation that for small values of $\sigma(t)$ the process $\hat{Z}(t)$ behaves roughly like a time-homogeneous process $\hat{X}$ with a constant noise intensity $\sigma \approx \sigma(t)$. The principal nonzero eigenvalue λ^1_σ determines the convergence rate of $\hat{X}$ to its invariant measure which is proportional to $\exp(-2U/\sigma)$. The weak limit of this invariant measure as $\sigma \to 0$ is a Dirac mass at the potential's global minimum. Thus, if $\sigma(t)$ is such that $t|\lambda^1_{\sigma(t)}| \to \infty$ as $t \to \infty$, then $\hat{Z}(t)$ has enough time to settle down in the deepest potential well.

The asymptotic properties of the annealed process $\hat{Z}$ have been studied mathematically by many authors. We mention here the papers by Chiang, Hwang and Sheu [1] and Hwang and Sheu [2,3] who developed a small-noise analysis of the corresponding forward Kolmogorov equation and applied the Freidlin–Wentzell theory of randomly perturbed dynamical systems (see [4]). On the other hand, Holley and Stroock in [5], and Holley, Kusuoka and Stroock [6] proposed a Dirichlet form approach. We also refer the reader to the review paper [7] and further references therein.

Simulated annealing has proved to be an effective optimisation tool with a solid theoretical justification. However, its practical implementation has several negative features. First, it is usually difficult to determine the cooling schedule $\hat{\theta}$ – or equivalently the logarithmic order of the first eigenvalue λ^1_σ – without a detailed information on the potential function U. In practice, the cooling schedule is set by the method of trials and errors so that the optimisation procedure yields satisfactory results. The second problem is the convergence rate. The logarithmic convergence of $\sigma(t)$ to zero leads to time-consuming optimisation procedures. Finally, due to the continuity of the trajectories of $\hat{Z}$, the search for the global minimum becomes slow if U has many wells and the initial value $\hat{Z}(0)$ is chosen 'far' from the domain of attraction of the global minimum. We refer the reader to the papers by Sorkin [8], Ingber [9] and Locatelli [10] for the discussion on the restrictions of simulated annealing.

There is a number of papers devoted to various modifications of the classical algorithm. For example, Szu and Hartley in their physical paper [11] introduced the so-called *fast simulated annealing* which allows to perform a non-local search of the deepest well. A fast simulated annealing process in the sense of [11] is a discrete time Markov chain where the next state is obtained from the Euler approximation of (1.1) driven not by Gaussian noise but by *Cauchy* noise. The new state is accepted according to the Metropolis algorithm introduced in [12]. The acceptance probability is 1 if the potential value in the new state is smaller, i.e. the new position is 'lower' in the potential landscape. If the new position is 'higher', it is accepted with the probability $\sim \exp(-\Delta U/\sigma)$, where ΔU is the difference of the potential values in the new and the old states, and σ is a decreasing temperature parameter. The advantage of this method consists in faster transitions between the potential wells due to the heavy tails of the Cauchy distribution. Moreover, the authors claim that the optimal cooling rate is algebraic, i.e. $\sigma(t) \approx t^{-1}$ which also accelerates convergence. There are several practical implementations of this algorithm, e.g. by Szu [13] and Nascimento et al. [14]. However, to our knowledge it is not proved whether it really converges to the global minimum of U.

In this paper we consider a continuous-time annealed jump-diffusion driven by a symmetric stable Lévy process with heavy tails. Our goal is to determine a cooling schedule $\sigma(\cdot)$ so that the jump-diffusion converges to some nontrivial limit. We alert the reader that this limit (when it exists) will not be a Dirac mass at the global minimum of U.

This paper can be seen as a sequel of [15–17] where the small-noise dynamics of heavy-tailed Lévy-driven jump-diffusions were studied. We emphasize that our methods are purely probabilistic.

2. 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 hypothesis in the sense of Protter [18], i.e. $\mathcal{F}_0$ contains all the $\mathbf{P}$-null sets of $\mathcal{F}$ and is right continuous. We consider a time-nonhomogeneous process $Z^\lambda = (Z^\lambda_{s,z}(t))_{t\geq s}$ satisfying the following one-dimensional stochastic differential equation

$$Z^\lambda_{s,z}(t) = z - \int_s^t U'(Z^\lambda_{s,z}(u-))\mathrm{d}u + \int_s^t \frac{\mathrm{d}L(u)}{(\lambda+u)^\theta}, \quad 0 \leq s \leq t,\ z \in \mathbb{R}, \lambda, \theta > 0, \tag{2.3}$$

where L is a Lévy process and U is a potential function. We make the following assumptions.
Assumptions on L:

L The process L is a symmetric α-stable process, $\alpha \in (0, 2)$, whose marginals have the Lévy-Hinchin representation

$$\ln \mathbf{E}\mathrm{e}^{\mathrm{i}\lambda L(t)} = t\int (\mathrm{e}^{\mathrm{i}\lambda y} - 1 - \mathrm{i}\lambda y \mathbb{I}\{|y| \leq 1\}) \frac{\mathrm{d}y}{|y|^{1+\alpha}} = -tc(\alpha)|\lambda|^\alpha,$$
$$c(\alpha) = 2\left|\cos\left(\frac{\pi\alpha}{2}\right)\Gamma(-\alpha)\right|. \tag{2.4}$$

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

$$-\infty = s_0 < m_1 < s_1 < m_2 < \cdots < s_{n-1} < m_n < s_n = +\infty. \tag{2.5}$$

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

U3 $|U'(x)| > |x|^{1+c}$ as $x \to \pm\infty$ for some $c > 0$.

As a time-homogeneous counterpart of Z^λ, we consider the process $X^\varepsilon = (X^\varepsilon_x(t))_{t\geq 0}$ satisfying the following stochastic differential equation

$$X^\varepsilon_x(t) = x - \int_0^t U'(X^\varepsilon_x(u-))\mathrm{d}u + \varepsilon L(t), \quad t \geq 0,\ x \in \mathbb{R}, \tag{2.6}$$

where ε is a positive parameter. For small values of ε, the behaviour of X^ε was studied in our previous papers [15–17]. We also refer the reader to the works by Godovanchuk [19] and Wentzell [20] for results on the large deviations due to large jumps.

Both processes X^ε and Z^λ can be seen as a perturbation of the deterministic dynamical system

$$X_x^0(t) = x - \int_0^t U'(X_x^0(u))\mathrm{d}u, \quad t \geq 0,\ x \in \mathbb{R}, \tag{2.7}$$

by $\varepsilon L(t)$ and $\int_0^t \frac{\mathrm{d}L(u)}{(\lambda+u)^\theta}$ respectively. Under the above assumptions on U, the underlying deterministic equation (2.7) 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 $X_t^0(x) \to m_i$ as $t \to \infty$. It is clear that the deterministic solution X^0 does not leave the domain of attraction in which it started.

From the physical point of view, Eqs. (2.3) and (2.6) describe the motion of an overdamped Lévy particle in a potential energy landscape. The jump magnitude is parameterised by the 'instant temperature' $(\lambda + t)^{-\theta}$ and ε respectively. In Eq. (2.3), the initial temperature $(t = s)$ equals $(\lambda + s)^{-\theta}$. A positive cooling rate θ determines the speed of temperature decrease.

Since the Lévy process L is a semimartingale, the stochastic differential equations (2.3) and (2.6) are well defined, see [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 (2.3), which can be done analogously to the time-homogeneous case considered in [17]. The processes X^ε and Z^λ are also strong Markov and Feller.

It is necessary to notice that the evolution of the process starting at time $s \geq 0$ is the same as that of the process starting at time zero with a different initial temperature, namely

$$(Z_{s,z}^\lambda(s+t))_{t\geq 0} \stackrel{d}{=} (Z_{0,z}^{\lambda+s}(t))_{t\geq 0}, \tag{2.8}$$

and thus the particular values of s or λ do not influence asymptotic properties of the jump-diffusion in the limit $t \to \infty$. However, since our theory will work for low temperatures, it is often convenient to study the jump-diffusion not for large values of s and t but for large values of λ, and we do not omit λ in our notation.

As we have seen in the Gaussian case, a good candidate for the limiting distribution of $Z^\lambda(t)$ as $t \to \infty$ can be found among the limiting distributions of the time-homogeneous jump-diffusion X^ε as $\varepsilon \to 0$. Although no closed-form formula for the invariant distribution of X^ε is known, one can use the metastability results for X^ε, which state that for small values of ε the jump-diffusion reminds of a continuous-time Markov chain on the set of stable attractors of U. Indeed, the following theorem holds true.

Theorem 2.1 (*metastability, [17]*). *Let X^ε be a solution of (2.6). If $x \in (s_{i-1}, s_i)$, $1 \leq i \leq n$, then for $t > 0$*

$$X_x^\varepsilon(\varepsilon^{-\alpha} t) \to Y_{m_i}(t), \quad \varepsilon \downarrow 0, \tag{2.9}$$

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

$$\begin{aligned} q_{ij} &= \alpha^{-1}||s_{j-1} - m_i|^{-\alpha} - |s_j - m_i|^{-\alpha}|, \quad i \neq j, \\ -q_{ii} &= q_i = \sum_{j\neq i} q_{ij} = \alpha^{-1}(|s_{i-1} - m_i|^{-\alpha} + |s_i - m_i|^{-\alpha}). \end{aligned} \tag{2.10}$$

It is easy to see from the form of Q that the limiting Markov chain Y is irreducible and positive-recurrent. Denote by $\pi^0(\mathrm{d}y) = \sum_{j=1}^n \pi_j^0 \delta_{m_j}(\mathrm{d}y)$ its unique stationary measure, $\delta_x(\mathrm{d}y)$ being a Dirac δ-function. This measure π^0 is the candidate for the limiting distribution of $Z^\lambda(t)$, $t \to +\infty$. However, as in the Gaussian case, the convergence depends on the cooling rate θ.

Our main result is formulated in the following theorems. For a measurable bounded function f denote $\pi^0 f = \int f(x)\pi^0(\mathrm{d}x) = \sum_{j=0}^n f(m_j)\pi_j^0$.

Theorem 2.2 (*slow cooling*). *Let Z^λ be a solution of (2.3). Let $\alpha\theta < 1$. Then for any $\lambda > 0$, $z \in \mathbb{R}$ and any continuous and bounded function f*

$$\mathbf{E}_{0,z} f(Z^\lambda(t)) \to \pi^0 f, \quad t \to \infty. \tag{2.11}$$

If the cooling rate θ is above the threshold $1/\alpha$, the solution Z^λ gets trapped in one of the wells and thus the convergence fails. Consider the first exit time from the ith well $\tilde{\sigma}_z^i(\lambda) = \inf\{t \geq 0 : Z_{0,z}^\lambda \notin (s_{i-1}, s_i)\}$. Then the following trapping result holds.

Theorem 2.3 (*fast cooling*). *Let Z^λ be a solution of (2.3). Let $\alpha\theta > 1$. There is $\Delta > 0$ such that for any $i = 1, \ldots, n$,*

$$\mathbf{P}_{0,z}(\tilde{\sigma}^i(\lambda) < \infty) = \mathcal{O}(\lambda^{1-\alpha\theta}), \quad \lambda \to \infty, \tag{2.12}$$

uniformly for $|z - m_i| \leq \Delta$.

As we see, the annealing procedure does not locate the global minimum of U, but reveals some information on the spatial structure of the potential. For instance, if the coordinates of the local minima are known or can be estimated, then with help of the invariant measure π^0 one can estimate the coordinates of the saddle points and thus the sizes of the domains of attraction. This information can be useful for the global optimisation, since often the spatially biggest well is the deepest at the same time, see [10,21].

To make calculations simpler, we consider only jump-diffusions driven by symmetric α-stable non-Gaussian Lévy processes in this paper. In fact, our results are also expect to hold for jump-diffusions driven by more general Lévy processes with generating triplet (G, ν, μ), where $G \geq 0$ is the Gaussian variance, $\mu \in \mathbb{R}$ is the drift, and the Lévy measure ν has regularly varying tails of the same index and is not necessarily symmetric. In [17], Theorem 2.1 is proven in such a setting.

The paper is organised as follows. In a preparatory Section 3 we decompose the driving process L into small- and big-jump parts and study a one-well dynamics of the process Z^λ for large values of λ. Here we use results from our paper [17]. Then we study the slow cooling case $\alpha\theta < 1$. Section 4 is devoted to the study of the mean first exit time from a single well. This is technically the most complex part of the proof. In Section 5 we determine mean transition times between the neighbourhoods of the wells' minima and the corresponding transition probabilities. In Section 6 we construct an embedded nonhomogeneous discrete-time Markov chain whose distribution is asymptotically close to π^0. In Section 7 we prove Theorem 2.2. The case of fast cooling is short and is considered in Section 8. Finally, some calculations of Laplace integrals are collected in the Appendix.

We use the notation $\mathbb{I}\{A\}$ for the indicator function of the set A and $\mathbb{I}\{A, B\} = \mathbb{I}\{A\}\mathbb{I}\{B\}$. The complement of a set A is denoted by A^c. In our estimates, we shall use c_1, c_2, etc. to denote positive constants without calculating their explicit values. On the other hand, c will denote the constant from the estimate (3.25). As usual, $\Delta X(t) = X(t) - X(t-)$.

Finally, we direct the reader's attention to our paper [16], where we explain the results of Theorems 2.2 and 2.3 on the physical level of rigour in the spirit of our previous work [22] and illustrate them with some simulations.

Of course, the most intriguing question remains, namely whether one can construct a time-nonhomogeneous jump-diffusion which settles down near the global minimum of the potential? The answer to this question is affirmative and the numerical algorithm can be found in [23]. The rigorous mathematical proof of its convergence will be presented in a forthcoming paper.

3. One-well dynamics of the annealed process with small jumps

3.1. Small and big jumps of L

For $\lambda > 0$, consider the decomposition $L = \xi^\lambda + \eta^\lambda$, where the Lévy processes ξ^λ and η^λ have generating triplets $(0, \nu_\xi^\lambda, 0)$ and $(0, \nu_\eta^\lambda, 0)$ with Lévy measures $\nu_\xi(\cdot) = \nu(\cdot \cap |y| \le \lambda^{\theta/2})$ and $\nu_\eta(\cdot) = \nu(\cdot \cap |y| > \lambda^{\theta/2})$. The process ξ^λ has a Lévy measure with compact support, and the Lévy measure of η^λ is finite. Denote

$$\beta = \beta_\lambda = \nu_\eta(\mathbb{R}) = \int_{\mathbb{R}\setminus[-\lambda^{\theta/2},\lambda^{\theta/2}]} \frac{\mathrm{d}y}{|y|^{1+\alpha}} = \frac{2}{\alpha}\lambda^{-\alpha\theta/2}. \tag{3.13}$$

Then, η^λ is a compound Poisson process with intensity β_λ and jumps distributed according to the law $\beta_\lambda^{-1}\nu_\eta^\lambda(\cdot)$. Denote by $(\tau_k)_{k\ge 0}$ the arrival times of η^λ with $\tau_0 = 0$. Let $T_k = \tau_k - \tau_{k-1}$ denote the successive inter-jump periods, and $W_k = \eta^\lambda_{\tau_k} - \eta^\lambda_{\tau_k-}$ the jump sizes of η^λ. Then, the three processes $(T_k)_{k\ge1}$, $(W_k)_{k\ge1}$ and $(\xi^\lambda)_{t\ge0}$ are independent. Moreover,

$$\mathbf{P}(T_k \ge u) = \int_u^\infty \beta \mathrm{e}^{-\beta s}\mathrm{d}s = \mathrm{e}^{-\beta u}, \quad u \ge 0, \quad \text{and} \quad \mathbf{E}T_k = \frac{1}{\beta}, \tag{3.14}$$

$$\mathbf{P}(W_k < w) = \frac{\nu_\eta^\lambda(-\infty, w)}{\nu_\eta^\lambda(\mathbb{R})} = \frac{1}{\beta}\int_{(-\infty,w)} \mathbb{I}\{|y| > \lambda^{\theta/2}\}\frac{\mathrm{d}y}{|y|^{1+\alpha}}, \quad w \in \mathbb{R}. \tag{3.15}$$

Of course, all random variables introduced above depend on λ. We shall usually suppress λ for brevity.

Due to the strong Markov property, for any stopping time τ the process $\xi^\lambda_{t+\tau} - \xi^\lambda_\tau$, $t \ge 0$, is also a Lévy process with the same law as ξ^λ. For $\lambda > 0$ and $\lambda_1 \ge 0$, consider a small jump-diffusion $z^{\lambda,\lambda_1} = (z^{\lambda,\lambda_1}_{s,z}(t))_{t\ge s}$, $s \ge 0$, $z \in \mathbb{R}$, satisfying the following equation

$$z^{\lambda,\lambda_1}_{s,z}(t) = z - \int_s^t U'(z^{\lambda,\lambda_1}_{s,z}(u-))\mathrm{d}u + \int_s^t \frac{\mathrm{d}\xi^\lambda_u}{(\lambda+\lambda_1+u)^\theta}. \tag{3.16}$$

For $j \ge 1$, consider the processes $\xi^{\lambda,j}_t = \xi^\lambda_{t+\tau_{j-1}} - \xi^\lambda_{\tau_{j-1}}$ and

$$z^\lambda_j(t,z) = z - \int_{\tau_{j-1}}^t U'(z^\lambda_j(u-,z))\mathrm{d}u + \int_{\tau_{j-1}}^t \frac{\mathrm{d}\xi^{\lambda,j}_u}{(\lambda+u)^\theta}, \quad t \ge \tau_{j-1}. \tag{3.17}$$

The processes $\xi^{\lambda,j}$ are independent copies of the ξ^λ and the strong Markov property implies that $(z^\lambda_j(t+\tau_{j-1},z))_{t\ge0} \overset{d}{=} (z^{\lambda,\tau_{j-1}}_{0,z}(t))_{t\ge0}$. In this notation, for $z \in \mathbb{R}$ and $k \ge 1$, the solution of (2.3) has the following representation:

$$Z_{0,z}^{\lambda}(t+\tau_{k-1}^{\lambda}) = z_k^{\lambda}(t+\tau_{k-1}, Z_{0,z}^{\lambda}(\tau_{k-1}^{\lambda})) + \frac{W_k}{(\lambda+\tau_k)^{\theta}}\mathbb{I}\{t=T_k\}, \quad t\in[0,T_k]. \tag{3.18}$$

The random path Z^{λ} can be represented as a sum of a Poisson process with big time-dependent jump sizes and small jump-diffusions living on the exponentially distributed independent inter-arrival time intervals. Since the jump sizes $|\Delta\xi_u^{\lambda,j}|/(\tau_{j-1}+\lambda+u)^{\theta} \to 0$ as $\lambda\to\infty$ for all $j\geq 0$ and $u\geq 0$, we can consider the processes z_j^{λ} and z^{λ,λ_1} as small random perturbations of the deterministic system X^0 defined by (2.7). We are going to estimate the distance between z^{λ,λ_1} and X^0 in terms of the small-jump process ξ^{λ}.

3.2. *Exponential estimate on ξ^{λ}*

Lemma 3.1. *Let $\theta>0$. Then there exist constants $\gamma_0>0$, $q_0>\alpha\theta/2$ and $w>0$ such that for any $0<\gamma\leq\gamma_0$ and $0\leq q\leq q_0$ there is $\lambda_0>0$ such that the inequality*

$$\mathbf{P}\left(\sup_{t\in[0,\lambda^q]}|(\lambda+\lambda_1)^{-\theta}\xi_t^{\lambda}|\geq(\lambda+\lambda_1)^{-\gamma}\right)\leq\exp(-(\lambda+\lambda_1)^{w})\leq\exp(-\lambda^{w}) \tag{3.19}$$

holds for all $\lambda\geq\lambda_0$ and $\lambda_1\geq 0$.

Proof. The statement is obtained with help of the reflection principle for symmetric Lévy processes and Chebyshev's inequality analogously to Lemma 2.1 in [17]. □

3.3. *Random perturbations by ξ^{λ}*

In this section we study the one-well dynamics of the small-jump process z^{λ,λ_1} defined in (3.16) in the limit of large values of λ. We are going to use the estimates from [17] for the time-homogeneous counterpart of z^{λ,λ_1}. Inspecting the argument of Section 3 in [17], one notes that path-wise estimates obtained there depend only on the size of random perturbations on finite time intervals. In the present time-dependent case, we can also estimate the perturbation in terms of a Lévy process ξ^{λ}. Indeed, fix $T\geq 0$. Then for any $0\leq t\leq T$, $\lambda>0$ and $\lambda_1\geq 0$, integration by parts yields

$$\left|\int_0^t\frac{\mathrm{d}\xi_s^{\lambda}}{(\lambda+\lambda_1+u)^{\theta}}\right|\leq\frac{2}{\lambda^{\theta}}\sup_{u\in[0,T]}|\xi_u^{\lambda}|. \tag{3.20}$$

Thus, we are in the setting of the paper [17] and can borrow the following results. For definiteness, we assume that the well's minimum is located at the origin and thus the corresponding domain of attraction for $X_x^0(\cdot)$ is a finite interval $I=(a,b)$, $-\infty<a<0<b<+\infty$, if the well is inner, and $I=(-\infty,b)$ if it is peripheral. In the first case, we also assume that a and b are nondegenerate local maxima of U. In the second case, b is a nondegenerate local maximum of U, and $|U'(x)|$ increases to infinity faster than linearly as $x\to-\infty$ according to assumption **U3**.

Let $\gamma>0$. For $\lambda>0$ sufficiently large, we define subintervals $I_1:=[a+\lambda^{-\gamma},b-\lambda^{-\gamma}]$ and $I_2:=[a+\lambda^{-\gamma}+\lambda^{-2\gamma},b-\lambda^{\gamma}-\lambda^{-2\gamma}]$ if $a>-\infty$ (inner well), and $I_1=(-\infty,b-\lambda^{-\gamma}]$ and

$I_2 = (-\infty, b - \lambda^{-\gamma} - \lambda^{-2\gamma}]$ if $a = -\infty$ (peripheral well). For $t \geq 0$ we introduce events

$$\mathcal{E}^{\lambda,\lambda_1}(t) := \left\{\omega : \sup_{s\in[0,t]} |(\lambda+\lambda_1)^{-\theta}\xi_s^\lambda| \leq (\lambda+\lambda_1)^{-4\gamma}\right\}, \qquad \mathcal{E}^{\lambda}(t) := \mathcal{E}^{\lambda,0}(t). \tag{3.21}$$

Obviously, if $0 < 4\gamma \leq \theta$, $\lambda \geq 1$ and $0 \leq \lambda_1 \leq \lambda_2$, then

$$\mathcal{E}^{\lambda}(t) \subseteq \mathcal{E}^{\lambda,\lambda_1}(t) \subseteq \mathcal{E}^{\lambda,\lambda_2}(t), \quad t \geq 0. \tag{3.22}$$

In the case of a finite interval (a, b), we consider an event

$$E_z^{\lambda,\lambda_1}(t) = \left\{\sup_{s\in[0,t]} |z_{0,z}^{\lambda,\lambda_1}(s) - X_z^0(s)| \leq \frac{1}{2(\lambda+\lambda_1)^{2\gamma}}\right\}. \tag{3.23}$$

In case of a peripheral well $(-\infty, b)$, we consider the dynamics of $z_{0,z}^{\lambda,\lambda_1}$ separately for initial values in $(-\infty, -A]$ and $[-A, b]$ for some A big enough. For $z \leq b$, consider the first entrance time to $[A, b]$, namely $\tau_A^{\lambda,\lambda_1}(z) = \inf\{t \geq 0 : z_{0,z}^{\lambda,\lambda_1}(t) \geq -A\}$. Then there exists a constant T_A such that $\tau_A^{\lambda,\lambda_1}(z) \leq T_A$ a.s. on $\mathcal{E}^{\lambda,\lambda_1}(T_A)$ for all $z \leq b$ [17, Lemma 2.3]. Then we define an event $E_z^{\lambda,\lambda_1}(t)$ by

$$\begin{aligned} E_z^{\lambda,\lambda_1}(t) = \Bigg\{ &\sup_{s\in[0,\tau_A\wedge t]} z_{0,z}^{\lambda,\lambda_1}(s) \leq -A+1 \text{ and} \\ &\sup_{s\in[\tau_A\wedge t,t]} |z_{0,z}^{\lambda,\lambda_1}(s) - X^0_{z^{\lambda,\lambda_1}(\tau_A)}(s-\tau_A\wedge t)| \leq \frac{1}{2(\lambda+\lambda_1)^{2\gamma}}\Bigg\}. \end{aligned} \tag{3.24}$$

Now as Section 2 in [17] we obtain the following result.

Lemma 3.2. *For any $\gamma > 0$, there is $\lambda_0 > 0$ such that the inclusion $\mathcal{E}^{\lambda,\lambda_1}(t) \subseteq E_z^{\lambda,\lambda_1}(t)$ holds almost surely for $\lambda \geq \lambda_0$ and $\lambda_1 \geq 0$ uniformly for $t \geq 0$ and $z \in [a+(\lambda+\lambda_1)^{-\gamma}, b-(\lambda+\lambda_1)^{-\gamma}]$ if the well is inner or $z \in (-\infty, b-(\lambda+\lambda_1)^{-\gamma}]$ if the well is peripheral.*

Lemma 3.2 compares the trajectories of the small jump-diffusion z_z^{λ,λ_1} with the underlying deterministic trajectory X_z^0 in terms of the driving process ξ^λ, particularly on the event $\mathcal{E}^{\lambda,\lambda_1}(t)$, when ξ^λ does not essentially deviate from zero. Indeed, if the well is inner, then the random path $z_{0,z}^{\lambda,\lambda_1}$ is contained in a $1/(2(\lambda+\lambda_1)^{2\gamma})$-neighbourhood of the deterministic trajectory X_z^0. If the well is peripheral, we have to take into account initial values z which are close to $-\infty$, see definition (3.24). If $z \leq -A$, the process $z_{0,z}^\lambda$ enters the compact $[-A, b]$ in a.s. finite time first and then follows the deterministic trajectory starting at the entrance point of $z_{0,z}^{\lambda,\lambda_1}$. If $z \geq -A$, the dynamics is the same as in the inner-well case.

3.4. *One-well behaviour of z^{λ,λ_1}*

In this section we exploit the properties of the deterministic trajectory X^0 to show that on appropriate time intervals the small jump-process z^{λ,λ_1} does not leave the well and settles near its local minimum with high probability.

From the nondegeneracy assumptions **U2** and **U3** on the potential we can easily see that for any $\gamma > 0$, any $p > 0$ and some $c > 0$ the estimates

$$\begin{aligned} &|X^0_z(t)| \leq \lambda^{-2\gamma}/2, \quad t \geq \lambda^p,\ z \in I_1, \\ &X^0_z(t) \in (a + \lambda^{-\gamma} + 2\lambda^{-2\gamma}, b - \lambda^{-\gamma} - 2\lambda^{-2\gamma}), \quad t \geq c\lambda^{-\gamma},\ z \in I_1, \end{aligned} \tag{3.25}$$

hold for λ sufficiently large.

For $j \geq 1$, define events

$$\begin{aligned} C^j_y &= \left\{ z^{\lambda,\tau_{j-1}}_{0,y}(s) \in I_1, s \in [0, T_j), z^{\lambda,\tau_{j-1}}_{0,y}(T_j) + \frac{W_j}{(\lambda + \tau_j)^\theta} \in I_1 \right\}, \\ C^{j,-}_y &= \left\{ z^{\lambda,\tau_{j-1}}_{0,y}(s) \in I_1, s \in [0, T_j), z^{\lambda,\tau_{j-1}}_{0,y}(T_j) + \frac{W_j}{(\lambda + \tau_j)^\theta} \in I_2 \right\}, \\ \widetilde{C}^j_y &= \left\{ z^{\lambda,\tau_{j-1}}_{0,y}(s) \in I_1, s \in [0, T_j), z^{\lambda,\tau_{j-1}}_{0,y}(T_j) + \frac{W_j}{(\lambda + \tau_j)^\theta} \notin I_1 \right\}. \end{aligned} \tag{3.26}$$

Recalling definitions (3.23) and (3.24) denote also $E^j_y = E^{\lambda,\tau_{j-1}}_y(T_j)$, $j \geq 1$.

Lemma 3.3. 1. *For any $\gamma > 0$, $p > 0$ and λ sufficiently large, the inclusions*

$$C^j_y E^j_y \cap \{T_j \geq \lambda^p\} \subseteq \left\{ \frac{W_j}{(\lambda + \tau_k)^\theta} \in I \right\} \cap \{T_j \geq \lambda^p\} \subseteq \left\{ \frac{W_k}{(\lambda + \tau_k)^\theta} \in I \right\}, \tag{3.27}$$

$$\widetilde{C}^k_y E^k_y \cap \{T_k \geq \lambda^p\} \subseteq \left\{ \frac{W_k}{(\lambda + \tau_k)^\theta} \notin I_2 \right\} \cap \{T_k \geq \lambda^p\} \subseteq \left\{ \frac{W_k}{(\lambda + \tau_k)^\theta} \notin I_2 \right\}, \tag{3.28}$$

hold uniformly for $y \in I_1$ and $1 \leq j \leq k$, $k \geq 1$.

2. *For any $\gamma > 0$ and λ sufficiently large the inclusion*

$$C^j_y E^j_y \cap \left\{ \frac{|W_j|}{\lambda^\theta} \leq \frac{1}{2\lambda^{2\gamma}} \right\} \cap \left\{ T_j \geq \frac{c}{\lambda^\gamma} \right\} \supseteq E^j_y \cap \left\{ \frac{|W_j|}{\lambda^\theta} \leq \frac{1}{2\lambda^{2\gamma}} \right\} \cap \left\{ T_j \geq \frac{c}{\lambda^\gamma} \right\} \tag{3.29}$$

holds uniformly for $y \in I_1$ and $j \geq 1$.

Proof. The statement follows directly from Lemma 3.2, estimates (3.25) and a.s. monotonicity of the arrival times τ_k, $k \geq 0$. □

Finally, the exponential estimate of Lemma 3.1, inclusion (3.22) and Lemma 3.2 imply that the probability of the event E^j_y is exponentially close to 1, if the interval T_j is not too long. For $q > 0$, let $\mathcal{E}_j = \{\omega : \sup_{s\in[0,\lambda^q]} |\lambda^{-\theta}\xi^{\lambda,j}_s| \leq \lambda^{-4\gamma}\}$.

Lemma 3.4. *There are constants $\gamma_0 > 0$, $q_0 > \alpha\theta/2$ and $w > 0$ such that for any $0 < \gamma \leq \gamma_0$ and $0 \leq q \leq q_0$, there is $\lambda_0 > 0$ such that for all $\lambda \geq \lambda_0$ we have the inclusion $(E^j_y)^c \cap \{T_j \leq \lambda^q\} \subseteq \mathcal{E}^c_j$. Consequently, the estimate $\mathbf{P}((E^j_y)^c, T_j \leq \lambda^q) \leq \mathbf{P}(\mathcal{E}^c_j) \leq \exp(-\lambda^w)$ holds uniformly for $y \in I_1$ and $j \geq 1$.*

4. Exit from a single well, $\alpha\theta < 1$

For $i = 1, \ldots, n$, consider the wells of the potential U with local minima at m_i and denote $\Omega^i = (s_{i-1}, s_i)$. For $\gamma > 0$ and $\lambda > 0$, consider the λ-dependent inner neighbourhoods of the wells $\Omega^i_\lambda = [s_{i-1} + 2\lambda^{-\gamma}, s_i - 2\lambda^{-\gamma}]$, where by convention $\Omega^1_\lambda = (-\infty, s_1 - 2\lambda^{-\gamma}]$ and $\Omega^n_\lambda = [s_{n-1} + 2\lambda^{-\gamma}, +\infty)$. We assume that λ is big enough so that $\Omega^i_\lambda \subset \Omega^i$.

Consider the following life times of the process Z^λ in the potential wells Ω^i:

$$\sigma^{i,\lambda}_{s,z} = \inf\{t \geq s : Z^\lambda_{s,z}(t) \notin [s_{i-1} + \lambda^{-\gamma}, s_i - \lambda^{-\gamma}]\}, \quad s \geq 0,\ 1 \leq i \leq n. \tag{4.30}$$

In this section we prove the following proposition.

Proposition 4.1. *Let $\alpha\theta < 1$. Then there is $\gamma_0 > 0$ such that for any $0 < \gamma \leq \gamma_0$ there is $\delta > 0$ such that for all $i, j = 1, \ldots, n$, $i \neq j$,*

$$\begin{aligned} &\mathbf{E}_{s,z}\sigma^{i,\lambda} - s = q_i^{-1}(\lambda + s)^{\alpha\theta}(1 + \mathcal{O}(\lambda^{-\delta})), \\ &\mathbf{P}_{s,z}(Z^\lambda(\sigma^{i,\lambda}) \in \Omega^j_\lambda) = q_{ij} q_i^{-1}(1 + \mathcal{O}(\lambda^{-\delta})), \quad \lambda \to +\infty, \end{aligned} \tag{4.31}$$

uniformly for $s \geq 0$ and $z \in \Omega^i_\lambda$.

Note that $\sigma^{i,\lambda}$ is the exit time from a well, which is at $\lambda^{-\gamma}$ 'bigger' than Ω^i_λ: to obtain uniform estimates over $z \in \Omega^i_\lambda$ for life times and transition probabilities we have to separate the initial points z from the well's boundary.

It follows from (2.8) that it suffices to prove Proposition 4.1 for $s = 0$. For simplicity, we consider a well $I = (a, b)$ with a local minimum at zero as defined in Section 3.3. Thus, in what follows we study the exit time $\sigma_{0,z}(\lambda) = \inf\{t \geq 0 : Z^\lambda_{0,z} \notin I_1\}$.

4.1. Mean exit time — Estimate from above

Proposition 4.2. *Let $\alpha\theta < 1$. Then there exists $\gamma_0 > 0$ such that for any $0 < \gamma \leq \gamma_0$ there is $\delta > 0$ such that for λ sufficiently large*

$$\mathbf{E}_{0,z}\sigma(\lambda) \leq \frac{\alpha}{|a|^{-\alpha} + b^{-\alpha}} \lambda^{\alpha\theta}(1 + \lambda^{-\delta}) \tag{4.32}$$

uniformly for $z \in I_2$.

4.1.1. Notation and technicalities

For $0 < p < q$, $\gamma > 0$, $\lambda \geq 1$ and $1 \leq i \leq k$ introduce events $A_k = \{\lambda^p \leq T_k \leq \lambda^q\}$, $B_k = \{\lambda^{-\gamma/2} \leq T_k < \lambda^p\}$, $D_k = \{k\lambda^p \leq \tau_k \leq k\lambda^q\}$, and $H_{i,k} = \{\frac{|W_i|}{(\lambda + k\lambda^p)^\theta} \leq \frac{1}{\lambda^{2\gamma}}\}$. Consider also events

$$G_{i,k} = A_i \sqcup \big(B_i \cap H_{i,k}\big) \quad \text{and} \quad G^c_{i,k} = (A_i \sqcup B_i)^c \sqcup (B_i \cap H^c_{i,k}), \tag{4.33}$$

where "$\sqcup$" emphasizes that we take a union of disjoint sets. Representation (4.33) follows from the disjointness of A_i and B_i. In what follows we shall need the estimates

$$r := \mathbf{P}((A_i \sqcup B_i)^c) \leq c_r \lambda^{-\alpha\theta/2 - \gamma/2} \quad \text{and} \quad \mathbf{P}(B_i)\mathbf{P}(H^c_{i,k}) \leq c_r \frac{\lambda^{p + 2\alpha\gamma}}{(\lambda + k\lambda^p)^{\alpha\theta}}, \tag{4.34}$$

which hold for some $c_r > 0$.

For $k \geq 1$ and $1 \leq i \leq k$, let $F_{i,k} = G_{i,k}$ or $G^c_{i,k}$. Then the total probability formula leads to the following inequality:

$$
\begin{aligned}
&\mathbf{E}_{0,z}\sigma(\lambda)\\
&\quad\le \sum_{k=1}^{\infty}\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{\sigma(\lambda)=\tau_k\}]+\sum_{k=1}^{\infty}\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{\sigma(\lambda)\in(\tau_{k-1},\tau_k)\}]+\sum_{k=1}^{\infty}\mathbf{E}[\tau_k\mathbb{I}\{D_k^c\}]\\
&\quad=\sum_{k=1}^{\infty}\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{\sigma(\lambda)=\tau_k\}\mathbb{I}\{G_{1,k},\dots,G_{k,k}\}]\\
&\qquad+\sum_{k=1}^{\infty}\sum_{(F_{1,k},\dots,F_{k-1,k})\neq(G_{1,k},\dots,G_{k-1,k})}\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{\sigma(\lambda)=\tau_k\}\mathbb{I}\{F_{1,k},\dots,F_{k-1,k},G_{k,k}\}]\\
&\qquad+\sum_{k=1}^{\infty}\sum_{F_{1,k},\dots,F_{k-1,k}}\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{D_k\}(\mathbb{I}\{\sigma(\lambda)=\tau_k\}+\mathbb{I}\{\sigma(\lambda)\in(\tau_{k-1},\tau_k)\})\mathbb{I}\{F_{1,k},\dots,F_{k-1,k},G_{k,k}^c\}]\\
&\qquad+\sum_{k=1}^{\infty}\sum_{F_{1,k},\dots,F_{k-1,k}}\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{\sigma(\lambda)\in(\tau_{k-1},\tau_k)\}\mathbb{I}\{F_{1,k},\dots,F_{k-1,k},G_{k,k}\}]\\
&\qquad+\sum_{k=1}^{\infty}\mathbf{E}[\tau_k\mathbb{I}\{D_k^c\}]=:M(z,\lambda)+R_1(z,\lambda)+R_2(z,\lambda)+R_3(z,\lambda)+R_4(\lambda).
\end{aligned}
\tag{4.35}
$$

We show that the main contribution to the mean value of $\sigma(\lambda)$ is made by the main term M, whereas the remainder terms R_i, $i=1,\dots 4$, are negligible as $\lambda\to\infty$.

To estimate the summands in (4.35) for $z\in I_1$ and $k\ge 1$, the following chain of inequalities is deduced, which results in a factorisation formula

$$
\begin{aligned}
&\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{\sigma=\tau_k\}\mathbb{I}\{F_{1,k},\dots,F_{k,k}\}]\\
&\quad=\mathbf{E}_{0,z}\left[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{Z^\lambda(s)\in I_1, s\in[0,\tau_k), Z^\lambda(\tau_k)\notin I_1\}\mathbb{I}\{F_{1,k},\dots,F_{k,k}\}\right]\\
&\quad=\mathbf{E}\left[\tau_k\mathbb{I}\{D_k\}\prod_{j=1}^{k-1}\mathbf{E}_{0,Z^\lambda(\tau_{j-1})}\mathbb{I}\left\{z_j^\lambda(s)\in I_1, s\in[0,T_j), z_j^\lambda(T_j)+\frac{W_j}{(\lambda+\tau_j)^\theta}\in I_1, F_{j,k}\right\}\right.\\
&\qquad\left.\times\ \mathbf{E}_{0,Z^\lambda(\tau_{k-1})}\mathbb{I}\{z_k^\lambda(s)\in I_1, s\in[0,T_k), z_k^\lambda(T_k)+\frac{W_k}{(\lambda+\tau_k)^\theta}\notin I_1, F_{j,k}\}\right]\\
&\quad\le\mathbf{E}\left[\tau_k\mathbb{I}\{D_k\}\prod_{j=1}^{k-1}\sup_{y\in I_1}\mathbf{E}\mathbb{I}\{C_y^j, F_{j,k}, D_k\}\times\sup_{y\in I_1}\mathbf{E}\mathbb{I}\{\widetilde{C}_y^k, F_{k,k}, D_k\}\right],
\end{aligned}
\tag{4.36}
$$

with events C_y^j, $\widetilde{C}_y^j$, $j\ge 1$, defined in (3.26).

Analogously,

$$
\begin{aligned}
&\mathbf{E}_{0,z}[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{\sigma\in(\tau_{k-1},\tau_k)\}\mathbb{I}\{F_{1,k},\dots,F_{k,k}\}]\\
&\quad\le\mathbf{E}\tau_k\mathbb{I}\{D_k\}\left[\prod_{j=1}^{k-1}\sup_{y\in I_1}\mathbf{E}_y\mathbb{I}\{C_y^j, F_{j,k}, D_k\}\times\sup_{y\in I_1}\mathbf{E}_y\mathbb{I}\{\widehat{C}_y^k, F_{k,k}, D_k\}\right],
\end{aligned}
\tag{4.37}
$$

with

$$
\widehat{C}_y^k=\{z_{0,y}^{\lambda,\tau_{k-1}}(s)\notin I_1 \text{ for some } s\in(0,T_k)\}.
\tag{4.38}
$$

The statement of the Proposition 4.2 will follow from the forthcoming lemmas where we estimate the main and the remainder terms.

4.1.2. Main term $M(z,\lambda)$

Lemma 4.3. *Let $\alpha\theta < 1$. Then there exist $\gamma_0 > 0$ and $q_0 > \alpha\theta/2$ such that for any $0 < \gamma \le \gamma_0$, any $0 < p < \alpha\theta/2 < q \le q_0$ there is $\delta > 0$ such that for λ sufficiently large*

$$M(z,\lambda) \le \frac{\alpha}{|a|^{-\alpha} + b^{-\alpha}}\lambda^{\alpha\theta} + \lambda^{\alpha\theta-\delta} \tag{4.39}$$

uniformly for $z \in I_1$.

Proof. For the main term M, the estimate (4.36) takes the form:

$$\begin{aligned}
&\mathbf{E}_{0,z}[\tau_k \mathbb{I}\{D_k\}\mathbb{I}\{\sigma = \tau_k\}\mathbb{I}\{G_{1,k},\dots,G_{k,k}\}] \\
&\quad\le \mathbf{E}\left[\tau_k \prod_{j=1}^{k-1} \sup_{y\in I_1} \mathbf{E}\mathbb{I}\{C_y^j, G_{j,k}, D_k\} \times \sup_{y\in I_1} \mathbf{E}\mathbb{I}\{\widetilde{C}_y^k, G_{k,k}, D_k\}\right].
\end{aligned} \tag{4.40}$$

Let Lemma 3.4 hold for $0 < q \le q_1$, $0 < \gamma \le \gamma_1$ and $w > 0$ for some $\gamma_1 > 0$, $q_1 > \alpha\theta/2$ and λ sufficiently large. Then for any $j \ge 1$, $(E_y^j)^c \cap (A_j \sqcup B_j) \subseteq \mathcal{E}_j^c$ and $\mathbf{P}(\mathcal{E}_j^c) \le \exp(-\lambda^w)$. Moreover, the events $\mathcal{E}_j^c$, $j \ge 1$, are mutually independent and independent from the jump sizes $(W_j)_{j\ge1}$ and arrival times $(\tau_j)_{j\ge0}$. Recall the definition of the sets $H_{i,k}$ form (4.33).

Step MA-1. Let $k \ge 1$. With help of (3.27) from Lemma 3.3, we obtain for $1 \le j \le k-1$ and $y \in I_1$ that

$$\begin{aligned}
\mathbb{I}\{C_y^j, G_{j,k}, D_k\} &= \mathbb{I}\{C_y^j, A_j, D_k\} + \mathbb{I}\{C_y^j, B_j, D_k\}\mathbb{I}\{H_{i,k}\} \\
&= \mathbb{I}\{C_y^j, A_j, D_k\}\mathbb{I}\{H_{i,k}^c\}\mathbb{I}\{E_y^j\} + \mathbb{I}\{C_y^j, A_j, D_k\}\mathbb{I}\{H_{i,k}\} \\
&\quad + \mathbb{I}\{C_y^j, A_j, D_k\}\mathbb{I}\{H_{i,k}^c\}\mathbb{I}\{(E_y^j)^c\} + \mathbb{I}\{C_y^j, B_j, D_k\}\mathbb{I}\{H_{i,k}\} \\
&\le \mathbb{I}\{A_j, D_k\}\mathbb{I}\left\{\frac{W_j}{(\lambda+\tau_k)^\theta} \in I\right\}\mathbb{I}\{H_{i,k}^c\} + \mathbb{I}\{A_j, D_k\}\mathbb{I}\{H_{i,k}\} \\
&\quad + \mathbb{I}\{A_j, (E_y^j)^c\}\mathbb{I}\{H_{i,k}^c\} + \mathbb{I}\{B_j, D_k\}\mathbb{I}\{H_{i,k}\} \\
&\le 1 - \mathbb{I}\left\{\frac{W_j}{(\lambda+\tau_k)^\theta} \notin I\right\} + \mathbb{I}\{H_{i,k}^c\}\mathbb{I}\{\mathcal{E}_j^c\}.
\end{aligned} \tag{4.41}$$

Step MA-2. With help of (3.28) from Lemma 3.3 for $k \ge 1$ and $y \in I_1$, we estimate

$$\begin{aligned}
&\mathbb{I}\{\widetilde{C}_y^k, G_{k,k}, D_k\} \le \mathbb{I}\{\widetilde{C}_y^k\}\mathbb{I}\{A_k\}\mathbb{I}\{E_y^k\} + \mathbb{I}\{\widetilde{C}_y^k\}\mathbb{I}\{B_k\}\mathbb{I}\{E_y^k\}(=0) + \mathbb{I}\{\widetilde{C}_y^k\}\mathbb{I}\{G_{k,k}\}\mathbb{I}\{(E_y^k)^c\} \\
&\quad\le \mathbb{I}\left\{\frac{W_k}{(\lambda+\tau_k)^\theta} \notin I_2\right\} + \mathbb{I}\{A_k \sqcup B_k\}\mathbb{I}\{(E_y^k)^c\} \le \mathbb{I}\left\{\frac{W_k}{(\lambda+\tau_k)^\theta} \notin I_2\right\} + \mathbb{I}\{\mathcal{E}_k^c\}.
\end{aligned} \tag{4.42}$$

Consequently, using the independence of τ_k, W_j and $\xi^{j,\lambda}$, $1 \le j \le k$, we get

$$\begin{aligned}
&\mathbf{E}_{0,z}[\tau_k \mathbb{I}\{D_k\}\mathbb{I}\{\sigma = \tau_k\}\mathbb{I}\{G_{1,k},\dots,G_{k,k}\}] \\
&\quad\le \mathbf{E}\left[\tau_k \mathbb{I}\{D_k\} \prod_{j=1}^{k-1}\left(1 - \mathbb{I}\left\{\frac{W_j}{(\lambda+\tau_k)^\theta} \notin I\right\} + \mathbb{I}\{H_{i,k}^c\}\mathbb{I}\{\mathcal{E}_j^c\}\right)\right. \\
&\qquad \left.\times \left(\mathbb{I}\left\{\frac{W_k}{(\lambda+\tau_k)^\theta} \notin I_2\right\} + \mathbb{I}\{\mathcal{E}_k^c\}\right)\right]
\end{aligned}$$

$$= \int_{k\lambda^p}^{k\lambda^q} \mathrm{e}^{-\beta t} \frac{(\beta t)^k}{(k-1)!} \left[1 - \mathbf{P}\left(\frac{W_1}{(\lambda+t)^\theta} \notin I \right) + \mathbf{P}\left(H_{i,k}^c\right) \mathbf{P}(\mathcal{E}_1^c) \right]^{k-1} \times \left[\mathbf{P}\left(\frac{W_1}{(\lambda+t)^\theta} \notin I_2 \right) + \mathbf{P}(\mathcal{E}_1^c) \right] \mathrm{d}t. \tag{4.43}$$

Then for any $\delta_1 < \gamma$ the following estimates hold for λ being sufficiently big:

$$\mathbf{P}\left(\frac{W_1}{(\lambda+t)^\theta} \notin I \right) = \frac{A}{\beta} \frac{1}{(\lambda+t)^{\alpha\theta}} \quad \text{with } A = \frac{|a|^{-\alpha} + b^{-\alpha}}{\alpha},$$

$$\mathbf{P}\left(\frac{W_1}{(\lambda+t)^\theta} \notin I_2 \right) = \frac{|a + \lambda^{-\gamma} + \lambda^{-2\gamma}|^{-\alpha} + (b - \lambda^{-\gamma} - \lambda^{-2\gamma})^{-\alpha}}{\alpha\beta} \frac{1}{(\lambda+t)^{\alpha\theta}} \le \frac{A}{\beta} \frac{1 + \lambda^{-\delta_1}}{(\lambda+t)^{\alpha\theta}},$$

$$\begin{aligned} \mathbf{P}\left(H_{i,k}^c\right) \mathbf{P}(\mathcal{E}_1^c) &\le \frac{2\exp(-\lambda^w)}{\alpha\beta} \frac{\lambda^{2\gamma\alpha}}{(\lambda + k\lambda^p)^{\alpha\theta}} \\ &\le \frac{2\exp(-\lambda^w)}{\alpha A} \left(\frac{\lambda + k\lambda^q}{\lambda + k\lambda^p} \right)^{\alpha\theta} \frac{A\lambda^{2\gamma\alpha}}{\beta(\lambda + k\lambda^q)^{\alpha\theta}} \\ &\le \frac{A\lambda^{-\delta_1}}{\beta(\lambda+t)^{\alpha\theta}}, \quad t \in [k\lambda^p, k\lambda^q]. \end{aligned} \tag{4.44}$$

Finally, with help of Lemma A.1, we obtain the estimate for the main term:

$$\begin{aligned} M(z,\lambda) &= \sum_{k=1}^{\infty} \mathbf{E}_{0,z} \left[\tau_k \mathbb{I}\{D_k\} \mathbb{I}\{\sigma = \tau_k\} \mathbb{I}\{G_{1,k}, \dots, G_{k,k}\} \right] \\ &\le \sum_{k=1}^{\infty} \int_{k\lambda^p}^{k\lambda^q} \mathrm{e}^{-\beta t} \frac{(\beta t)^k}{(k-1)!} \left[1 - \frac{A}{\beta} \frac{1 - \lambda^{-\delta_1}}{(\lambda+t)^{\alpha\theta}} \right]^{k-1} \left[\frac{A}{\beta} \frac{1 + \lambda^{-\delta_1}}{(\lambda+t)^{\alpha\theta}} + \exp(\lambda^{-w}) \right] \mathrm{d}t \\ &\le \int_0^\infty t \exp\left[-\frac{At}{(\lambda+t)^{\alpha\theta}} (1 - \lambda^{-\delta_1}) \right] \left[A \frac{1 + \lambda^{-\delta_1}}{(\lambda+t)^{\alpha\theta}} + \beta \exp(\lambda^{-w}) \right] \mathrm{d}t \\ &\le A^{-1} \lambda^{\alpha\theta} + \lambda^{\alpha\theta - \delta}, \end{aligned} \tag{4.45}$$

for any $0 < \delta < \delta_1$ and λ sufficiently big. □

4.1.3. Remainder terms $R_1(z,\lambda)$, $R_2(z,\lambda)$ and $R_3(z,\lambda)$

Lemma 4.4. *Let $\alpha\theta < 1$. Then there exist $\gamma_0 > 0$ and $0 < p_0 < \alpha\theta/2$ such that for any $0 < \gamma \le \gamma_0$, there is $q_0 = q_0(\gamma) > \alpha\theta/2$ such that for any $0 < \gamma \le \gamma_0$, $\alpha\theta/2 < q \le q_0(\gamma)$ and any $0 < p \le p_0$, there is $\delta > 0$ such that for λ sufficiently large*

$$R_1(z,\lambda),\ R_2(z,\lambda),\ R_3(z,\lambda) \le \lambda^{\alpha\theta - \delta} \tag{4.46}$$

uniformly for $z \in I_1$.

Estimate of $R_1(z,\lambda)$. First, let γ_0 and q_0 be such that Lemma 3.4 holds for $0 < \gamma \le \gamma_0$, $\alpha\theta/2 < q \le q_0$ and some $w > 0$. Let $k \ge 1$. To estimate the expectation $\mathbf{E}_{0,z}[\tau_k \mathbb{I}\{D_k\} \mathbb{I}\{\sigma = \tau_k\} \mathbb{I}\{F_{1,k}, \dots, F_{k-1,k}, G_{k,k}\}]$, we note that $k\lambda^p \le \tau_k \mathbb{I}\{D_k\} \le k\lambda^q$. Further, we apply (4.36) and consider the following cases.

Step R1-1. Let $1 \leq j \leq k-1$, $F_{j,k} = G_{j,k}$ and $y \in I_1$. Similarly to the step **MA-1** of the previous section, we obtain

$$\begin{aligned}
\mathbb{I}\{C_y^j, G_{j,k}, D_k\} &= \mathbb{I}\{C_y^j, A_j, D_k\} + \mathbb{I}\{C_y^j, B_j, D_k\}\mathbb{I}\{H_{j,k}\} \\
&= \mathbb{I}\{C_y^j, A_j, D_k\}\mathbb{I}\{H_{j,k}^c\}\mathbb{I}\{E_y^j\} + \mathbb{I}\{C_y^j, A_j, D_k\}\mathbb{I}\{H_{j,k}\} \\
&\quad + \mathbb{I}\{C_y^j, A_j, D_k\}\mathbb{I}\{H_{j,k}^c\}\mathbb{I}\{(E_y^j)^c\} + \mathbb{I}\{C_y^j, B_j, D_k\}\mathbb{I}\{H_{j,k}\} \\
&\leq \mathbb{I}\{A_j, D_k\}\mathbb{I}\left\{\frac{W_j}{(\lambda+\tau_k)^\theta} \in I\right\}\mathbb{I}\{H_{j,k}^c\} + \mathbb{I}\{A_j, D_k\}\mathbb{I}\{H_{j,k}\} \\
&\quad + \mathbb{I}\{A_j, D_k\}\mathbb{I}\{H_{j,k}^c\}\mathbb{I}\{(E_y^j)^c\} + \mathbb{I}\{B_j, D_k\}\mathbb{I}\{H_{j,k}\} \\
&\leq \mathbb{I}\{A_j \sqcup B_j\}\left[1 - \mathbb{I}\left\{\frac{W_j}{(\lambda+k\lambda^q)^\theta} \notin I\right\} + \mathbb{I}\{H_{j,k}^c\}\mathbb{I}\{\mathcal{E}_j^c\}\right].
\end{aligned} \tag{4.47}$$

Denote $\varphi_{\lambda,k} := \lambda^{\alpha\theta/2}(\lambda+k\lambda^q)^{-\alpha\theta}$. Then, using the independence of T_j, W_j and $\mathbb{I}\{\mathcal{E}_j^c\}$, we obtain for λ sufficiently large, uniformly for all $k \geq 1$

$$\begin{aligned}
\mathbf{E}\mathbb{I}\{C_y^j, G_{j,k}, D_k\} &\leq \mathbf{P}(A_j \sqcup B_j)\left[1 - \mathbf{P}\left(\frac{W_j}{(\lambda+k\lambda^q)^\theta} \notin I\right) + \mathbf{P}(H_{j,k}^c)\mathbf{P}(\mathcal{E}_j^c)\right] \\
&\leq (1-r)\left[1 - c_1\varphi_{\lambda,k} + \frac{c_2 e^{-\lambda^w}\lambda^{2\gamma\alpha}}{\lambda^{-\alpha\theta/2}(\lambda+k\lambda^p)^{\alpha\theta}}\right] \leq (1-r)(1-c_3\varphi_{\lambda,k}).
\end{aligned} \tag{4.48}$$

Here we took into account that $\frac{\lambda+k\lambda^q}{\lambda+k\lambda^p} \leq 2\lambda^{q-p}$ for all $0 \leq p \leq q$, $k \geq 0$ and $\lambda \geq 1$.

Step R1-2. Let $1 \leq j \leq k-1$, $F_{j,k} = G_{j,k}^c$ and $y \in I_1$. Here we obtain directly from (4.34) that

$$\begin{aligned}
\mathbf{E}\mathbb{I}\{C_y^j, G_{j,k}^c, D_k\} &\leq \mathbf{P}(G_{j,k}^c) \leq r + \frac{c_r\lambda^{p+2\alpha\gamma}}{(\lambda+k\lambda^p)^{\alpha\theta}} \\
&\leq r + c_4\lambda^{-\alpha\theta/2+p+2\alpha\gamma+\alpha\theta(q-p)} \cdot \varphi_{\lambda,k}.
\end{aligned} \tag{4.49}$$

We note that the exponent in the latter summand is strictly negative for γ, p and $q - \alpha\theta/2$ small enough. Indeed,

$$\begin{aligned}
-\frac{\alpha\theta}{2} + p + 2\alpha\gamma + \alpha\theta(q-p) &= -\frac{\alpha\theta}{2}(1-\alpha\theta) + \alpha\theta\left(q - \frac{\alpha\theta}{2}\right) + (1-\alpha\theta)p + 2\alpha\gamma \\
&< -\frac{\alpha\theta}{4}(1-\alpha\theta) = -\rho < 0,
\end{aligned} \tag{4.50}$$

if we demand additionally to previous assumptions that $\gamma_0 \leq \frac{2}{15}\theta(1-\alpha\theta)$, $p_0 \leq \frac{1}{15}\alpha\theta$ and $q_0 - \frac{\alpha\theta}{2} \leq \frac{1}{15}(1-\alpha\theta)$. This yields the estimate

$$\mathbf{E}\mathbb{I}\{C_y^j, G_{j,k}^c, D_k\} \leq r + \lambda^{-\rho}\varphi_{\lambda,k} \tag{4.51}$$

for λ sufficiently big.

Step R1-3. As in **MA-2**, for $y \in I_1$ and $k \geq 1$, we have for λ sufficiently big that

$$\mathbf{E}\mathbb{I}\{\widetilde{C}_y^k, G_{k,k}, D_k\} \leq \mathbf{P}\left(\frac{W_k}{(\lambda+k\lambda^p)^\theta} \notin I_2\right) + \mathbf{P}(\mathcal{E}_k^c) \leq c_5\lambda^{-\alpha\theta/2}. \tag{4.52}$$

Now we combine Steps **R1-1**, **R1-2** and **R1-3** and make summation over all $(F_{1,k},\dots,F_{k-1,k})\neq(G_{1,k},\dots,G_{k-1,k})$ to obtain the estimate:

$$\begin{aligned}
&\sum_{(F_{1,k},\dots,F_{k-1,k})\neq(G_{1,k},\dots,G_{k-1,k})}\mathbf{E}_{0,z}\left[\tau_k\mathbb{I}\{D_k\}\mathbb{I}\{\sigma=\tau_k\}\mathbb{I}\{F_{1,k},\dots,F_{k-1,k},G_{k,k}\}\right]\\
&\leq k\lambda^q\sum_{m=0}^{k-2}\binom{k-1}{m}(1-c_3\varphi_{\lambda,k})^m(1-r)^m(r+\lambda^{-\rho}\varphi_{\lambda,k})^{k-1-m}\cdot c_5\lambda^{-\alpha\theta/2}\\
&=c_5k\lambda^{q-\alpha\theta/2}[(1-(c_3(1-r)-\lambda^{-\rho})\varphi_{\lambda,k})^{k-1}-(1-c_3\varphi_{\lambda,k})^{k-1}(1-r)^{k-1}]\\
&=c_5k\lambda^{q-\alpha\theta/2}\left(1-c_6\varphi_{\lambda,k}\right)^{k-1}\left[1-\left(1-\frac{r+\lambda^{-\rho}\varphi_{\lambda,k}}{1-c_6\varphi_{\lambda,k}}\right)^{k-1}\right]\\
&\leq c_5k\lambda^{q-\alpha\theta/2}\left(1-c_6\varphi_{\lambda,k}\right)^{k-1}[1-(1-c_7\lambda^{-\alpha\theta/2-\rho\wedge(\gamma/2)})^{k-1}]\\
&\leq c_8k(k-1)\lambda^{q-\alpha\theta-\rho\wedge(\gamma/2)}(1-c_6\varphi_{\lambda,k})^{k-1},
\end{aligned}\tag{4.53}$$

where we used an elementary inequality $1-(1-x)^{k-1}\leq(k-1)x$, $x\in[0,1]$, $k\geq 1$.

Demand now that $q_0<1-\alpha\theta/2$ to satisfy the assumptions of Lemma A.2, and $q_0-\alpha\theta/2<\rho\wedge(\gamma/2)$. With help of the estimate (A.5), we finish the proof with δ such that $0<\delta<\rho\wedge(\gamma/2)-(q-\alpha\theta/2)$. □

Estimate of $R_2(z,\lambda)$**.** Let γ_0 and p_0 and $q_0(\gamma)$ be as in the estimate of $R_1(z,\lambda)$, and $0<\gamma\leq\gamma_0$, $0<p\leq p_0$ and $\alpha\theta/2<q\leq q_0(\gamma)$. We estimate the factors in the factorisation formulae (4.36) and (4.37).

Step R2-1. For $k\geq 1$, $y\in I_1$, we use the estimates from (4.48) and (4.51) from Steps **R1-1** and **R1-2** for the factors $\mathbf{E}\mathbb{I}\{C_y^j,G_{j,k},D_k\}$ and $\mathbf{E}\mathbb{I}\{C_y^j,G_{j,k}^c,D_k\}$, $1\leq j\leq k-1$.

Step R2-2. For $k\geq 1$ and $y\in I_1$, we use again the estimate (4.51) of Step **R1-2** to obtain

$$\mathbf{E}\mathbb{I}\{C_y^k,G_{k,k}^c,D_k\}+\mathbf{E}\mathbb{I}\{\widehat{C}_y^k,G_{k,k}^c,D_k\}\leq\mathbf{P}(G_{k,k}^c)\leq\lambda^{-\alpha\theta/2-\rho\wedge(\gamma/2)},\tag{4.54}$$

with ρ defined in (4.50) and λ being sufficiently big. Then, with help of Lemma A.2, we estimate

$$\begin{aligned}
R_2(z,\lambda)&=\sum_{k=1}^{\infty}\sum_{F_{1,k},\dots,F_{k-1,k}}\mathbf{E}_{0,z}[\tau_k(\mathbb{I}\{\sigma=\tau_k\}\\
&\quad+\mathbb{I}\{\sigma\in(\tau_{k-1},\tau_k)\})\mathbb{I}\{D_k\}\mathbb{I}\{F_{1,k},\dots,F_{k-1,k},G_{k,k}^c\}]\\
&\leq\lambda^{q-\alpha\theta/2-\rho\wedge(\gamma/2)}\sum_{k=1}^{\infty}k\sum_{m=0}^{k-1}\binom{k-1}{m}(1-c_1\varphi_{\lambda,k})^m(1-r)^m\left(r+\lambda^{-\rho}\varphi_{\lambda,k}\right)^{k-1-m}\\
&\leq\lambda^{q-\alpha\theta/2-\rho\wedge(\gamma/2)}\sum_{k=1}^{\infty}k[1-(c_1(1-r)+\lambda^{-\rho})\varphi_{\lambda,k}]^{k-1}\\
&\leq\lambda^{q-\alpha\theta/2-\rho\wedge(\gamma/2)}\sum_{k=1}^{\infty}k(1-c_2\varphi_{\lambda,k})^{k-1}\leq c_3\lambda^{\alpha\theta}\lambda^{q-\alpha\theta/2-\rho\wedge(\gamma/2)}\leq\lambda^{\alpha\theta-\delta},
\end{aligned}\tag{4.55}$$

for $0<\delta<\rho\wedge(\gamma/2)-(q-\alpha\theta/2)$ and λ sufficiently big. □

Estimate of $R_3(z,\lambda)$**.** Let γ_0 and p_0 and $q_0(\gamma)$ be as in the estimate of $R_1(z,\lambda)$, and $0<\gamma\leq\gamma_0$, $0<p\leq p_0$, and $\alpha\theta/2<q\leq q_0(\gamma)$. For $k\geq 2$ and $y\in I_1$ denote $\hat{z}(y)=z_{0,y}^{\lambda,\tau_{k-2}}(T_{k-1})+\frac{W_{k-1}}{(\lambda+\tau_{k-1})^\theta}$ and consider a factorisation

$$\mathbf{E}_{0,z}[\tau_k \mathbb{I}\{D_k\}\mathbb{I}\{\sigma \in (\tau_{k-1}, \tau_k)\}\mathbb{I}\{F_{1,k}, \dots, F_{k-1,k}, G_{k,k}\}]$$
$$\leq \mathbf{E}\left[\tau_k \mathbb{I}\{D_k\}\prod_{j=1}^{k-2}\sup_{y\in I_1}\mathbf{E}\mathbb{I}\{C_y^j, F_{j,k}, D_k\} \times \sup_{y\in I_1}\mathbf{E}\mathbb{I}\{C_y^{k-1}, F_{k-1,k}, \widehat{C}_{\hat{z}(y)}, G_{k,k}, D_k\}\right]. \tag{4.56}$$

Step R3-1. For $k \geq 3$ and $y \in I_1$, we use the estimates from (4.48) and (4.51) from Steps **R1-1** and **R1-2** for the factors $\mathbf{E}\mathbb{I}\{C_y^j, G_{j,k}, D_k\}$ and $\mathbf{E}\mathbb{I}\{C_y^j, G_{j,k}^c, D_k\}$, $1 \leq j \leq k-2$.

Let us estimate the last factor in (4.56). Indeed, for $y \in I_1$ we have

$$\begin{aligned}\mathbb{I}\{C_y^{k-1}, F_{k-1,k}, \widehat{C}^k_{\hat{z}(y)}, G_{k,k}, D_k\} &\leq \mathbb{I}\{C_y^{k-1}, F_{k-1,k}, D_k\}\mathbb{I}\{\hat{z}(y) \in I_1 \setminus I_2\}\\ &\quad + \mathbb{I}\{C_y^{k-1}, F_{k-1,k}, D_k\}\mathbb{I}\{\hat{z}(y) \in I_2\}\sup_{y\in I_2}\mathbb{I}\{\widehat{C}_y^k, G_{k,k}\}\\ &\leq \mathbb{I}\{C_y^{k-1}, F_{k-1,k}, D_k\}\mathbb{I}\{\hat{z}(y) \in I_1 \setminus I_2\} + \sup_{y\in I_2}\mathbb{I}\{\widehat{C}_y^k, G_{k,k}\}.\end{aligned} \tag{4.57}$$

The last supremum in the previous formula can be estimated easily for $y \in I_2$ as $\mathbb{I}\{\widehat{C}_y^k, G_{k,k}\} \leq \mathbb{I}\{\widehat{C}_y^k, G_{k,k}, E_y^k\} + \mathbb{I}\{(E_y^k)^c\} \leq \mathbb{I}\{\mathcal{E}_k^c\}$. To estimate the first summand in the final estimate in (4.57) we consider two cases.

Step R3-2. $F_{k-1,k} = G_{k-1,k}$. For $y \in I_1$, analogously to Step **MA-2**, we obtain

$$\begin{aligned}\mathbb{I}\{C_y^{k-1}, G_{k-1,k}, D_k\}\mathbb{I}\{\hat{z}(y) \in I_1 \setminus I_2\} &\leq \mathbb{I}\{C_y^{k-1}, A_{k-1}, E_y^{k-1}\}\mathbb{I}\{\hat{z}(y) \in I_1 \setminus I_2\}\\ &\quad + \mathbb{I}\{C_y^{k-1}, B_{k-1}, E_y^{k-1}, D_k\}\mathbb{I}\{H_{k-1,k}\}\mathbb{I}\{\hat{z}(y) \in I_1 \setminus I_2\}(=0) + \mathbb{I}\{\mathcal{E}_{k-1}^c\}\\ &\leq \mathbb{I}\left\{\frac{W_{k-1}}{(\lambda+\tau_{k-1})^\theta} - a - \lambda^{-\gamma} \in [-\lambda^{-\gamma}, 2\lambda^{-2\gamma}]\right\}\\ &\quad + \mathbb{I}\left\{\frac{W_{k-1}}{(\lambda+\tau_{k-1})^\theta} - b - \lambda^{-\gamma} \in [-2\lambda^{-2\gamma}, \lambda^{-2\gamma}]\right\} + \mathbb{I}\{\mathcal{E}_{k-1}^c\}\\ &\leq \mathbb{I}\left\{\frac{W_{k-1}}{\lambda^\theta} - a - \lambda^{-\gamma} \in [-\lambda^{-\gamma}, 2\lambda^{-2\gamma}]\right\}\\ &\quad + \mathbb{I}\left\{\frac{W_{k-1}}{\lambda^\theta} - b - \lambda^{-\gamma} \in [-2\lambda^{-2\gamma}, \lambda^{-2\gamma}]\right\} + \mathbb{I}\{\mathcal{E}_{k-1}^c\}.\end{aligned} \tag{4.58}$$

Under conditions B_{k-1}, E_y^{k-1} and $H_{k-1,k}$, the event $\{\hat{z}(y) \in I_1 \setminus I_2\}$ is empty due to (3.25) because $X_y^0(t) \in [a + \lambda^{-\gamma} + \frac{3}{2}\lambda^{-2\gamma}, b - \lambda^{-\gamma} - \frac{3}{2}\lambda^{-2\gamma}]$ for all $y \in I_1$ and $t \geq c\lambda^{-\gamma}$, and $|X_y^0(t) - z_{0,y}^{\lambda,\tau_{k-2}}(t)| \leq \frac{1}{2}\lambda^{-2\gamma}$ (in case of the inner well).

Taking the expectation yields,

$$\mathbf{E}\mathbb{I}\{C_y^{k-1}, G_{k-1,k}, \widehat{C}^k_{\hat{z}(y)}, G_{k,k}\} \leq c_1\lambda^{-\alpha\theta/2-2\gamma}. \tag{4.59}$$

Step R3-3. $F_{k-1,k} = G_{k-1,k}^c$. For $y \in I_1$, we use the estimate (4.54) from **R2-2**:

$$\mathbf{E}\mathbb{I}\{C_y^{k-1}, G_{k-1,k}^c\}\mathbb{I}\{\hat{z}(y) \in I_1 \setminus I_2\} \leq \mathbf{P}(G_{k-1,k}^c) \leq \lambda^{-\alpha\theta/2-\rho\wedge(\gamma/2)}. \tag{4.60}$$

Combining the estimates (4.59) and (4.60) we obtain for $y \in I_1$ and $F_{k-1,k} = G_{k-1,k}$ or $G_{k-1,k}^c$ that

$$\mathbf{E}\mathbb{I}\{C_y^{k-1}, F_{k-1,k}, \widehat{C}^k_{\hat{z}(y)}, G_{k,k}\} \leq \lambda^{-\alpha\theta/2-\rho\wedge(\gamma/2)}. \tag{4.61}$$

Finally, as in (4.55), with help of Lemma A.2, we obtain for $z \in I_1$ that

$$\begin{aligned} R_3(\lambda, z) &= \mathbf{E}_{0,z}[\tau_1 \mathbb{I}\{D_1\}\mathbb{I}\{\sigma \in (0, \tau_1)\}\mathbb{I}\{G_{1,1}\}] \\ &\quad + \sum_{k=2}^{\infty} \sum_{F_1,\dots,F_{k-1}} \mathbf{E}_{0,z}\left[\tau_k \mathbb{I}\{D_k\}\mathbb{I}\{\sigma \in (\tau_{k-1}, \tau_k)\}\mathbb{I}\{F_{1,k}, \dots, F_{k-1,k}, G_{k,k}\}\right] \\ &\le \lambda^q + \lambda^{q-\alpha\theta/2-\rho\wedge(\gamma/2)} \sum_{k=2}^{\infty} k(1 - c_2\varphi_{\lambda,k})^{k-2} \le \lambda^{\alpha\theta-\delta}, \end{aligned} \tag{4.62}$$

with (as in Lemma 4.4) $0 < \delta < q - \alpha\theta/2 - \rho \wedge (\gamma/2)$ and for λ sufficiently big. □

4.1.4. Remainder term $R_4(\lambda)$

Lemma 4.5. *Let $\alpha\theta < 1$. Then for any p and q such that $0 < p < \frac{\alpha\theta}{2} < q$, any $0 < \delta < \alpha\theta/2$, the estimate*

$$R_4(\lambda) = \sum_{k=1}^{\infty} \mathbf{E}[\tau_k(\mathbb{I}\{\tau_k \le k\lambda^p\} + \mathbb{I}\{\tau_k \ge k\lambda^q\})] \le \lambda^{\alpha\theta-\delta} \tag{4.63}$$

holds for λ sufficiently large.

Proof. 1. Let $0 < p < \alpha\theta/2$ and $0 < \delta < \alpha\theta/2$. With help of Stirling's formula, we estimate

$$\begin{aligned} \mathbf{E}[\tau_k \mathbb{I}\{\tau_k \le k\lambda^p\}] &= \int_0^{k\lambda^p} \mathrm{e}^{-\beta t} \frac{(\beta t)^k}{(k-1)!} \mathrm{d}t \quad \left(u = \frac{t}{k\lambda^p}\right) \\ &= \int_0^1 [\beta u\lambda^p \mathrm{e}^{-\beta u\lambda^p}]^k \frac{k^k}{(k-1)!} k\lambda^p \mathrm{d}u \le \int_0^1 [\beta u\lambda^p \mathrm{e}^{-\beta u\lambda^p}]^k \frac{k!}{(k-1)!} \frac{\mathrm{e}^k \mathrm{e}^{\frac{1}{12k}}}{\sqrt{2\pi k}} k\lambda^p \mathrm{d}u \\ &\le c_1\lambda^p k^2 \int_0^1 [\beta u\lambda^p \mathrm{e}^{1-\beta u\lambda^p}]^k \mathrm{d}u \le c_1\lambda^p k[\beta\lambda^p \mathrm{e}]^k. \end{aligned} \tag{4.64}$$

Hence, for λ sufficiently big so that $\beta\lambda^p\mathrm{e} = \frac{2\mathrm{e}}{\alpha}\lambda^{-\alpha\theta/2+p} \le 1/2$, we have

$$\begin{aligned} \sum_{k=1}^{\infty} \mathbf{E}[\tau_k \mathbb{I}\{\tau_k \le k\lambda^p\}] &\le c_1\lambda^p \sum_{k=1}^{\infty} k\left[\beta\lambda^p \mathrm{e}\right]^k \\ &= c_1\lambda^p \frac{\beta\lambda^p \mathrm{e}}{(1-\beta\lambda^p\mathrm{e})^2} \le c_2\lambda^{2p-\alpha\theta/2} \le \lambda^{\alpha\theta-\delta}. \end{aligned} \tag{4.65}$$

2. Let $q > \alpha\theta/2$ and $\delta > 0$. We estimate

$$\begin{aligned} \mathbf{E}[\tau_k \mathbb{I}\{\tau_k \ge k\lambda^q\}] &= \int_{k\lambda^q}^{\infty} \mathrm{e}^{-\beta t} \frac{(\beta t)^k}{(k-1)!} \mathrm{d}t \quad \left(u = \frac{k\lambda^q}{t}\right) \\ &= \frac{1}{\beta} \frac{1}{(k-1)!} \int_0^1 \exp\left(-\frac{\beta k\lambda^q}{u}\right) \frac{(\beta k\lambda^q)^{k+1}}{u^{k+2}} \mathrm{d}u \le \frac{1}{\beta} \frac{1}{(k-1)!} \mathrm{e}^{-\beta k\lambda^q} [\beta k\lambda^q]^{k+1}, \end{aligned} \tag{4.66}$$

where the latter inequality follows from the monotonicity of the integrand on $u \in (0, 1]$ for $\beta\lambda^q > 2$, which holds for λ being big enough. Let also λ be such that $\mathrm{e}^{-\beta\lambda^q}\mathrm{e}\beta\lambda^q < 1$. Applying Stirling's formula yields

$$\sum_{k=1}^{\infty} \mathbf{E}[\tau_k \mathbb{I}\{\tau_k \ge k\lambda^q\}] \le \frac{1}{\beta} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \mathrm{e}^{-\beta k\lambda^q} (\beta k\lambda^q)^{k+1}$$

$$\leq \sum_{k=1}^{\infty} \frac{k^2\lambda^q}{\sqrt{2\pi k}} (\mathrm{e}^{-\beta\lambda^q} \mathrm{e}\beta\lambda^q)^k \mathrm{e}^{\frac{1}{12k}} \leq c_3 \lambda^{2q-\alpha\theta/2} \mathrm{e}^{-\beta\lambda^q} \leq \lambda^{\alpha\theta-\delta}. \quad \square \tag{4.67}$$

Proof of Proposition 4.2. The statement of the proposition follows directly from Lemmas 4.3–4.5 with γ_0 being the minimal value for which these lemmas hold. $\square$

4.2. Mean exit time — Estimate from below

Proposition 4.6. *Let $\alpha\theta < 1$. There exists $\gamma_0 > 0$ such that for any $0 < \gamma \leq \gamma_0$, there is $\delta > 0$ such that for all λ sufficiently large*

$$\mathbf{E}_{0,z}\sigma(\lambda) \geq \frac{\alpha}{|a|^{-\alpha} + b^{-\alpha}} \lambda^{\alpha\theta}(1 - \lambda^{-\delta}) \tag{4.68}$$

uniformly for $z \in I_2$. Moreover, for any $0 < \delta < \alpha\theta$

$$\mathbf{P}_{0,z}(\sigma(\lambda) \geq \lambda^{\alpha\theta-\delta}) \to 1, \tag{4.69}$$

uniformly for $z \in I_2$ as $\lambda \to \infty$.

Proof. 1. To obtain an estimate from below for the mean value, it is enough to consider the exit at the arrival times of the compound Poisson process η^λ, i.e. to use a simple inequality $\mathbf{E}_{0,z}\sigma(\lambda) \geq \sum_{k=1}^{\infty} \mathbf{E}_{0,z}[\tau_k \mathbb{I}\{\sigma = \tau_k\}]$.

Let $\gamma > 0$. For $z \in I_2$ and $k \geq 1$, consider the chain of inequalities analogous to (4.36), which leads to the following factorisation formula:

$$\begin{aligned}
&\mathbf{E}_{0,z}[\tau_k \mathbb{I}\{\sigma = \tau_k\}] = \mathbf{E}_{0,z}[\tau_k \mathbb{I}\{Z^\lambda(s) \in I_1, s \in [0,\tau_k), Z^\lambda(\tau_k) \notin I_1\}] \\
&\quad \geq \mathbf{E}\left[\tau_k \prod_{j=1}^{k-1} \mathbf{E}_{0,Z^\lambda(\tau_{j-1})} \mathbb{I}\left\{z_j^\lambda(s) \in I_1, s \in [0,T_j), z_j^\lambda(T_j) + \frac{W_j}{(\lambda+\tau_j)^\theta} \in I_2\right\}\right. \\
&\qquad \left. \times\ \mathbf{E}_{0,Z^\lambda(\tau_{k-1})} \mathbb{I}\left\{z_k^\lambda(s) \in I_1, s \in [0,T_k), z_k^\lambda(T_k) + \frac{W_k}{(\lambda+\tau_k)^\theta} \notin I_1\right\}\right] \\
&\quad \geq \mathbf{E}\tau_k \left[\prod_{j=1}^{k-1} \inf_{y \in I_2} \mathbf{E}_{0,y} \mathbb{I}\left\{z_j^\lambda(s) \in I_1, s \in [0,T_j), z_j^\lambda(T_j) + \frac{W_j}{(\lambda+\tau_j)^\theta} \in I_2\right\}\right. \\
&\qquad \left. \times\ \inf_{y \in I_2} \mathbf{E}_{0,y} \mathbb{I}\left\{z_k^\lambda(s) \in I_1, s \in [0,T_k), z_k^\lambda(T_k) + \frac{W_k}{(\lambda+\tau_k)^\theta} \notin I_1\right\}\right] \\
&\quad = \mathbf{E}\left[\tau_k \prod_{j=1}^{k-1} \inf_{y \in I_2} \mathbf{E}\mathbb{I}\{C_y^{j,-}\} \times \inf_{y \in I_2} \mathbf{E}\mathbb{I}\{\widetilde{C}_y^k\}\right]
\end{aligned} \tag{4.70}$$

with events $C_y^{k,-}$ and $\widetilde{C}_y^k$, $k \geq 1$, defined in Eq. (3.26).

Let Lemma 3.1 hold for $0 < \gamma \leq \gamma_0$ and $\alpha\theta/2 < q \leq q_0$ for sufficiently big λ. Fix γ and q satisfying the above inequalities and let $0 < p < \alpha\theta/2$. Then there is $w > 0$ such that $\mathbf{P}(\mathcal{E}_1^c) \leq \exp(-\lambda^w)$ and $\mathbf{P}(T_1 > \lambda^q) \leq \exp(-\lambda^w)$ for λ big enough. Let us estimate the factors in the last expression in (4.70).

Step MB-1. Let $k \geq 2$, $1 \leq j \leq k-1$ and $y \in I_2$. Let the estimate (3.25) hold with a constant $c > 0$ for λ being big enough. Denote $H_j := H_{j,0} = \{\frac{|W_j|}{\lambda^\theta} \leq \frac{1}{\lambda^{2\gamma}}\}$. Then

$$\mathbb{I}\{C_y^{j,-}\} \geq \mathbb{I}\{C_y^{j,-}, E_y^j\}$$

$$
\begin{aligned}
&\geq \mathbb{I}\{C_y^{j,-}, E_y^j, H_j\}\mathbb{I}\left\{\frac{c}{\lambda^\gamma} \leq T_j \leq \lambda^q\right\} + \mathbb{I}\{C_y^{j,-}, E_y^j, H_j^c\}\mathbb{I}\{\lambda^p \leq T_j \leq \lambda^q\}\\
&\geq \mathbb{I}\{E_y^j, H_j\}\mathbb{I}\left\{\frac{c}{\lambda^\gamma} \leq T_j \leq \lambda^q\right\} + \mathbb{I}\left\{\frac{W_j}{(\lambda+\tau_j)^\theta} \in I_2\right\}\mathbb{I}\{E_y^j, H_j^c\}\mathbb{I}\{\lambda^p \leq T_j \leq \lambda^q\}\\
&\geq \mathbb{I}\{H_j\}\mathbb{I}\left\{T_j \geq \frac{c}{\lambda^\gamma}\right\} + \mathbb{I}\left\{\frac{W_j}{\lambda^\theta} \in I_2\right\}\mathbb{I}\{H_j^c\}\mathbb{I}\{T_j \leq \lambda^q\} - 2\mathbb{I}\{(E_y^j)^c\}\mathbb{I}\{T_j \leq \lambda^q\}\\
&\geq \mathbb{I}\{H_j\} + \mathbb{I}\left\{\frac{W_j}{\lambda^\theta} \in I_2\right\}\mathbb{I}\{H_j^c\} - \mathbb{I}\left\{T_j < \frac{c}{\lambda^\gamma}\right\} - \mathbb{I}\{T_j > \lambda^q\} - 2\mathbb{I}\{\mathcal{E}_j^c\}\\
&= 1 - \mathbb{I}\left\{\frac{W_j}{\lambda^\theta} \notin I_2\right\} - \mathbb{I}\left\{T_j < \frac{c}{\lambda^\gamma}\right\} - \mathbb{I}\{T_j > \lambda^q\} - 2\mathbb{I}\{\mathcal{E}_j^c\}. \qquad (4.71)
\end{aligned}
$$

Hence, taking the expectation yields

$$
\begin{aligned}
\mathbf{E}\mathbb{I}\{C_y^{j,-}\} &\geq 1 - \frac{|a+\lambda^{-\gamma}+\lambda^{-2\gamma}|^{-\alpha} + (b-\lambda^{-\gamma}-\lambda^{-2\gamma})^{-\alpha}}{\alpha\beta}\frac{1}{\lambda^{\alpha\theta}}\\
&\quad - \frac{c_1}{\lambda^{\alpha\theta/2+\gamma}} - 3\exp(-\lambda^w)\\
&\geq 1 - \frac{A}{\beta}\frac{1+c_2\lambda^{-\gamma}}{\lambda^{\alpha\theta}}, \qquad A = \frac{|a|^{-\alpha}+b^{-\alpha}}{\alpha} > 0, \qquad (4.72)
\end{aligned}
$$

for λ sufficiently big.

Step MB-2. Let $k \geq 1$ and $y \in I_2$. Then we estimate

$$
\begin{aligned}
\mathbb{I}\{\widetilde{C}_y^k\} &\geq \mathbb{I}\{\widetilde{C}_y^k\}\mathbb{I}\{E_y^k\}\mathbb{I}\{\lambda^p \leq T_k \leq \lambda^q\}\\
&\geq \mathbb{I}\left\{\frac{W_k}{(\lambda+\tau_k)^\theta} \notin I_1\right\} - \mathbb{I}\left\{\frac{W_k}{\lambda^\theta} \notin I_1\right\}\mathbb{I}\{T_k \leq \lambda^p\} - \mathbb{I}\{T_k \geq \lambda^q\} - \mathbb{I}\{\mathcal{E}_k^c\},
\end{aligned}
$$

$$
\mathbf{E}[\mathbb{I}\{C_y^k\}|\tau_k = t] \geq \frac{A}{\beta}\frac{1-c_3\lambda^{-\gamma}}{(\lambda+t)^{\alpha\theta}} - c_4\lambda^{-\alpha\theta+p}, \qquad (4.73)
$$

for λ sufficiently big.

Finally, with help of equality (A.3) in Lemma A.1, we obtain the statement of the Proposition with $\delta < \gamma \wedge \delta_1$, λ being big and δ_1 being a constant from Lemma A.1:

$$
\begin{aligned}
\mathbf{E}_{0,z}\sigma(\lambda) &\geq \sum_{k=1}^\infty \int_0^\infty e^{-\beta t}\frac{(\beta t)^k}{(k-1)!}\left[1 - \frac{A}{\beta}\frac{1+c_2\lambda^{-\gamma}}{\lambda^{\alpha\theta}}\right]^{k-1}\left[\frac{A}{\beta}\frac{1-c_3\lambda^{-\gamma}}{(\lambda+t)^{\alpha\theta}} - c_4\lambda^{-\alpha\theta+p}\right]\mathrm{d}t\\
&= \int_0^\infty t\exp\left[-A\frac{t(1+c_2\lambda^{-\gamma})}{\lambda^{\alpha\theta}}\right]\left[A\frac{1-c_3\lambda^{-\gamma}}{(\lambda+t)^{\alpha\theta}} - c_4\beta\lambda^{-\alpha\theta+p}\right]\mathrm{d}t\\
&\geq \frac{\lambda^{\alpha\theta}}{A}(1-c_5\lambda^{-\gamma}) \geq \frac{\lambda^{\alpha\theta}}{A}(1-\lambda^{-\delta}). \qquad (4.74)
\end{aligned}
$$

2. Analogously, we estimate the probability (4.69) for any $0 < \delta < \alpha\theta$. Indeed, doing some algebra yields

$$
\begin{aligned}
\mathbf{P}_{0,z}(\sigma(\lambda) \geq \lambda^{\alpha\theta-\delta}) &\geq \sum_{k=1}^\infty \mathbf{E}_{0,z}\mathbb{I}\{\sigma \geq \lambda^{\alpha\theta-\delta}\}\mathbb{I}\{\sigma = \tau_k\}\\
&\geq \sum_{k=1}^\infty \int_{\lambda^{\alpha\theta-\delta}}^\infty \beta e^{-\beta t}\frac{(\beta t)^{k-1}}{(k-1)!}\left[1 - \frac{A}{\beta}\frac{1+c_2\lambda^{-\gamma}}{\lambda^{\alpha\theta}}\right]^{k-1}\left[\frac{A}{\beta}\frac{1-c_3\lambda^{-\gamma}}{(\lambda+t)^{\alpha\theta}} - c_4\lambda^{-\alpha\theta+p}\right]\mathrm{d}t
\end{aligned}
$$

$$= \int_{\lambda^{\alpha\theta-\delta}}^{\infty} \exp\left[-A\frac{t(1+c_2\lambda^{-\gamma})}{\lambda^{\alpha\theta}}\right]\left[A\frac{1-c_3\lambda^{-\gamma}}{(\lambda+t)^{\alpha\theta}} - c_4\beta\lambda^{-\alpha\theta+p}\right]\mathrm{d}t \geq 1-\lambda^{-\delta_1}, \quad (4.75)$$

for big λ and $0 < \delta_1 < \gamma \wedge (\alpha\theta/2 - p) \wedge \delta$ subject to the condition $0 < \gamma < 1-\alpha\theta$. □

4.3. Transition probability

Let $a' < b'$ be such that $I \cap [a', b'] = \emptyset$. For $\gamma > 0$, consider the interval $I_1' = [a'+\lambda^{-\gamma}, b'-\lambda^{-\gamma}]$.

Proposition 4.7. *Let* $\alpha\theta < 1$. *There exists* $\gamma_0 > 0$ *such that for any* $0 < \gamma \leq \gamma_0$, *there is* $\delta > 0$ *such that for* λ *sufficiently large*

$$\mathbf{P}_{0,z}(Z^\lambda(\sigma(\lambda)) \in I_1') \geq \frac{||a'|^{-\alpha} - |b'|^{-\alpha}|}{a^{-\alpha}+b^{-\alpha}}(1-\lambda^{-\delta}) \quad (4.76)$$

uniformly for $z \in I_2$.

Proof. Assume for definiteness that $b < a' < b' \leq +\infty$. As in the previous section, we use the following simple estimate:

$$\mathbf{P}_{0,z}(Z^\lambda(\sigma(\lambda)) \in I_1') \geq \sum_{k=1}^{\infty} \mathbf{E}_{0,z}\mathbb{I}\{Z^\lambda(\tau_k) \in I_1'\}\mathbb{I}\{\sigma = \tau_k\}. \quad (4.77)$$

Let $\gamma > 0$. For $z \in I_2$, similarly to the estimate (4.70) we have

$$\mathbf{E}_{0,z}\mathbb{I}\{Z^\lambda(\tau_k) \in I_1'\}\mathbb{I}\{\sigma = \tau_k\} \geq \prod_{j=1}^{k-1} \inf_{y\in I_2} \mathbf{E}\mathbb{I}\{C_y^{j,-}\} \times \inf_{y \in I_2} \mathbf{E}\mathbb{I}\{F_y^k\}, \quad (4.78)$$

with $C_y^{j,-}$ defined in (3.26) and

$$F_y^k = \left\{z_{0,y}^{\lambda,\tau_{k-1}}(s) \in I_1, s \in [0, T_k), z_{0,z}^{\lambda,\tau_{k-1}}(T_k) + \frac{W_k}{(\lambda+\tau_k)^\theta} \in I_1'\right\}. \quad (4.79)$$

Choose $0 < \gamma \leq \gamma_0$, $0 < p < \alpha\theta/2 < q \leq q_0$ as in Proposition 4.6. Then the estimate (4.72) holds with some constant c_1 for λ sufficiently big. Let $k \geq 1$ and $y \in I_2$. As in (4.73), we estimate

$$\begin{aligned} \mathbb{I}\{F_y^k\} &\geq \mathbb{I}\{F_y^k\}\mathbb{I}\{E_y^k\}\mathbb{I}\{\lambda^p \leq T_k \leq \lambda^q\} \\ &\geq \mathbb{I}\left\{\frac{W_k}{(\lambda+\tau_k)^\theta} \in I_1'\right\} - \mathbb{I}\left\{\frac{W_k}{\lambda^\theta} > c\right\}\mathbb{I}\{T_k \leq \lambda^p\} - \mathbb{I}\{T_k > \lambda^q\} - \mathbb{I}\{\mathcal{E}_k^c\}, \end{aligned}$$

$$\mathbf{E}[\mathbb{I}\{F_y^k\}|\tau_k = t] \geq \frac{A'}{\beta}\frac{1-c_2\lambda^{-\gamma}}{(\lambda+t)^{\alpha\theta}} - c_3\lambda^{-\alpha\theta+p}, \qquad A' = \frac{||b'|^{-\alpha} - |a'|^{-\alpha}|}{\alpha}, \quad (4.80)$$

for λ sufficiently big. Combining these estimates we obtain

$$\begin{aligned} &\mathbf{P}_{0,z}(Z^\lambda(\sigma(\lambda)) \in I_1') \\ &\geq \sum_{k=1}^{\infty}\int_0^\infty \beta e^{-\beta t}\frac{(\beta t)^{k-1}}{(k-1)!}\left[1 - \frac{A}{\beta}\frac{1+c_1\lambda^{-\gamma}}{\lambda^{\alpha\theta}}\right]^{k-1}\left[\frac{A'}{\beta}\frac{1-c_2\lambda^{-\gamma}}{(\lambda+t)^{\alpha\theta}} - c_3\lambda^{-\alpha\theta+p}\right]\mathrm{d}t \\ &= \int_0^\infty \exp\left[-A\frac{t(1+c_1\lambda^{-\gamma})}{(\lambda+t)^{\alpha\theta}}\right]\left[A'\frac{1-c_2\lambda^{-\delta}}{(\lambda+t)^{\alpha\theta}} - c_3\beta\lambda^{-\alpha\theta+p}\right]\mathrm{d}t \geq \frac{A'}{A}(1-\lambda^{-\delta}) \end{aligned} \quad (4.81)$$

for $0 < \delta < \gamma$ and $\lambda \to \infty$. $\square$

Proof of Proposition 4.1. The first statement of Proposition 4.1 follows immediately from Propositions 4.2 and 4.6. The estimate from below for the transition probability from the potential well Ω_λ^i to the well Ω_λ^j, $i \neq j$, follows from Proposition 4.7. Further, it follows from the conditions $\sum_{j\neq i} \mathbf{P}_{0,z}(Z^\lambda(\sigma^i(\lambda)) \in \Omega_\lambda^j) \leq 1$ and $\sum_{j\neq i} \frac{q_{ij}}{q_i} = 1$ that $\mathbf{P}_{0,z}(Z^\lambda(\sigma^i(\lambda)) \in \Omega_\lambda^j) \leq \frac{q_{ij}}{q_i}\left(1 + \frac{q_i}{\min_{j\neq i} q_{ij}}\lambda^{-\delta}\right)$. Hence, the estimate from above holds for any $0 < \delta' < \delta$ and λ sufficiently big. Finally, we note that all estimates are uniform over $1 \leq i \leq n$. $\square$

5. Transition times

Let $\Delta_0 = \frac{1}{2}\min_{1\leq i\leq n}(|m_i - s_{i-1}| \wedge |m_i - s_i|)$. For $0 < \Delta \leq \Delta_0$, consider Δ-neighbourhoods of the potential's local minima $B_i = \{y : |y - m_i| \leq \Delta\}$, $i = 1, \ldots, n$ and the stopping times $\tau_{s,z}^{i,\lambda} = \inf\{u \geq s : Z_{s,z}^\lambda(u) \in \cup_{j\neq i} B_j\}$. If $z \in B_i$, then $\tau_{s,z}^{i,\lambda}$ denotes the *transition time* from the Δ-neighbourhood of m_i to a Δ-neighbourhood of some other potential's minimum. For all $j \neq i$, we also consider the corresponding *transition probabilities* $\mathbf{P}_{s,z}(Z^\lambda(\tau^{i,\lambda}) \in B_j)$. The goal of this section is the following theorem.

Theorem 5.1. *For any $0 < \Delta \leq \Delta_0$, there is $\delta > 0$ such that*

$$\begin{aligned} &\mathbf{E}_{0,z}\tau^{i,\lambda} = q_i^{-1}\lambda^{\alpha\theta}(1 + \mathcal{O}(\lambda^{-\delta})) \quad \textit{and} \\ &\mathbf{P}_{0,z}(Z^\lambda(\tau^{i,\lambda}) \in B_j) = q_{ij}q_i^{-1}(1 + \mathcal{O}(\lambda^{-\delta})), \end{aligned} \tag{5.82}$$

as $\lambda \to +\infty$ uniformly for $z \in B_i$, $i, j = 1, \ldots, n$, $i \neq j$.

Remark 5.2. Under the conditions of Theorem 5.1, we easily obtain from (2.8) that uniformly for $s \geq 0$

$$\begin{aligned} &\mathbf{E}_{s,z}\tau^{i,\lambda} = s + \mathbf{E}_{0,z}\tau^{i,\lambda+s} = s + q_i^{-1}(s+\lambda)^{\alpha\theta}\left(1 + \mathcal{O}(\lambda^{-\delta})\right), \\ &\mathbf{P}_{s,z}(Z^\lambda(\tau^{i,\lambda}) \in B_j) = q_{ij}q_i^{-1}\left(1 + \mathcal{O}(\lambda^{-\delta})\right), \quad \lambda \to +\infty. \end{aligned} \tag{5.83}$$

To prove Theorem 5.1, we consider some auxiliary stopping times. Let $\gamma > 0$. For $i = 1, \ldots, n-1$, denote $U_\lambda^i = \{y : |y - s_i| \leq 2\lambda^{-\gamma}\}$ the small neighbourhood of the saddle point s_i, and $U_\lambda = \cup_{i=1}^{n-1} U_\lambda^i$. Consider the following stopping times

$$\begin{aligned} &T_{s,z}^{i,\lambda} = \inf\left\{u \geq s : Z_{s,z}^\lambda(u) \in \cup_{j\neq i}\Omega_\lambda^j\right\}, \quad i = 1, \ldots, n, \\ &S_{s,z}^{i,\lambda} = \inf\{u \geq s : Z_{s,z}^\lambda(u) \notin U_\lambda^i\}, \quad i = 1, \ldots, n-1, \\ &S_{s,z}^\lambda = \inf\{u \geq s : Z_{s,z}^\lambda(u) \notin U_\lambda\}. \end{aligned} \tag{5.84}$$

Lemma 5.3. *Let $\alpha\theta < 1$. There is $\gamma_0 > 0$ such that for any $0 < \gamma \leq \gamma_0$, there is $\delta > 0$ such that the estimate*

$$\mathbf{E}_{s,z}S^\lambda \leq (s + \lambda^{-\delta}(s+\lambda)^{\alpha\theta})(1 + \mathcal{O}(\lambda^{-\delta})) \tag{5.85}$$

holds uniformly for $z \in U_\lambda$ and $s \geq 0$ for λ sufficiently large.

Proof. 1. Let $0 < \gamma \leq \gamma_0 < \theta$ and $0 < \delta < \alpha\gamma$. First, we estimate the mean exit time $\mathbf{E}_{s,z} S^{i,\lambda}$ for $1 \leq i \leq n$. We note that for $z \in U_\lambda^i$,

$$S_{s,z}^{\lambda,i} - s \leq \inf \left\{ u \geq 0 : \frac{|\Delta L(u)|}{(\lambda + s + u)^\theta} > \frac{4}{\lambda^\gamma} \right\} \text{a.s.,} \tag{5.86}$$

i.e. the process leaves the neighbourhood U_λ^i if the absolute value of a jump exceeds the size of the neighbourhood. Denote for brevity $\mu := \lambda + s$, $\psi_{\mu,\lambda}(k) := (\mu + k\lambda^{\alpha\theta-\delta})^{-\alpha\theta}$, and $J_\lambda^k := ((k-1)\lambda^{\alpha\theta-\delta}, k\lambda^{\alpha\theta-\delta}]$. For $k \geq 1$, introduce events

$$\begin{aligned} A_k &:= \left\{ \sup_{u \in J_\lambda^k} \frac{|\Delta L(u)|}{(\mu + u)^\theta} \leq \frac{4}{\lambda^\gamma} \right\} \subseteq \left\{ \sup_{u \in J_\lambda^k} \psi_{\mu,\lambda}(k)^{1/\alpha} |\Delta L(u)| \leq \frac{4}{\lambda^\gamma} \right\} =: B_k, \\ A_k^c &:= \left\{ \sup_{u \in J_\lambda^k} \frac{|\Delta L(u)|}{(\mu + u)^\theta} > \frac{4}{\lambda^\gamma} \right\} \subseteq \left\{ \sup_{u \in J_\lambda^k} \psi_{\mu,\lambda}(k-1)^{1/\alpha} |\Delta L(u)| > \frac{4}{\lambda^\gamma} \right\} =: \hat{B}_k. \end{aligned} \tag{5.87}$$

Hence we can estimate

$$\begin{aligned} \max_{1 \leq i \leq n-1} \sup_{z \in U_\lambda^i} \mathbf{E}_{s,z} S^{i,\lambda} - s &\leq \sum_{k=1}^\infty k\lambda^{\alpha\theta-\delta} \mathbf{P}(A_1 \cdots A_{k-1} A_k^c) \\ &\leq \lambda^{\alpha\theta-\delta} \sum_{k=1}^\infty k \mathbf{P}(B_1) \cdots \mathbf{P}(B_{k-1}) \mathbf{P}(\hat{B}_k), \end{aligned} \tag{5.88}$$

where the factorisation in the latter formula is due to the independence of increments of L. Further, recalling that the number of jumps on the interval $(0, t]$, $t > 0$, with the absolute value bigger than a positive a is a Poisson random variable with mean $t \int_{|y|>a} \nu(\mathrm{d}y)$, we calculate the probabilities

$$\begin{aligned} \mathbf{P}(B_k) &= \exp\left(-2\lambda^{\alpha\theta-\delta} \int_{\frac{4\psi_{\mu,\lambda}(k)^{1/\alpha}}{\lambda^\gamma}}^\infty \frac{\mathrm{d}y}{y^{1+\alpha}} \right) = \exp\left(-c_1 \psi_{\mu,\lambda}(k) \lambda^{\alpha\gamma+\alpha\theta-\delta} \right), \\ \mathbf{P}(\hat{B}_k) &= 1 - \exp\left(-2\lambda^{\alpha\theta-\delta} \int_{\frac{4\psi_{\mu,\lambda}(k)^{1/\alpha}}{\lambda^\gamma}}^\infty \frac{\mathrm{d}y}{y^{1+\alpha}} \right) = 1 - \exp\left(-c_1 \psi_{\mu,\lambda}(k-1) \lambda^{\alpha\gamma+\alpha\theta-\delta} \right) \end{aligned} \tag{5.89}$$

with $c_1 = \frac{2}{4^\alpha \alpha}$. Thus, with help of Lemma A.3, we obtain the uniform estimate

$$\max_{1 \leq i \leq n-1} \sup_{z \in U_\lambda^i} \mathbf{E}_{s,z} S^{i,\lambda} \leq s + c_2 (\lambda + s)^{\alpha\theta} \lambda^{-\alpha\gamma} \tag{5.90}$$

for λ sufficiently large.

2. Leaving U_λ^i, the process Z^λ either enters $\cup_{i=1}^n \Omega_\lambda^i$, or jumps to some U_λ^j, $j \neq i$. Consider the probability $\mathbf{P}_{s,z}(Z^\lambda(S^{i,\lambda}) \in \cup_{j \neq i} U_\lambda^j)$. If $a_{ij} = s_j - s_i$, $1 \leq i, j \leq n-1$, $j \neq i$, then the jump

of Z^λ at the exit time must be of the order a_{ij}. Let $a := \frac{1}{2}\min_{1\le i<j\le n-1} a_{ij}$, and define events

$$\begin{aligned}\tilde{A}_k^{ij} &:= \begin{cases}\left\{\exists u\in J_\lambda^k : \dfrac{\Delta L(u)}{(\mu+u)^\theta}\in[a_{ij}-\frac{2}{\lambda^\gamma}, a_{ij}+\frac{2}{\lambda^\gamma}]\right\}, & a_{ij}>0\\ \left\{\exists u\in J_\lambda^k : \dfrac{\Delta L(u)}{(\mu+u)^\theta}\in[a_{ij}+\frac{2}{\lambda^\gamma}, a_{ij}-\frac{2}{\lambda^\gamma}]\right\}, & a_{ij}<0\end{cases}\\ &\subseteq \left\{\sup_{u\in((k-1)\lambda^{\alpha\theta-\delta},k\lambda^{\alpha\theta-\delta}]} : |\Delta L(u)|\ge a\psi_{\mu,\lambda}(k-1)^{1/\alpha}\right\} =: \tilde{B}_k.\end{aligned}\tag{5.91}$$

Analogously to (5.89), we get for $c_3=\frac{2}{a^\alpha\alpha}$, that $\mathbf{P}(\tilde{B}_k)=1-\exp(-c_3\lambda^{\alpha\theta-\delta}\psi_{\mu,\lambda}(k-1))$. Then with help of Lemma A.3, we obtain for all $s\ge 0$ and λ sufficiently big that

$$\begin{aligned}\max_{1\le i\le n-1}\sup_{z\in U_\lambda^i}\mathbf{P}_{s,z}\left(Z^\lambda(S^{i,\lambda})\in\cup_{j\ne i}U_\lambda^j\right) &\le \max_{1\le i\le n-1}\sum_{k=1}^\infty\sum_{j\ne i}\mathbf{P}(A_1,\dots,A_{k-1},\tilde{A}_k^{ij})\\ &\le (n-2)\sum_{k=1}^\infty \mathbf{P}(B_1,\dots,B_{k-1},\tilde{B}_k)\le c_4\lambda^{-\alpha\theta+\delta-\alpha\gamma}.\end{aligned}\tag{5.92}$$

3. Before exiting U_λ, the process Z^λ may jump several times between different neighbourhoods U_λ^j. Denote these times $S^\lambda(0)<S^\lambda(1)<\dots<S^\lambda(j)<\cdots$. If $z\in U_\lambda^i$, then it is clear that $\mathbf{E}_{s,z}S^\lambda(0)\le\mathbf{E}_{s,z}S^{i,\lambda}$. For a nonnegative random variable σ, the elementary inequality $(x+y)^a\le x^a+y^a\le x+1+y^a$, $x,y\ge 0$, $0\le a\le 1$, yields that $\mathbf{E}(\sigma+\lambda)^a\le\mathbf{E}\sigma+1+\lambda^a$. For $j\ge 1$, we apply the strong Markov property to obtain with help of (5.90) that

$$\begin{aligned}\mathbf{E}_{s,z}S^\lambda(j) &\le \mathbf{E}_{s,z}\sup_{z\in U_\lambda}\mathbf{E}_{S^\lambda(j-1),z}S^\lambda(0)\le\mathbf{E}_{s,z}\left[S^\lambda(j-1)+c_2\lambda^{-\alpha\gamma}(S^\lambda(j-1)+\lambda)^{\alpha\theta}\right]\\ &\le (1+c_2\lambda^{-\alpha\gamma})\mathbf{E}_{s,z}S^\lambda(j-1)+c_2\lambda^{-\alpha\gamma}(1+\lambda^{\alpha\theta}),\end{aligned}\tag{5.93}$$

and thus for $j\ge 1$, we obtain for $z\in U_\lambda$ that

$$\mathbf{E}_{s,z}S^\lambda(j)\le(1+c_2\lambda^{-\alpha\theta})^j\left(\max_{1\le i\le n}\sup_{z\in U_\lambda^i}\mathbf{E}_{s,z}S^{i,\lambda}+c_2\lambda^{-\alpha\gamma}(1+\lambda^{\alpha\theta})\right).\tag{5.94}$$

The probability that Z^λ jumps j times between neighbourhoods U_λ^i before it leaves U_λ is estimated by $(c_4\lambda^{-\alpha\theta+\delta-\alpha\gamma})^j$. Finally, for λ sufficiently big such that $c_4\lambda^{-\alpha\theta+\delta-\alpha\gamma}(1+c_2\lambda^{-\alpha\theta})\le 1/2$ we get in the limit of large λ that

$$\begin{aligned}\mathbf{E}_{s,z}S^\lambda &\le \max_{1\le i\le n}\sup_{z\in U_\lambda^i}\mathbf{E}_{s,z}S^{i,\lambda}+\left(\max_{1\le i\le n}\sup_{z\in U_\lambda^i}\mathbf{E}_{s,z}S^{i,\lambda}+c_2\lambda^{-\alpha\gamma}(1+\lambda^{\alpha\theta})\right)\\ &\quad\times\sum_{j=1}^\infty[c_4\lambda^{-\alpha\theta+\delta-\alpha\gamma}(1+c_2\lambda^{-\alpha\theta})]^j\\ &\le s+c_2\lambda^{-\alpha\gamma}(s+\lambda)^{\alpha\theta}+c_5\lambda^{-\alpha\theta+\delta-\alpha\gamma}(s+c_2\lambda^{-\alpha\gamma}(s+\lambda)^{\alpha\theta}+c_2\lambda^{-\alpha\gamma}(1+\lambda^{\alpha\theta}))\\ &\le (s+\lambda^{-\delta}(s+\lambda)^{\alpha\theta})(1+\mathcal{O}(\lambda^{-\delta})).\quad\square\end{aligned}\tag{5.95}$$

Lemma 5.4. *Let $\alpha\theta < 1$. There is $\gamma_0 > 0$ such that for any $0 < \gamma \le \gamma_0$, there is $\delta > 0$ such that*

$$\begin{aligned}
&\mathbf{E}_{s,z}T^{i,\lambda} = \left(s + q_i^{-1}(\lambda+s)^{\alpha\theta}\right)(1+\mathcal{O}(\lambda^{-\delta})),\\
&\mathbf{P}_{s,z}(Z^\lambda(T^{i,\lambda}) \in \Omega_\lambda^j) = q_{ij}q_i^{-1}(1+\mathcal{O}(\lambda^{-\delta})), \quad \lambda \to +\infty,
\end{aligned} \tag{5.96}$$

holds uniformly for $z \in \Omega^i$, $s \ge 0$ and $1 \le i, j \le n$, $i \ne j$.

Proof. 1. Let γ_0 satisfy conditions of Proposition 4.1 and Lemma 5.3, $s \ge 0$ and $z \in \Omega_\lambda^i$. Obviously, $T_{s,z}^{i,\lambda} \ge \sigma_{s,z}^{i,\lambda}$ a.s. Then for λ big enough and some $\delta_1 > 0$ such that Proposition 4.1 holds, we have

$$\mathbf{P}_{s,z}(T^{i,\lambda} = \sigma^{i,\lambda}) = \sum_{j\ne i}\mathbf{P}_{s,z}(Z^\lambda(\sigma^{i,\lambda}) \in \Omega_\lambda^j) \ge 1 - c_1\lambda^{-\delta_1}. \tag{5.97}$$

Consequently,

$$\begin{aligned}
\mathbf{P}_{s,z}(Z^\lambda(T^{i,\lambda}) \in \Omega_\lambda^j) &= \mathbf{P}_{s,z}(Z^\lambda(\sigma^{i,\lambda}) \in \Omega_\lambda^j, T^{i,\lambda} = \sigma^{i,\lambda})\\
&\quad + \mathbf{P}_{s,z}(Z^\lambda(T^{i,\lambda}) \in \Omega_\lambda^j, T^{i,\lambda} > \sigma^{i,\lambda})\\
&\ge \mathbf{P}_{s,z}(Z^\lambda(\sigma^\lambda) \in \Omega_\lambda^j) - \mathbf{P}(T^{i,\lambda} > \sigma^{i,\lambda}) \ge q_{ij}q_i^{-1} - c_1\lambda^{-\delta_1}.
\end{aligned} \tag{5.98}$$

The converse inequality follows easily from the identity $q_i^{-1}\sum_{j\ne i} q_{ij} = 1$ (see also the proof of Proposition 4.1).

2. Before entering $\cup_{j\ne i}\Omega_\lambda^j$, the process Z^λ may repeatedly visit U_λ and Ω_λ^i. There are two possibilities, namely

$$\begin{aligned}
&\Omega_\lambda^i \to U_\lambda \to \cdots \to \Omega_\lambda^i \to U_\lambda \to \quad \Omega_\lambda^i \to \cup_{j\ne i}\Omega_\lambda^j,\\
&\Omega_\lambda^i \to U_\lambda \to \cdots \to \Omega_\lambda^i \to U_\lambda \to \Omega_\lambda^i \to \cup_{j\ne i}\Omega_\lambda^j.
\end{aligned} \tag{5.99}$$

We estimate the length of the second longer transition. Denote $\sigma(j)$ the jth hitting time of U_λ in this cycle, and $\rho(j)$ the first hitting time of Ω_λ^i after $\sigma(j)$, where by convention $\rho(0) = s$, and $T(j)$ be the transition time from Ω_λ^i to $\cup_{j\ne i}\Omega_\lambda^j$ on the event $\{Z^\lambda$ visits U_λ j times$\}$.

Let Proposition 4.1 and Lemma 5.3 hold for some δ_1 and λ sufficiently large. Due to Proposition 4.1, the probability to jump into a set U_λ from Ω_λ^i is bounded by $c_1\lambda^{-\delta_1}$, and the strong Markov property implies that the probability of making the cycle $\Omega_\lambda^i \leftrightarrows U_\lambda$ exactly j times is bounded by $(c_1\lambda^{-\delta_1})^j$. There is a constant $q > 0$ such that for all $s \ge 0$ and λ sufficiently large $\sup_{z\in\Omega_\lambda^i}\mathbf{E}_{s,z}\sigma^{i,\lambda} \le s + q(s+\lambda)^{\alpha\theta}$. Due to Lemma 5.3, we have the estimate $\sup_{z\in U_\lambda}\mathbf{E}_{s,z}S^\lambda \le 2(s + (s+\lambda)^{\alpha\theta})$. Then the strong Markov property yields

$$\begin{aligned}
\mathbf{E}_{s,z}T(j) &\le \mathbf{E}_{s,z}\sup_{z'\in\Omega_\lambda^i}\mathbf{E}_{\rho(j),z'}\sigma^{i,\lambda}\\
&\le \mathbf{E}_{s,z}[\rho(j) + q(\rho(j)+\lambda)^{\alpha\theta}] \le (1+q)\mathbf{E}_{s,z}\rho(j) + q(1+\lambda^{\alpha\theta}),\\
\mathbf{E}_{s,z}\sigma(j) &\le \mathbf{E}_{s,z}\sup_{z'\in\Omega_\lambda^i}\mathbf{E}_{\rho(j-1),z'}\sigma^{i,\lambda} \le \mathbf{E}_{s,z}[\rho(j-1) + q(\rho(j-1)+\lambda)^{\alpha\theta}]\\
&\le (1+q)\mathbf{E}_{s,z}\rho(j-1) + q(1+\lambda^{\alpha\theta}),\\
\mathbf{E}_{s,z}\rho(j) &\le \mathbf{E}_{s,z}\sup_{z'\in U_\lambda}\mathbf{E}_{\sigma(j),z'}S^\lambda \le 2\mathbf{E}_{s,z}[\sigma(j) + (\sigma(j)+\lambda)^{\alpha\theta}]\\
&\le 3\mathbf{E}_{s,z}\sigma(j) + 2(1+\lambda^{\alpha\theta}).
\end{aligned} \tag{5.100}$$

Thus for $j \geq 1$, the previous recursive estimates give

$$\mathbf{E}_{s,z}\rho(j) \leq 3(1+q)\mathbf{E}_{s,z}\rho(j-1) + (3q+2)(1+\lambda^{\alpha\theta}) \leq c_2\mathbf{E}_{s,z}\rho(j-1) + c_2\lambda^{\alpha\theta},$$

$$\mathbf{E}_{s,z}\rho(j) \leq c_2^j\mathbf{E}_{s,z}\rho(0) + c_2\frac{c_2^{j-1}-1}{c_2-1}\lambda^{\alpha\theta} \leq c_2^j(s+c_3\lambda^{\alpha\theta}),$$

$$\mathbf{E}_{s,z}T(j) \leq (1+q)c_2^j s + c_3(1+q)c_2^j\lambda^{\alpha\theta} + q(1+\lambda^{\alpha\theta}) \leq c_4c_2^j(s+\lambda^{\alpha\theta}). \tag{5.101}$$

Finally, we obtain the statement of the lemma

$$\mathbf{E}_{s,z}T^{i,\lambda} \leq \mathbf{E}_{s,z}\sigma^{i,\lambda} + c_4(s+\lambda^{\alpha\theta})\sum_{j=1}^{\infty}(c_1c_2\lambda^{-\delta_1})^j \leq (s+q_i^{-1}(s+\lambda)^{\alpha\theta})(1+\mathcal{O}(\lambda^{-\delta})) \tag{5.102}$$

with $0 < \delta < \delta_1$ and λ sufficiently big. $\square$

Proof of the Theorem 5.1. 1. Let Lemma 5.4 hold for $0 < \gamma \leq \gamma_0$ and $\delta_1 > 0$. It is clear that for $z \in B_i$ the inequalities $\sigma_{0,z}^{i,\lambda} \leq T_{0,z}^{i,\lambda} \leq \tau_{0,z}^{i,\lambda}$ hold a.s. The main contribution to $\tau_{0,z}^{i,\lambda}$ is made by the switching time $T_{0,z}^{i,\lambda}$, for if the trajectory enters Ω_λ^j for some $j \neq i$, it follows the deterministic trajectory with high probability and thus reaches the set B_j in a short time due to (3.25). Namely for $0 < p < \alpha\theta$,

$$\mathbf{P}_{0,z}(\tau^{i,\lambda} \leq T^{i,\lambda} + \lambda^p) \to 1 \tag{5.103}$$

as $\lambda \to \infty$. Indeed, on the event $A_\lambda = \{\omega : \sup_{t\in[0,\lambda^p]} |\int_0^t \frac{\mathrm{d}L(u)}{(T^{i,\lambda}+\lambda+u)^\theta}| \leq \lambda^{-4\gamma}\}$, the trajectory $Z_{0,z}^\lambda(t+T^{i,\lambda})$ follows the deterministic trajectory $X^0_{Z_{0,z}^\lambda(T^{i,\lambda})}(t)$ in the sense of (3.23) or (3.24) and (5.103) follows from the estimate $\mathbf{P}_{0,z}(A_\lambda^c) \leq \lambda^{-\delta_2}$ for some positive δ_2. Further,

$$\begin{aligned}\mathbf{P}_{0,z}(Z^\lambda(\tau^{i,\lambda}) \in B_j) &\geq \mathbf{P}_{0,z}(Z^\lambda(\tau^{i,\lambda}) \in B_j, Z^\lambda(T^{i,\lambda}) \in \Omega_\lambda^j, A_\lambda)\\ &\geq \mathbf{P}(Z^\lambda(T^{i,\lambda}) \in \Omega_\lambda^j) - P(A_\lambda^c) \geq q_{ij}q_i^{-1}(1-\lambda^{-\delta_1}) - \lambda^{-\delta_2}.\end{aligned} \tag{5.104}$$

The converse estimate is obtained analogously to Proposition 4.1.

2. Before entering $\cup_{j\neq i} B_j$, the process Z^λ may repeatedly make cycles $\Omega_\lambda^i \leftrightarrows \cup_{j\neq i}\Omega_\lambda^j$. Taking into account (5.103) and that $p < \alpha\theta$, we obtain the estimate for the mean transition time $\mathbf{E}_{0,z}\tau^{i,\lambda}$ analogously to that of $\mathbf{E}_{0,z}T^{i,\lambda}$. $\square$

6. Embedded Markov chain

Let $0 < \Delta \leq \Delta_0$ and $B := \cup_{i=1}^n B_i$. For $s \geq 0$ and $z \in B$, consider the stopping times and indices $\tau_{s,z}^\lambda(0) = s$, $m_{s,z}^\lambda(0) = \sum_{i=1}^n i \cdot \mathbb{I}\{z \in B_i\}$, and

$$\begin{aligned}\tau_{s,z}^\lambda(k) &= \inf\left\{t \geq \tau_{s,z}^\lambda(k-1) : Z_{s,z}^\lambda(t) \in \cup_{j\neq m_{s,z}^\lambda(k-1)} B_j\right\},\\ m_{s,z}^\lambda(k) &= \sum_{i=1}^n i \cdot \mathbb{I}\{Z_{s,z}^\lambda(\tau_{s,z}^\lambda(k)) \in B_i\}, \quad k \geq 1.\end{aligned} \tag{6.105}$$

Consider a Markov chain $(U_{s,z}^\lambda(k))_{k\geq 0}$ with the state space B such that

$$U_{s,z}^\lambda(k) = Z_{s,z}^\lambda(\tau_{s,z}^\lambda(k)). \tag{6.106}$$

Since for all $k \geq 0$ and $s \geq 0$, $\mathbf{E}_{s,z}\tau^\lambda(k) < \infty$, the Markov chain $U^\lambda(k)$ is well-defined. Note that U^λ is not stationary since its one-step transition probabilities depend on the position of Z^λ.

The goal of this section is to compare the distribution of $U^\lambda(k)$ for $k \to +\infty$ with the invariant distribution π^0 of the limiting Markov chain from Theorem 2.1. Denote $p^{(k)}_{s,z}(j) := \mathbf{P}_{s,z}(U^\lambda(k) \in B_j)$ and $p_{s,z}(j) := \mathbf{P}_{s,z}(U^\lambda(1) \in B_j)$.

Lemma 6.1. *For any $0 < \Delta \leq \Delta_0$ and $\varepsilon > 0$, there are $\lambda_0 \geq 0$ and $k_0 \geq 0$ such that for $\lambda \geq \lambda_0$ and $k \geq k_0$ the estimate $\max_{1\leq j\leq n} |p^{(k)}_{s,z}(j) - \pi^0_j| \leq \varepsilon$ holds uniformly for $z \in B$ and $s \geq 0$.*

Proof. To prove Lemma 6.1, we compare the time-nonhomogeneous Markov chain U with a time-homogeneous irreducible Markov chain which has the stationary distribution $\pi^0 = \sum_{j=0}^n \pi^0_j \delta_{m_j}$. We adapt the argument by Kartashov [24], which was developed for more general continuous-time Markov processes.

1. Denote $p_{ij} = q_{ij}/q_i$ if $i \neq j$ and $p_{ii} = 0$. Then $(p_{ij})^n_{i,j=1}$ is a stochastic matrix. Let $(V_z(k))_{k\geq 0}$ be a time-homogeneous Markov chain on B with the transition kernel $Q(z, A) = \sum_{i=1}^n \left[\mathbb{I}\{z \in B_i\} \sum_{j=1}^n p_{ij}\delta_{m_j}(A)\right]$. Then V is an irreducible recurrent Markov chain with the invariant measure π^0. Denote $Q^{(k)}_z(A) \equiv Q^{(k)}_{s,z}(A) := \mathbf{P}_{s,z}(V(k) \in A)$ and $Q_{s,z}(A) := Q^{(1)}_z(A)$. The Markov chain V is geometrically ergodic, and we denote

$$\rho = \sup_{z\in B} \max_{1\leq j\leq n} \sum_{k=0}^{\infty} |Q^{(k)}_z(B_j) - \pi^0_j|. \tag{6.107}$$

Let $0 < \Delta \leq \Delta_0$ and $0 < \varepsilon < 1$. According to Theorem 5.1, we choose $\lambda_0 > 0$ such that for $\lambda \geq \lambda_0$

$$\sup_{z\in B_i} |\mathbf{P}_{s,z}(Z^\lambda(\tau^{i,\lambda}) \in B_j) - p_{ij}| < \frac{\varepsilon}{4n^2\rho}, \tag{6.108}$$

uniformly for $s \geq 0$ and $1 \leq i, j \leq n$. For $k \geq 0$ consider the distance $d(k) := \sup_{z\in B} \max_{1\leq j\leq n} \left|P^{(k)}_{0,z}(B_j) - Q^{(k)}_z(B_j)\right|$. Fix $k \geq 1$ and $1 \leq j \leq n$. For $z \in B$, consider a decomposition

$$P^{(k)}_{0,z}(B_j) - Q^{(k)}_{0,z}(B_j) = \sum_{i=1}^k \left[\mathbf{E}_{0,z} Q^{(k-i)}_{\tau(i),U(i)}(B_j) - \mathbf{E}_{0,z} Q^{(k-i+1)}_{\tau(i-1),U(i-1)}(B_j)\right]. \tag{6.109}$$

For $i = 1, \ldots, k$, we transform the summands in the latter formula

$$\begin{aligned}
&\mathbf{E}_{0,z} Q^{(k-i)}_{\tau(i),U(i)}(B_j) - \mathbf{E}_{0,z} Q^{(k-i+1)}_{\tau(i-1),U(i-1)}(B_j) \\
&\quad = \mathbf{E}_{0,z}\mathbf{E}_{\tau(i-1),U(i-1)} Q^{(k-i)}_{U(1)}(B_j) - \mathbf{E}_{0,z}\mathbf{E}_{\tau(i-1),U(i-1)} Q^{(k-i)}_{V(1)}(B_j) \\
&\quad = \sum_{l=1}^n [\mathbf{E}_{0,z} P_{\tau(i-1),U(i-1)}(B_l) - \mathbf{E}_{0,z} Q_{\tau(i-1),U(i-1)}(B_l)] Q^{(k-i)}_{m_l}(B_j) \\
&\quad = \sum_{l=1}^n [\mathbf{E}_{0,z} P_{\tau(i-1),U(i-1)}(B_l) - \mathbf{E}_{0,z} Q_{\tau(i-1),U(i-1)}(B_l)][Q^{(k-i)}_{m_l}(B_j) - \pi^0_j] \\
&\quad = \sum_{l=1}^n [(\mathbf{E}_{0,z} P_{\tau(i-1),U(i-1)}(B_l) - \mathbf{E}_{0,z} P_{\tau(i-1),V(i-1)}(B_l))
\end{aligned}$$

$$
\begin{aligned}
&- (\mathbf{E}_{0,z} Q_{\tau(i-1),U(i-1)}(B_l) - \mathbf{E}_{0,z} Q_{\tau(i-1),V(i-1)}(B_l))][Q^{(k-i)}_{m_l}(B_j) - \pi^0_j] \\
&+ \sum_{l=1}^{n} [\mathbf{E}_{0,z} P_{\tau(i-1),V(i-1)}(B_l) - \mathbf{E}_{0,z} Q_{\tau(i-1),V(i-1)}(B_l)][Q^{(k-i)}_{m_l}(B_j) - \pi^0_j] \\
=: &\sum_{l=1}^{n} A_z(i,l)[Q^{(k-i)}_{m_l}(B_j) - \pi^0_j] + \sum_{l=1}^{n} B_z(i,l)[Q^{(k-i)}_{m_l}(B_j) - \pi^0_j]. \qquad (6.110)
\end{aligned}
$$

For $z \in B_i$, $1 \le i \le k$ and $1 \le j \le n$, we estimate

$$
\begin{aligned}
|A_z(i,l)| &:= |\mathbf{E}_{0,z} P_{\tau(i-1),U(i-1)}(B_l) - \mathbf{E}_{0,z} P_{\tau(i-1),V(i-1)}(B_l) \\
&\quad - \mathbf{E}_{0,z} Q_{\tau(i-1),U(i-1)}(B_l) - \mathbf{E}_{0,z} Q_{\tau(i-1),V(i-1)}(B_l)| \\
&= \left| \sum_{m=1}^{n} \int_{B_m} (P^{(i-1)}_{0,z}(\mathrm{d}u) - Q^{(i-1)}_{0,z}(\mathrm{d}u)) P_{\tau(i-1),u}(B_l) \right. \\
&\quad \left. - \sum_{m=1}^{n} \int_{B_m} (P^{(i-1)}_{0,z}(\mathrm{d}u) - Q^{(i-1)}_{0,z}(\mathrm{d}u)) Q_{\tau(i-1),u}(B_l) \right| \\
&\le \sum_{m=1}^{n} |Q^{(i-1)}_{0,z}(B_m) - Q^{(i-1)}_{0,z}(B_m)| \times \max_{1\le m\le n} \sup_{s\ge 0, u\in B_m} |P_{s,u}(B_l) - Q_u(B_l)| \\
&< nd(i-1) \cdot \frac{\varepsilon}{4n^2\rho} = d(i-1) \cdot \frac{\varepsilon}{4n\rho}. \qquad (6.111)
\end{aligned}
$$

Analogously,

$$
\begin{aligned}
|B_z(i,l)| &:= |\mathbf{E}_{0,z} P_{\tau(i-1),V(i-1)}(B_l) - \mathbf{E}_{0,z} Q_{\tau(i-1),V(i-1)}(B_l)| \\
&\le \max_{1\le m\le n} \sup_{s\ge 0, u\in B_m} |P_{s,u}(B_l) - Q_u(B_l)| \le \frac{\varepsilon}{4n^2\rho}. \qquad (6.112)
\end{aligned}
$$

Taking into account estimates (6.110) and (6.112), we obtain from (6.109) that

$$
\begin{aligned}
d(k) &< \frac{\varepsilon}{4n\rho} \sum_{i=1}^{k} d(i-1) \sum_{l=1}^{n} |Q^{(k-i)}_{m_l}(B_j) - \pi^0_j| + \frac{\varepsilon}{4n^2\rho} \sum_{i=1}^{k} \sum_{l=1}^{n} |Q^{(k-i)}_{m_l}(B_j) - \pi^0_j| \\
&< \frac{\varepsilon}{4} \left(\max_{0\le i\le k-1} d(i) + 1 \right). \qquad (6.113)
\end{aligned}
$$

Finally, $d(0) = 0$ implies that $d(k) < \varepsilon/2$ for all $k \ge 1$. The time-homogeneous Markov chain $V(k)$ converges to its stationary distribution geometrically fast, and hence there is $k_0 \ge 1$ such that for $k \ge k_0$

$$
\max_{z\in B} \max_{1\le j\le n} |Q^{(k)}_z(B_j) - \pi^0_j| < \frac{\varepsilon}{2}. \qquad (6.114)
$$

The statement of the lemma follows from the triangle inequality. $\square$

7. Slow cooling, $\alpha\theta < 1$. Proof of Theorem 2.2

Lemma 7.1. *Let $\alpha\theta < 1$. There is $\gamma_0 > 0$ such that for any $0 < \gamma \le \gamma_0$, $0 < \delta < \alpha\theta$ and $0 < \Delta \le \Delta_0$*

$$\max_{1\le i\le n} \sup_{z\in\Omega^i_\lambda} \mathbf{P}_{0,z}(Z^\lambda(t) \notin B_i \text{ for some } t \in [\lambda^\delta, 2\lambda^\delta]) \to 0, \quad \lambda \to +\infty. \tag{7.115}$$

Proof. Choose $\gamma_0 > 0$ such that Proposition 4.1 holds and let $0 < \Delta \le \Delta_0$. Then for $0 < \gamma \le \gamma_0$ and some $0 < p < \delta$ consider an event $A_\lambda = \left\{\sup_{0\le t\le\lambda^p} \left|\int_0^t \frac{\mathrm{d}L(u)}{(\lambda+u)^\theta}\right| \le \frac{1}{\lambda^{4\gamma}}\right\}$ as in the proof of Theorem 5.1. Fix $i = 1,\dots,n$. Due to (3.25), the inequality $|Z^\lambda_{0,z}(\lambda^p) - m_i| \le \Delta/2$ holds on A_λ a.s. for all $z \in \Omega^i_\lambda$. For any $\varepsilon > 0$, one has the probability $\mathbf{P}(A^c_\lambda) < \varepsilon/2$ for λ sufficiently large (see estimates Eqs. (5.87) and (5.89)). Then

$$\begin{aligned}
&\sup_{z\in\Omega^i_\lambda} \mathbf{P}_{0,z}\left(Z^\lambda(u) \notin B_i \text{ for some } u \in [\lambda^\delta, 2\lambda^\delta]\right)\\
&\quad\le \sup_{|z-m_i|\le\Delta/2} \mathbf{P}_{\lambda^p,z}(Z^\lambda(u) \notin B_i \text{ for some } u \in [\lambda^p, 2\lambda^\delta]) + \mathbf{P}(A^c_\lambda)\\
&\quad\le \sup_{|z-m_i|\le\Delta/2} \mathbf{P}_{0,z}(Z^{\lambda+\lambda^p}(u) \notin B_i \text{ for some } u \in [0, 2\lambda^\delta - \lambda^p]) + \mathbf{P}(A^c_\lambda)\\
&\quad\le \sup_{|z-m_i|\le\Delta/2} \mathbf{P}_{0,z}(\sigma^{i,\lambda+\lambda^p}_\Delta \le 2\lambda^\delta - \lambda^p) + \mathbf{P}(A^c_\lambda),
\end{aligned} \tag{7.116}$$

where $\sigma^{i,\lambda}_\Delta$ is the first exit time from B_i of the process $Z^\lambda_{0,z}$. Due to the estimate (4.69) in Proposition 4.6, we obtain for λ sufficiently large that

$$\sup_{|z-m_i|\le\Delta/2} \mathbf{P}_{0,z}(\sigma^{i,\lambda+\lambda^p}_\Delta \le 2\lambda^\delta - \lambda^p) \le \sup_{|z-m_i|\le\Delta/2} \mathbf{P}_{0,z}(\sigma^{i,\lambda+\lambda^p}_\Delta \le (\lambda+\lambda^p)^{\delta_1}) < \frac{\varepsilon}{2}, \tag{7.117}$$

for $\delta < \delta_1 < \alpha\theta$ and λ sufficiently large so that $(2\lambda^\delta - \lambda^p)/(\lambda+\lambda^p)^{\delta_1} \le 1$. Since the number of wells is finite, the statement of the lemma holds uniformly for all $1 \le i \le n$ for λ large enough. $\square$

Lemma 7.2. *Let $\alpha\theta < 1$. For any $\lambda > 0$ and $0 < \Delta \le \Delta_0$, we have*

$$\mathbf{P}_{0,z}(Z^\lambda(t) \notin B) \to 0, \quad t \to +\infty, \tag{7.118}$$

uniformly for all $z \in \mathbb{R}$.

Proof. Let $\varepsilon > 0$. The Markov property implies that

$$\mathbf{P}_{0,z}(Z^\lambda(t) \notin B) = \mathbf{E}_{0,z}\left[\mathbf{P}_{0,Z^\lambda(s)}(Z^{\lambda+s}(t-s) \notin B)\right] \tag{7.119}$$

for any $0 \le s \le t$. Fix $\gamma > 0$ and $0 < \delta < \alpha\theta$ such that Lemmas 5.3 and 7.1 hold for $\lambda \ge \lambda_1$, λ_1 being sufficiently big. In particular, $\sup_{z\in U_\lambda} \mathbf{E}_{0,z} S^\lambda \le \lambda^{\alpha\theta-\delta}$ for $\lambda \ge \lambda_1$. For $t > 0$ and $\lambda > 0$, let s_t be a solution of the equation $t - s_t = 2(\lambda+s_t)^{\alpha\theta-\delta/2}$. It is clear that s_t exists and is positive for large values of t, and $s_t \to \infty$ as $t \to \infty$. Let t_1 be big enough such that $\mu := \lambda + s_t \ge \lambda_1$ for

$t \geq t_1$. Denote $\hat{\mu} := \mu^{\alpha\theta-\delta/2}$. Then due to the Chebyshev inequality, there is $t_2 \geq t_1$ such that

$$\sup_{z\in U_\mu} \mathbf{P}_{0,z}\left(S^\mu \geq \tfrac{1}{2}\hat{\mu}\right) \leq 2\mu^{-\delta/2} < \tfrac{1}{2}\varepsilon, \quad t \geq t_2. \tag{7.120}$$

Moreover, due to Lemma 7.1, there is $t_3 \geq t_2$ such that the right-hand side of the Eq. (7.119) with $s = s_t$ and $z \in \cup_{i=1}^n \Omega_{\lambda+s_t}$ is estimated as

$$\begin{aligned}\sup_{z\in\bigcup_{i=1}^n \Omega_{\lambda+s_t}} \mathbf{P}_{0,z}(Z^{\lambda+s_t}(t-s_t) \notin B) &= \sup_{z\in\bigcup_{i=1}^n \Omega_\mu} \mathbf{P}_{0,z}(Z^\mu(2\hat{\mu}) \notin B)\\ &\leq \max_{1\leq i\leq n} \sup_{z\in\Omega_\mu^i} \mathbf{P}_{0,z}\left(Z^\mu(u) \notin B_i \text{ for some } u \in [\hat{\mu}, 2\hat{\mu}]\right) < \tfrac{1}{2}\varepsilon, \quad t \geq t_3.\end{aligned} \tag{7.121}$$

Finally, consider the right-hand side of the Eq. (7.119) for $s = s_t$ and $z \in U_{\lambda+s_t}$:

$$\begin{aligned}\mathbf{P}_{0,z}(Z^{\lambda+s_t}(t-s_t) \notin B) &\leq \mathbf{P}_{0,z}\left(Z^\mu(2\hat{\mu}) \notin B, S^\mu \leq \tfrac{1}{2}\hat{\mu}\right) + \mathbf{P}_{0,z}\left(S^\mu > \tfrac{1}{2}\hat{\mu}\right)\\ &< \mathbf{E}_{0,z}\left[\mathbb{I}\left\{S^\mu \leq \tfrac{1}{2}\hat{\mu}\right\}\mathbf{P}_{S^\mu, Z^\mu(S^\mu)}(Z^\mu(2\hat{\mu}) \notin B)\right] + \tfrac{1}{2}\varepsilon\\ &\leq \mathbf{E}_{0,z}\left[\mathbb{I}\left\{S^\mu \leq \tfrac{1}{2}\hat{\mu}\right\}\max_{1\leq i\leq n}\sup_{z\in\Omega_\mu^i}\mathbf{P}_{0,z}(Z^{\mu+S^\mu}(2\hat{\mu}-S^\mu) \notin B_i)\right] + \tfrac{1}{2}\varepsilon\\ &\leq \max_{1\leq i\leq n}\sup_{z\in\Omega_\mu^i}\mathbf{P}_{0,z}(Z^{\mu+S^\mu}(u) \notin B_i, u \in [(\mu+S^\mu)^{\alpha\theta-\delta/2}, 2(\mu+S^\mu)^{\alpha\theta-\delta/2}])\\ &\quad + \tfrac{1}{2}\varepsilon < \varepsilon,\end{aligned} \tag{7.122}$$

where we used the inequality $(\mu+s)^a \leq 2\mu^a - s \leq 2(\mu+s)^a$ for $0 \leq s \leq \mu^a/2$, $\mu \geq 0$, $0 < a \leq 1$, and applied Lemma 7.1. □

Proof of the Theorem 2.2. Let $\varepsilon > 0$, f be a bounded continuous real function, and $|f(x)| \leq C_f$ for all $x \in \mathbb{R}$. We show that for any $z \in \mathbb{R}$ and $\lambda > 0$ there is $t_0 > 0$ large enough such that $|\mathbf{E}_{0,z} f(Z_{0,z}^\lambda(t)) - \sum_{i=j}^n \pi_i^0 f(m_j)| \leq \varepsilon$ for $t \geq t_0$.

Choose $0 < \Delta \leq \Delta_0$ such that $\max_{1\leq j\leq n}\sup_{y\in B_j}|f(y) - f(m_j)| \leq \frac{\varepsilon}{8}$. Choose $t_1 \geq 0$ such that according to Lemma 7.2 the estimate $\mathbf{P}_{0,z}(Z^\lambda(t) \notin B) \leq \frac{\varepsilon}{8C_f}$ holds for $t \geq t_1$, and according to Lemma 6.1 the estimate $\sup_{z\in B}\max_{1\leq j\leq n}|\mathbf{P}(U_{s,z}^{\lambda+t}(k) \in B_j) - \pi_j^0| \leq \frac{\varepsilon}{8nC_f}$ holds for $t \geq t_1$, $k \geq k_0$ and all $s \geq 0$. Denote $\lambda_1 := \lambda + t_1$ and apply the Markov property to obtain for $t \geq 0$

$$\begin{aligned}\mathbf{E}_{0,z} f(Z^\lambda(t+t_1)) = \mathbf{E}_{0,z}\mathbf{E}_{0,Z^\lambda(t_1)} f(Z^{\lambda_1}(t)) &\leq \sup_{z\in B}\mathbf{E}_{0,z} f(Z^{\lambda_1}(t))\\ &\quad + C_f\mathbf{P}_{0,z}(Z^\lambda(t_1) \notin B),\end{aligned} \tag{7.123}$$

with the last summand bounded by $\varepsilon/8$. Let $(\tau_{0,z}^{\lambda_1}(k))_{k\geq 0}$ be transition times defined in (6.105). For $z \in B$, consider a (non-Markovian) jump process $(\tilde{Z}_{0,z}^{\lambda_1}(t))_{t\geq 0}$ defined by

$$\tilde{Z}_{0,z}^{\lambda_1}(t) = \sum_{k=0}^\infty U_{0,z}^{\lambda_1}(k)\mathbb{I}\{t \in [\tau^{\lambda_1}(k), \tau^{\lambda_1}(k+1))\}. \tag{7.124}$$

For $z \in B$ and $t \geq 0$, we estimate

$$\begin{aligned} |\mathbf{E}_{0,z} f(Z^{\lambda_1}(t)) - \pi^0 f| &\leq \sum_{i=1}^{n} \mathbf{E}_{0,z} |f(Z^{\lambda_1}(t)) - f(m_i)| \mathbb{I}\{Z^{\lambda_1}(t) \in B_i\} \\ &+ \sum_{i=1}^{n} |f(m_i)| \cdot |\mathbf{P}_{0,z}(Z^{\lambda_1}(t) \in B_i) - \mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i)| \\ &+ \sum_{i=1}^{n} |f(m_i)| \cdot |\mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i) - \pi_i^0| + C_f \mathbf{P}_{0,z}(Z^{\lambda_1}(t) \notin B). \end{aligned} \tag{7.125}$$

The first summand in (7.125) does not exceed $\varepsilon/8$ due to the definition of Δ. To estimate the second summand in (7.125), we apply the total probability formula to obtain for any $1 \leq i \leq n$ that

$$\begin{aligned} &|\mathbf{P}_{0,z}(Z^{\lambda_1}(t) \in B_i) - \mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i)| \\ &\quad = |\mathbf{P}_{0,z}(Z^{\lambda_1}(t) \in B_i, \tilde{Z}^{\lambda_1}(t) \in B_i) + \mathbf{P}_{0,z}(Z^{\lambda_1}(t) \in B_i, \tilde{Z}^{\lambda_1}(t) \notin B_i)[=0] \\ &\quad\quad - \mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i, Z^{\lambda_1}(t) \in B_i) - \mathbf{P}_{0,z}\left(\tilde{Z}^{\lambda_1}(t) \in B_i, Z^{\lambda_1}(t) \in \cup_{j \neq i} B_i\right)[=0] \\ &\quad\quad - \mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i, Z^{\lambda_1}(t) \notin B)| \leq \mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i, Z^{\lambda_1}(t) \notin B) \end{aligned} \tag{7.126}$$

and thus

$$\sum_{i=1}^{n} |f(m_i)| \mathbf{P}_{0,z}(Z^{\lambda_1}(t) \in B_i) - \mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i)| \leq C_f \mathbf{P}_{0,z}(Z^{\lambda_1}(t) \notin B). \tag{7.127}$$

According to Lemma 7.2 we choose $t_2 > 0$ such that $\mathbf{P}_{0,z}(Z^{\lambda_1}(t) \notin B) \leq \frac{\varepsilon}{4C_f}$ for $t \geq t_2$. Finally, to estimate the third summand in (7.125) consider for $1 \leq i \leq n$

$$\begin{aligned} |\mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i) - \pi_i^0| &\leq |\mathbf{P}_{0,z}(\tilde{Z}^{\lambda_1}(t) \in B_i, t \geq \tau^{\lambda_1}(k_0)) - \pi_i^0| + \mathbf{P}_{0,z}(t < \tau^{\lambda_1}(k_0)) \\ &\leq \sup_{k \geq k_0} |\mathbf{P}_{0,z}(U^{\lambda_1}(k) \in B_i) - \pi_i^0| + \mathbf{P}_{0,z}(t < \tau^{\lambda_1}(k_0)). \end{aligned} \tag{7.128}$$

Since $\mathbf{E}_{0,z} \tau^{\lambda_1}(k_0) < \infty$, we apply Chebyshev's inequality to get

$$\sup_{z \in B} \mathbf{P}_{0,z}(\tau^{\lambda_1}(k_0) > t) \leq \frac{\varepsilon}{8nC_f} \tag{7.129}$$

for $t \geq t_3 \geq t_2$. Thus, we obtain the statement of the theorem for all $t \geq t_1 + t_3$. □

8. Fast cooling, $\alpha\theta > 1$

Proof of the Theorem 2.3. Consider a well Ω^i and assume for definiteness that $m_i = 0$. We prove the statement of the theorem for the stopping time $\sigma(\lambda) = \{t \geq 0 : |Z_{0,z}^{\lambda}(t)| > a\}$ for some $0 < a < |s_{i-1}| \wedge s_i$. Since the properties of $Z^{\lambda}(t)$ for $t \in [0, \sigma(\lambda)]$ are determined by the jump process L and the values of the potential $U(x)$ for $x \in [-a, a]$, we assume that $U'(\cdot)$ is globally Lipschitz with constant $K > 0$.

It is clear that

$$\mathbf{P}_{0,z}(\sigma(\lambda) < \infty) = \sum_{k=1}^{\infty} \mathbf{P}_{0,z}\left(\sup_{t \in (0,k-1]} |Z_{0,z}^{\lambda}(t)| \leq a, \sup_{t \in (k-1,k]} |Z_{0,z}^{\lambda}(t)| > a\right). \tag{8.130}$$

Let $X_z^0(t)$ be the deterministic trajectory of the underlying dynamical system. Denote $\delta = \frac{1}{2}\min\{a - X_a^0(1), a - |X_{-a}^0(1)|\}$ and $\Delta = a - \delta$. Then for any $|z| \le a$, Gronwall's lemma yields

$$\sup_{t\in[0,1]} |Z_{0,z}^\lambda(t) - X_z^0(t)| \le \mathrm{e}^K \sup_{t\in[0,1]} \left| \int_0^t \frac{\mathrm{d}L(s)}{(\lambda+s)^\theta} \right| \le 2\mathrm{e}^K \frac{\sup_{t\in[0,1]} |L(s)|}{\lambda^\theta}. \tag{8.131}$$

For $\lambda \ge \lambda_0$, λ_0 being sufficiently large, and $|z| \le \Delta$, we estimate

$$\begin{aligned} \mathbf{P}_{0,z}\left(\sup_{t\in(0,1]} |Z^\lambda(t)| > a\right) &\le \mathbf{P}_{0,z}\left(\sup_{t\in(0,1]} |Z^\lambda(t) - X_z^0(t)| > \delta\right) \\ &\le \mathbf{P}\left(\sup_{t\in[0,1]} |L(s)| \ge \lambda^\theta \frac{\delta}{2\mathrm{e}^K}\right) \le 4\mathbf{P}\left(L(1) \ge \lambda^\theta \frac{\delta}{2\mathrm{e}^K}\right) \le c_1 \lambda^{-\alpha\theta}. \end{aligned} \tag{8.132}$$

For $k \ge 2$, we apply (2.8) and (8.132) to obtain

$$\begin{aligned} &\mathbf{P}_{0,z}\left(\sup_{t\in(0,k-1]} |Z^\lambda(t)| \le a, \sup_{t\in(k-1,k]} |Z^\lambda(t)| > a\right) \\ &\quad\le \sup_{|z|\le a} \mathbf{P}_{0,z}\left(\sup_{t\in(0,1]} |Z^{\lambda+k-2}(t)| \le a, \sup_{t\in(1,2]} |Z^{\lambda+k-2}(t)| > a\right) \\ &\quad\le \sup_{|z|\le a} \mathbf{P}_{0,z}\left(\sup_{t\in(0,1]} |Z^{\lambda+k-2}(t) - X_z^0(t)| > \delta\right) + \sup_{|z|\le\Delta} \mathbf{P}_{0,z}\left(\sup_{t\in(0,1]} |Z^{\lambda+k-1}(t)| > a\right) \\ &\quad\le c_1(\lambda+k-2)^{-\alpha\theta} + c_1(\lambda+k-1)^{-\alpha\theta}. \end{aligned} \tag{8.133}$$

Finally, estimating the sum by an integral we find that

$$\mathbf{P}_{0,z}(\sigma(\lambda) < \infty) \le 2c_1 \sum_{k=0}^\infty \frac{1}{(\lambda+k)^{\alpha\theta}} \le \frac{2c_1}{\alpha\theta - 1} \frac{1}{(\lambda-1)^{\alpha\theta-1}}. \quad \square \tag{8.134}$$

Acknowledgements

The author is grateful to Peter Imkeller for stimulating discussions and to the anonymous referees for their careful reading and helpful suggestions which lead to a significant improvement of the manuscript.

Appendix. Asymptotic estimates of sums and integrals

The arguments of Sections 4 and 5 involve asymptotic estimates of sums and integrals which can be reduced to the so-called Laplace type integrals of the form $I(\lambda) = \int_a^b q(u)\exp(-\lambda p(u))\mathrm{d}u$, where $\lambda \to +\infty$ is a big parameter. For the general theory, we refer the reader to the book by Olver [25, Chapter 3, Sections 7–9]. Roughly speaking, the evaluation method of $I(\lambda)$ is based on the observation, that in the limit of big λ the main contribution to

$I(\lambda)$ comes from the neighbourhood of the global minimum of $p(u)$ over $[a, b]$. For example, if $p(u)$ attains its minimum at a, and $p'(a) > 0$ and $q(a) \neq 0$, then we can approximate

$$I(\lambda) \approx \mathrm{e}^{\lambda p(a)} \int_a^\infty q(a)\mathrm{e}^{-\lambda p'(a)(u-a)}\mathrm{d}u = \mathrm{e}^{\lambda p(a)} \frac{q(a)}{\lambda p'(a)}, \quad \lambda \to +\infty. \tag{A.1}$$

The precise results and estimates of the error terms can be found in [25], Theorems 7.1 and 8.1 and Section 9.4 in Chapter 3. Below we formulate the estimates that we use in our proofs.

Lemma A.1. *Let $\alpha\theta < 1$, $A > 0$, $\delta > 0$ and $A_\lambda = A(1 + \mathcal{O}(\lambda^{-\delta}))$. Then there is $\delta_1 > 0$ such that*

$$\int_0^\infty \frac{At}{(\lambda+t)^{\alpha\theta}} \exp\left(-\frac{A_\lambda t}{(\lambda+t)^{\alpha\theta}}\right) \mathrm{d}t = \frac{\lambda^{\alpha\theta}}{A}(1 + \mathcal{O}(\lambda^{-\delta_1})), \tag{A.2}$$

$$\int_0^\infty \frac{At}{(\lambda+t)^{\alpha\theta}} \exp\left(-\frac{A_\lambda t}{\lambda^{\alpha\theta}}\right) \mathrm{d}t = \frac{\lambda^{\alpha\theta}}{A}(1 + \mathcal{O}(\lambda^{-\delta_1})), \quad \lambda \to +\infty. \tag{A.3}$$

Proof. First, we note that condition $\alpha\theta < 1$ guarantees convergence of the integrals for all positive A, δ and λ. To prove (A.2), we introduce a new variable $u = \frac{\lambda+t}{\lambda}$ to obtain an integral of Laplace type:

$$\int_0^\infty \frac{At}{(\lambda+t)^{\alpha\theta}} \exp\left(-\frac{A_\lambda t}{(\lambda+t)^{\alpha\theta}}\right) \mathrm{d}t = A\lambda^{2-\alpha\theta} \int_1^\infty \frac{u-1}{u^{\alpha\theta}} \exp\left(-\frac{u-1}{u^{\alpha\theta}} A_\lambda \lambda^{1-\alpha\theta}\right) \mathrm{d}u. \tag{A.4}$$

The big parameter here is equal to $A_\lambda \lambda^{1-\alpha\theta}$. Applying the Laplace method, we obtain the equality (A.2) for some $\delta_1 > 0$. The second equality is proved analogously. □

Recall the notation $\varphi_{\lambda,k} = \lambda^{\alpha\theta/2}(\lambda + k\lambda^q)^{-\alpha\theta}$ introduced in the Step **R1-1** of Lemma 4.4. It is clear that $\varphi_{\lambda,k} \to 0$ as $\lambda, k \to \infty$.

Lemma A.2. *Let $\alpha\theta < 1$, $c > 0$ and $0 < q < 1 - \alpha\theta/2$. Then there are $C > 0$ and $\lambda_0 > 0$ such that for $\lambda > \lambda_0$*

$$\sum_{k=2}^\infty k(k-1)\left(1 - c\varphi_{\lambda,k}\right)^{k-1} \le C\lambda^{3\alpha\theta/2}, \tag{A.5}$$

$$\sum_{k=1}^\infty k\left(1 - c\varphi_{\lambda,k}\right)^{k-1} \le \sum_{k=2}^\infty k\left(1 - c\varphi_{\lambda,k}\right)^{k-2} \le C\lambda^{\alpha\theta}. \tag{A.6}$$

Proof. First we prove (A.5). Using the elementary inequality $\ln(1-x) \le -x$, $x < 1$, we estimate uniformly for all $k \ge 2$ and λ such that $c\lambda^{-\alpha\theta/2} < 1$:

$$\left(1 - c\varphi_{\lambda,k}\right)^{k-1} = \mathrm{e}^{(k-1)\ln(1-c\varphi_{\lambda,k})} \le \mathrm{e}^{-c(k-1)\varphi_{\lambda,k}} \le c_1 \mathrm{e}^{-ck\varphi_{\lambda,k}}. \tag{A.7}$$

We also note, that due to the condition $\alpha\theta < 1$, the sum (A.5) converges. Further, for λ sufficiently large, we can estimate the sum by an integral, which can be transformed to an integral of Laplace type. Indeed,

$$\sum_{k=2}^\infty k(k-1)\left(1 - c\varphi_{\lambda,k}\right)^{k-1} \le c_1 \sum_{k=1}^\infty k^2 \mathrm{e}^{-ck\varphi_{\lambda,k}}$$

$$\le c_2 \int_0^\infty x^2 \exp\left(-\frac{cx}{\lambda^{-\alpha\theta/2}(\lambda+x\lambda^q)^{\alpha\theta}}\right)\mathrm{d}x \quad \left(u=\frac{\lambda+x\lambda^q}{\lambda}\right)$$
$$= c_2\lambda^{3(1-q)}\int_1^\infty (u-1)^2 \exp\left(-c\lambda^{1-\alpha\theta/2-q}\frac{u-1}{u^{\alpha\theta}}\right)\mathrm{d}u \le C\lambda^{3\alpha\theta/2}. \tag{A.8}$$

Here we used Theorem 7.1 in [25] to evaluate the latter Laplace integral with the big parameter $\lambda^{1-\alpha\theta/2-q}$.

The estimate (A.6) is obtained analogously by reducing the second sum to a Laplace integral

$$\lambda^{2(1-q)}\int_1^\infty (u-1)\exp\left(-c\lambda^{1-\alpha\theta/2-q}\frac{u-1}{u^{\alpha\theta}}\right)\mathrm{d}u \le C\lambda^{\alpha\theta}, \quad \lambda\to+\infty. \quad \square \tag{A.9}$$

Recall the notation $\psi_{\mu,\lambda}(k) = (\mu + k\lambda^{\alpha\theta-\delta})^{-\alpha\theta}$, $\mu,\lambda > 0$, $k \ge 0$, from the proof of Lemma 5.3.

Lemma A.3. *Let $\alpha\theta < 1$. For any $c > 0$, $\gamma > 0$, $0 < \delta < \alpha\theta$, there are $C > 0$ and $\lambda_0 > 0$ such that for all $\mu \ge \lambda \ge \lambda_0$*

$$\lambda^{\alpha\theta-\delta}\sum_{k=1}^\infty k\exp\left(-c\lambda^{\alpha\gamma+\alpha\theta-\delta}\sum_{j=1}^{k-1}\psi_{\mu,\lambda}(j)\right)\lambda^{\alpha\gamma+\alpha\theta-\delta}\psi_{\mu,\lambda}(k-1) \le C\mu^{\alpha\theta}\lambda^{-\alpha\gamma}, \tag{A.10}$$

$$\sum_{k=1}^\infty \exp\left(-c\lambda^{\alpha\gamma+\alpha\theta-\delta}\sum_{j=1}^{k-1}\psi_{\mu,\lambda}(j)\right)\psi_{\mu,\lambda}(k-1) \le C\lambda^{-\alpha\theta+\delta-\alpha\gamma}. \tag{A.11}$$

Proof. We prove the estimate (A.10). For $\mu \ge \lambda$ big enough and $k \ge 1$, we can consider an integral instead of a sum, namely

$$\sum_{j=1}^{k-1}\psi_{\mu,\lambda}(j) \ge c_1\int_0^{k-1}\psi_{\mu,\lambda}(x)\mathrm{d}x = c_2\lambda^{-\alpha\theta+\delta}\left[(\mu+(k-1)\lambda^{\alpha\theta-\delta})^{1-\alpha\theta} - \mu^{1-\alpha\theta}\right]. \tag{A.12}$$

This leads to the following inequality:

$$\lambda^{\alpha\theta-\delta}\sum_{k=1}^\infty k\exp\left(-c\lambda^{\alpha\gamma+\alpha\theta-\delta}\sum_{j=1}^{k-1}\psi_{\mu,\lambda}(j)\right)\lambda^{\alpha\gamma+\alpha\theta-\delta}\psi_{\mu,\lambda}(k-1)$$
$$\le \lambda^{\alpha\gamma+2\alpha\theta-2\delta}\exp[c_3\lambda^{\alpha\gamma}\mu^{1-\alpha\theta}]\sum_{k=0}^\infty (k+1)\psi_{\mu,\lambda}(k)\exp\left[-c_3\lambda^{\alpha\gamma}(\mu+k\lambda^{\alpha\theta-\delta})^{1-\alpha\theta}\right]. \tag{A.13}$$

The latter sum can be again approximated by an integral with a big parameter and evaluated with help of Laplace's method. Indeed,

$$\sum_{k=0}^\infty \frac{k+1}{(\mu+k\lambda^{\alpha\theta-\delta})^{\alpha\theta}}\exp[-c_3\lambda^{\alpha\gamma}(\mu+k\lambda^{\alpha\theta-\delta})^{1-\alpha\theta}]$$
$$\le c_4\int_0^\infty \frac{x+1}{(\mu+x\lambda^{\alpha\theta-\delta})^{\alpha\theta}}\exp[-c_3\lambda^{\alpha\gamma}(\mu+x\lambda^{\alpha\theta-\delta})^{1-\alpha\theta}]\mathrm{d}x \quad \left(u=\frac{\mu+x\lambda^{\alpha\theta-\delta}}{\mu}\right)$$

$$
\begin{aligned}
&\le c_5\lambda^{-2\alpha\theta+2\delta}\mu^{2-\alpha\theta}\int_1^\infty \frac{u-1}{u^{\alpha\theta}}\exp[-c_3\lambda^{\alpha\gamma}\mu^{1-\alpha\theta}u^{1-\alpha\theta}]\mathrm{d}u\\
&\le c_6\lambda^{-2\alpha\theta+2\delta}\mu^{2-\alpha\theta}\exp[-c_3\lambda^{\alpha\gamma}\mu^{1-\alpha\theta}]\\
&\quad\times\int_1^\infty (u-1)\exp[-c_3(1-\alpha\theta)\lambda^{\alpha\gamma}\mu^{1-\alpha\theta}(u-1)]\mathrm{d}u\\
&\le c_7\lambda^{-2(\alpha\gamma+\alpha\theta-\delta)}\mu^{\alpha\theta}\exp[-c_3\lambda^{\alpha\gamma}\mu^{1-\alpha\theta}]. \qquad \text{(A.14)}
\end{aligned}
$$

Combining (A.13) and (A.14) yields that the sum under consideration is less than $c_7\mu^{\alpha\theta}\lambda^{-\alpha\gamma}$. The estimate (A.11) is obtained analogously. □

References

[1] T.-S. Chiang, C.-R. Hwang, S.-J. Sheu, Diffusion for global optimization in $\mathbb{R}^n$, SIAM Journal on Control and Optimization 25 (3) (1987) 737–753.

[2] C.-R. Hwang, S.-J. Sheu, Large-time behavior of perturbed diffusion Markov processes with applications to the second eigenvalue problem for Fokker–Planck operators and simulated annealing, Acta Applicandae Mathematicae 19 (3) (1990) 253–295.

[3] C.-R. Hwang, S.-J. Sheu, Singular perturbed Markov chains and exact behaviors of simulated annealing processes, Journal of Theoretical Probability 5 (2) (1992) 223–249.

[4] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems, 2nd edition, in: Grundlehren der Mathematischen Wissenschaften, vol. 260, Springer, New York, NY, 1998.

[5] R. Holley, D. Stroock, Simulated annealing via Sobolev inequalities, Communications in Mathematical Physics 115 (4) (1988) 553–569.

[6] R. Holley, S. Kusuoka, D. Stroock, Asymptotics of the spectral gap with applications to the theory of simulated annealing, Journal of Functional Analysis 83 (2) (1989) 333–347.

[7] B. Gidas, Metropolis-type Monte Carlo simulation algorithms and simulated annealing, in: J.L. Snell (Ed.), Topics in Contemporary Probability and Its Applications, in: Probability and Stochastics Series, CRC Press, Boca Raton, FL, 1995, pp. 159–232.

[8] G.B. Sorkin, Efficient simulated annealing on fractal energy landscapes, Algorithmica 6 (1) (1991) 367–418.

[9] L. Ingber, Simulated annealing: Practice versus theory, Mathematical and Computer Modelling 18 (11) (1993) 29–57.

[10] M. Locatelli, Simulated annealing algorithms for continuous global optimization, in: P.M. Pardalos, H.E. Romeijn (Eds.), Handbook of Global Optimization II, in: Nonconvex Optimization and its Applications, vol. 62, Kluwer Academic Publishers, Dordrecht, 2002, pp. 179–229.

[11] H. Szu, R. Hartley, Fast simulated annealing, Physics Letters A 122 (3–4) (1987) 157–162.

[12] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, E. Teller, Equation of state calculations by fast computing machines, The Journal of Chemical Physics 21 (5) (1953) 1087–1092.

[13] H. Szu, Automated fault recognition by image correlation neutral network technique, IEEE Transactions on Industrial Electronics 40 (2) (1993) 197–208.

[14] V.B. Nascimento, V.E. de Carvalho, C.M.C. de Castilho, B.V. Costa, E.A. Soares, The fast simulated annealing algorithm applied to the search problem in LEED, Surface Science 487 (2001) 15–27.

[15] P. Imkeller, I. Pavlyukevich, First exit times of SDEs driven by stable Lévy processes, Stochastic Processes and Their Applications 116 (4) (2006) 611–642.

[16] P. Imkeller, I. Pavlyukevich, Lévy flights: Transitions and meta-stability, Journal of Physics A: Mathematical and General 39 (2006) L237–L246.

[17] P. Imkeller, I. Pavlyukevich, Metastable behaviour of small noise Lévy-driven diffusions (submitted for publication) arXiv Mathematics e-prints math.PR/0601771.

[18] P.E. Protter, Stochastic Integration and Differential Equations, 2nd edition, in: Applications of Mathematics, vol. 21, Springer, Berlin, 2004.

[19] V.V. Godovanchuk, Asymptotic probabilities of large deviations due to large jumps of a Markov process, Theory of Probability and its Applications 26 (1982) 314–327.

[20] A.D. Wentzell, Limit Theorems on Large Deviations for Markov Stochastic Processes, in: Mathematics and its Applications (Soviet Series), vol. 38, Kluwer Academic Publishers, Dordrecht, 1990.

[21] J.C. Schön, Preferential trapping on energy landscapes in regions containing deep-lying minima: The reason for the success of simulated annealing? Journal of Physics A: Mathematical and General 30 (7) (1997) 2367–2389.
[22] I. Pavlyukevich, Cooling down Lévy flights (submitted for publication). arXiv e-prints cond-mat/0701651.
[23] I. Pavlyukevich, Lévy flights, non-local search and simulated annealing, Journal of Computational Physics, in press (doi:10.1016/j.jcp.2007.06.008).
[24] M.V. Kartashov, The stability of almost homogeneous in time Markov semigroups of operators, Theory of Probability and Mathematical Statistics 71 (2005) 119–128.
[25] F.W.J. Olver, Asymptotics and special functions, in: Computer Science and Applied Mathematics, Academic Press (a subsidiary of Harcourt Brace Jovanovich Publishers), New York, London, 1974.