The Annals of Applied Probability
2022, Vol. 32, No. 6, 4611–4641
https://doi.org/10.1214/22-AAP1797

LARGE DEVIATIONS FOR MARKOV JUMP PROCESSES IN PERIODIC AND LOCALLY PERIODIC ENVIRONMENTS

BY ANDREY PIATNITSKI[1,a], SERGEY PIROGOV[2,b] AND ELENA ZHIZHINA[2,c]

[1]*The Arctic University of Norway, UiT, Campus Narvik,* [a]*apiatnitski@gmail.com*
[2]*Institute for Information Transmission Problems, RAS,* [b]*s.a.pirogov@bk.ru,* [c]*ejj@iitp.ru*

The paper deals with a family of jump Markov process defined in a medium with a periodic or locally periodic microstructure. We assume that the generator of the process is a zero order convolution type operator with rapidly oscillating locally periodic coefficient and, under natural ellipticity and localization conditions, show that the family satisfies the large deviation principle in the path space equipped with Skorokhod topology. The corresponding rate function is defined in terms of a family of auxiliary periodic spectral problems. It is shown that the corresponding Lagrangian is a convex function of velocity that has a superlinear growth at infinity. However, neither the Lagrangian nor the corresponding Hamiltonian need not be strictly convex, we only claim their strict convexity in some neighbourhood of infinity. It then depends on the profile of the generator kernel whether the Lagrangian is strictly convex everywhere or not.

1. Introduction. The goal of this work is to show that for a family of jump Markov process defined in a d-dimensional medium with a (locally) periodic microstructure the large deviation principle holds. We assume that the generators of these processes are of the form

$$A^\varepsilon u(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} a\left(\frac{x-y}{\varepsilon}\right) \Lambda^\varepsilon(x,y)\big(u(y)-u(x)\big)\,dy, \tag{1}$$

where ε is a small positive parameter that characterizes the microscopic length scale, $a(\cdot)$ is a nonnegative integrable convolution kernel that decays super exponentially at infinity, and a positive bounded function Λ^ε represents the local characteristics of the medium. We consider both the case of a periodic function Λ^ε, and the case of a locally periodic one. In the former case, $\Lambda^\varepsilon(x,y) = \Lambda(\frac{x}{\varepsilon}, \frac{y}{\varepsilon})$, where $\Lambda(\xi,\eta)$ is a periodic function in $\mathbb{R}^{2d}$. In the latter case, $\Lambda^\varepsilon(x,y) = \Lambda(x,y,\frac{x}{\varepsilon}, \frac{y}{\varepsilon})$, where $\Lambda(x,y,\xi,\eta)$ is periodic in ξ and η.

Previously, the large deviation principle for trajectories of a diffusion process with a small diffusion coefficient has been justified in [10, 11]. It was shown that the large deviation principle holds in the space of continuous functions and that the corresponding rate function is defined as an integral along the curve of an appropriate Lagrangian. The Lagrangian is explicitly given in terms of the coefficients of the process generator.

Large deviation problem for a diffusion in environments with a periodic microstructure was studied for the first time in [1], where a pure diffusion without drift has been considered. The case of a small diffusion with a drift in locally periodic media was studied in [9]. Here the Lagrangian is defined in terms of an auxiliary PDE problem on the torus.

Large deviation problems for jump processes with independent increments have been investigated in [2, 16, 17, 19] and other works. In [2] the author considered the one-dimensional case. The LDP was obtained in the Skorokhod space with a weak topology under the Cramer condition on the convolution kernel. These results were improved in [16, 17], where the LDP

Received September 2021.
MSC2020 subject classifications. Primary 60F10, 60J76; secondary 47A10.
Key words and phrases. Large deviations, jump Markov processes, locally periodic microstructure, spectral properties of convolution type operators.

was proved in the Skorokhod space with strong topology and the topology of uniform convergence. In the multidimensional case similar results were obtained in [19].

A number of interesting results on large deviations for Markov processes that combine a diffusive behaviour and many small jumps can be found in [20].

The monograph [6] focuses on LDP for rather general classes of Markov processes in metric spaces. The approaches developed in this book rely on exponential tightness, convergence of nonlinear contraction semigroups and theory of viscosity solutions of nonlinear equations. In particular, this allows to consider the case of processes whose rate function might be finite for sample paths with discontinuities.

A rather general approach to obtaining the LDP for Markov processes with a quasi-compact Markov generator has been developed in [13]. This approach was developed further in the works [7, 8], where the higher order large deviation asymptotics were obtained. Notice however that, at least for a certain class of kernels, this approach does not apply to the processes studied in the present paper. In particular, the operators obtained by the exponential transformation need not be quasi-compact and might have only a continuous spectrum.

To our best knowledge, large deviation problems for jump Markov processes with convolution type generators in periodic environments have not been studied in the existing literature.

In the present paper we consider a family of jump Makov processes $\xi^\varepsilon(t)$, $0 \le t \le T$, with the generator defined in (1). Under the assumptions that the convolution kernel $a(\cdot)$ is integrable and decays super exponentially at infinity, and that the function Λ^ε is strictly positive, bounded and has a periodic or locally periodic microstructure we prove that the family $\{\xi^\varepsilon(t)\}$ satisfies the large deviation principle in the Skorokhod space $\mathbf{D}([0,T];\mathbb{R}^d)$ equipped with the strong topology. The corresponding rate function is good, it is finite only for absolutely continuous functions and is given by

$$I(\gamma(\cdot)) = \int_0^T L(\gamma(t), \dot\gamma(t))\,dt,$$

where the Lagrangian $L(x,\zeta)$ is convex and has a super linear growth as a function of ζ while in x it is continuous. This Lagrangian is constructed in terms of a family of auxiliary periodic spectral problems on the torus for operators which are derived from the generator of the process by the exponential transformation.

Our strategy is natural for a family of Markov process. We approximate trajectories of the process by piece-wise linear functions and, for each segment of this function, apply finite-dimensional techniques. This leads to the said family of spectral problems which is of special interest.

Let us recall that in the case of diffusion processes with a small diffusion in an environment with a periodic microstructure the mentioned exponential transformation leads to a family of elliptic operators with a compact resolvent defined on the torus. The Krein–Rutman theorem applies to these operators, and the upper border of their spectrum coincides with the principal eigenvalue. Then, using for example, the arguments of the homogenization theory, one can show that this eigenvalue is a strictly convex function of the parameter. This implies that the corresponding Lagrangian is strictly convex and the Gärtner–Ellis theorem applies. In contrast with the case of diffusion processes for the jump Markov processes considered in the present paper the exponential transformation leads to a family of bounded operators which, in a typical situation, have a nontrivial continuous spectrum. The essential spectrum of these operators is real and does not depend on the transformation parameter. The Krein–Rutman theorem in this case does not apply, at least directly, and we need more delicate techniques to study the spectral properties of this family. It turns out that the discrete spectrum of these operators need not be real and might be empty. However, the eigenvalue with the largest real part, if it exists, is real and simple. It is shown that for large values of the parameter of exponential transformation the discrete spectrum is not empty and that the upper border of the spectrum coincides with the principal eigenvalue. Using the arguments of homogenization theory for jump Markov processes (see [18]) one can prove that this eigenvalue is a smooth strictly convex function of the parameter. This implies that in the vicinity of infinity both the Hamiltonian and the Lagrangian are strictly convex and smooth functions.

However, in the complement to the mentioned neighbourhood of infinity there might be a region where the discrete spectrum is empty. In this region the upper bound of the spectrum coincides with the border of the continuous spectrum. Then the Hamiltonian is equal to a constant, while at the border of this region it need not be differentiable so that the graph of the corresponding Lagrangian might belong to a conical surface in a neighbourhood of the origin. This is one of the interesting features of the studied problem. The structure of the Lagrangian is studied in detail in Sections 4.3–4.5.

A possible behaviour of the Hamiltonian and Lagrangian is illustrated with two examples. The first one shows that the Hamiltonian indeed might have a flat area with a nonempty interior.

In the second example we construct an operator for which the Hamiltonian is not C^1 function, and the corresponding Lagrangian is not strictly convex.

Due to the lack of strict convexity the Gärtner–Ellis theorem does not give a complete picture for finite-dimensional distributions. However, the results on the large deviations asymptotitcs remain valid. Our arguments rely on the particular structure of the Hamiltonian and the Lagrangian and on the Markov property of the process.

The paper is organized as follows. In Section 2 we introduce the studied family of jump Markov processes, provide all our assumptions and formulate the main result.

In Section 3 we recall some of the existing large deviation results for jump process with independent increments.

The case of purely periodic environment is considered in Section 4. First we introduce a family of auxiliary operators with periodic coefficients, consider the corresponding spectral problems on the torus and study the structure of their spectrum. Then we define the Hamiltonian and the Lagrangian that are required for formulating the large deviation results, and investigate their properties. In the last part of this section we formulate and prove the large deviation theorems, first for the distribution of the process in $\mathbb{R}^d$ at a fixed time, and then in the path space.

Section 5 deals with the media that do not depend on fast variables. Here we combine the results obtained for the processes with independent increments and perturbation theory arguments. Although this idea is very natural and not new, its realization requires a number of quite delicate technical statements.

Finally, in Section 6 we consider the generic case of locally periodic media.

2. Problem setup and main result. We consider a family of continuous time jump Markov processes $\xi^\varepsilon_{x_0}(t)$ in environments with locally periodic microstructure that depend on a small parameter $\varepsilon > 0$; the subindex x_0 indicates the starting point: $\xi^\varepsilon_{x_0}(0) = x_0$. The generator of this process has the form

$$A^\varepsilon u(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} a\Big(\frac{x-y}{\varepsilon}\Big) \Lambda\Big(x, y, \frac{x}{\varepsilon}, \frac{y}{\varepsilon}\Big)\big(u(y) - u(x)\big)\, dy, \tag{2}$$

$u \in L^2(\mathbb{R}^d)$. We call x, y slow variables and $\frac{x}{\varepsilon}, \frac{y}{\varepsilon}$ fast variables.

Our goal is to show that, under proper ellipticity and exponential moment conditions, the large deviation principle holds for this family of Markov processes. In this section we introduce these conditions.

For the function $a(z)$ we assume that

$$a(z) \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d), \quad a(z) \ge 0, \qquad \|a\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} a(z)\,dz = 1, \tag{3}$$

and the convolution kernel $a(z)$ satisfies the following upper bound with some $p > 1$, $k > 0$, $\mathrm{C} > 0$:

$$0 \le a(z) \le \mathrm{C} e^{-k|z|^p}. \tag{4}$$

The latter condition implies in particular that all exponential moments are bounded.

We assume furthermore that for all α from the unit sphere S^{d-1} we have

$$\int_{\Pi_\alpha} a(z)\,dz > 0 \quad \text{with } \Pi_\alpha = \{z \in \mathbb{R}^d, z \cdot \alpha > 0\}. \tag{5}$$

Observe, that the integral $\int_{\Pi_\alpha} a(z)\,dz$ is a continuous function of $\alpha \in S^{d-1}$ and, therefore,

$$\min_{\alpha \in S^{d-1}} \int_{\Pi_\alpha} a(z)\,dz \ge C_0 \tag{6}$$

for some $C_0 > 0$.

The function $\Lambda(x, y, \xi, \eta)$ describes the locally periodic environment. We assume that the function Λ is periodic in ξ and η,

$$\begin{gathered} \Lambda(x, y, \xi + j', \eta + j'') = \Lambda(x, y, \xi, \eta) \\ \text{for all } j', j'' \in \mathbb{Z}^d \text{ and for all } x, y, \xi, \eta \in \mathbb{R}^d, \end{gathered} \tag{7}$$

and that

$$\begin{aligned} &\Lambda(x, y, \xi, \eta) \\ &\quad \text{is uniformly continuous in } x \text{ and } y \text{and measurable in } (\xi, \eta) \text{ for each } x \text{ and } y. \end{aligned} \tag{8}$$

We assume furthermore that Λ is bounded from above and from below:

$$0 < \Lambda^- \le \Lambda(x, y, \xi, \eta) \le \Lambda^+ < \infty. \tag{9}$$

At the end of this section we formulate our main result in a vague form, for the detailed version of this theorem see Sections 4–6.

THEOREM 2.1. *Under assumptions* (3)–(5) *and* (7)–(9) *for any $T > 0$ the family of processes $\xi^\varepsilon_{x_0}(\cdot)$ satisfies the large deviation principle in the Skorokhod space $\mathbb{D}[0, T]$ with the rate function*

$$I(\gamma(\cdot)) = \begin{cases} \displaystyle\int_0^T L(\gamma(t), \dot\gamma(t))\,dt & \text{if } \gamma \text{ is absolutely continuous}, \gamma(0) = x_0, \\ +\infty & \text{otherwise}; \end{cases}$$

where the Lagrangian $L(x, \zeta)$ possesses the following properties:

- *$L(x, \zeta)$ is a continuous function of x and ζ*;
- *for each x the function $L(x, \cdot)$ is convex*;
- *$|\zeta|^{-1} L(x, \zeta)$ tends to ∞ as $|\zeta| \to \infty$ uniformly in x*;
- *there exists $R_0 > 0$ such that in the set $|\zeta| \ge R_0$ the function $L(x, \zeta)$ is strictly convex in ζ.*

Notice that the properties of the Lagrangian listed in the theorem ensure that the rate function $I(\cdot)$ is good that is, for any $T > 0$ and any $c \ge 0$ the sub-level set $\{\gamma \in \mathbb{D}[0, T] : I(\gamma) \le c\}$ is compact.

REMARK. The effective drift of the jump Markov process governed by the nonlocal operator A^ε that was defined in (2) is of order one, while the asymptotic diffusion of this process is small, of order ε. Hence, this process can be interpreted as a small random perturbation of a dynamical system.

REMARK. The key results of this work are those obtained for a periodic medium. Here we introduce the corresponding Hamiltonian and Lagrangian and study their properties. In particular we show that, in contrast with a homogeneous medium, both the Hamiltonian and the Lagrangian need not be smooth, neither strictly convex. The locally periodic case is then reduced to the periodic one with the help of perturbation theory arguments.

3. Processes with independent increments. We start with the case of constant Λ: $\Lambda^\varepsilon(x, y) \equiv \Lambda$. In this case $\xi_x^\varepsilon(\cdot)$ is a continuous time process with independent increments, or equivalently a compound Poisson process. The results on large deviations under condition (4) are well known; see for example, [2]. In [15–17] the authors considered a wider class of the compound Poisson processes that have exponential moments only in a neighborhood of zero. Let us shortly repeat the construction of the rate function and the Lagrangian for this process.

In this section the dependence of Λ is indicated explicitly, $\xi_{x,\Lambda}^\varepsilon(t)$ stands for a continuous time process with independent increments whose generator is defined by

$$A_\Lambda^\varepsilon u(x) = \frac{\Lambda}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} a\left(\frac{x-y}{\varepsilon}\right)\big(u(y) - u(x)\big)\,dy, \quad u \in L^2\big(\mathbb{R}^d\big). \tag{10}$$

To apply the Gärtner–Ellis theorem we consider the family of probability measures $\mu_{\Lambda,t}^{\varepsilon,x}$ in $\mathbb{R}^d$ defined as the law of the random variables $\xi_{x,\Lambda}^\varepsilon(t)$. In what follows we assume without loss of generality that $x = 0$ and drop the index x. We also consider the process $\xi_\Lambda(t)$ generated by

$$A_\Lambda u(x) = \Lambda \int_{\mathbb{R}^d} a(x-y)\big(u(y) - u(x)\big)\,dy, \quad u \in L^2\big(\mathbb{R}^d\big). \tag{11}$$

It is worth noticing that

$$\xi_\Lambda^\varepsilon(t) = \varepsilon \xi_\Lambda\left(\frac{t}{\varepsilon}\right).$$

We have

$$\mathbb{E}e^{\lambda \xi_\Lambda(T)} = e^{T H_\Lambda(\lambda)}$$

with

$$H_\Lambda(\lambda) = \Lambda\left(\int a(z)e^{-\lambda z}\,dz - 1\right) = \Lambda H(\lambda). \tag{12}$$

Representation (10) for the generator A_Λ^ε yields

$$\mathbb{E}e^{\frac{\lambda}{\varepsilon}\xi_\Lambda^\varepsilon(t)} = e^{\frac{t}{\varepsilon}H_\Lambda(\lambda)}. \tag{13}$$

Thus, we get

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{E}e^{\frac{\lambda}{\varepsilon}\xi_\Lambda^\varepsilon(t)} = t H_\Lambda(\lambda) = t\Lambda H(\lambda). \tag{14}$$

Relation (12) readily implies that the function $H_\Lambda(\lambda)$ is a smooth, strictly convex and of super-linear growth at infinity. Denote by $L(\zeta)$ the Legendre transform of $H(\lambda)$:

$$L(\zeta) = \sup_\lambda\big\{\lambda\zeta - H(\lambda)\big\}. \tag{15}$$

Then the function $t\Lambda L(\frac{\zeta}{t\Lambda})$ is the Legendre transform of $tH_\Lambda(\lambda)$:

$$\sup_\lambda\{\lambda\zeta - tH_\Lambda(\lambda)\} = t\Lambda \sup_\lambda\left\{\lambda\frac{\zeta}{t\Lambda} - H(\lambda)\right\} = t\Lambda L\left(\frac{\zeta}{t\Lambda}\right).$$

The function $L(\zeta)$ is nonnegative, strictly convex and finite for any $\zeta \in \mathbb{R}^d$. Consequently, by the Gärtner–Ellis theorem LDP holds in this case:

(1) for every closed set $C \subset \mathbb{R}^d$

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \mathbb{P}\big(\xi_\Lambda^\varepsilon(t) \in C\big) \le -\inf_{\zeta\in C}\left[t\Lambda L\left(\frac{\zeta}{t\Lambda}\right)\right]; \tag{16}$$

(2) for every open set $O \subset \mathbb{R}^d$

$$\liminf_{\varepsilon\to 0} \varepsilon \ln \mathbb{P}\big(\xi_\Lambda^\varepsilon(t) \in O\big) \ge -\inf_{\zeta\in O}\left[t\Lambda L\left(\frac{\zeta}{t\Lambda}\right)\right]. \tag{17}$$

REMARK. The case when $a(z)$ is a symmetric kernel, that is, $a(-z) = a(z)$, and $\Lambda(\cdot) \equiv 1$ has been studied in [12]. In particular, the large deviation result for the density $v(x,t)$ of the transition probability $\Pr(\xi(t) = x | \xi(0) = 0)$ has been proved with the rate function $\Phi(\zeta), x = \zeta t(1 + o(1)), t \to \infty$; see Theorems 3.4 and 3.8, [12]. The rate function $\Phi(\zeta)$ possesses the following properties:

$\Phi(0) = 0$, $\Phi(\zeta) > 0$ for $\zeta \neq 0$, Φ is a convex function, and

$$\Phi(\zeta) = \frac{1}{2}\big(\sigma^{-1}\zeta, \zeta\big)\big(1 + o(1)\big) \quad \text{as } |\zeta| \to 0, \tag{18}$$

where σ is the covariance matrix, $\sigma_{ij} = \int_{\mathbb{R}^d} x_i x_j a(x)\,dx$.

If the function $a(x)$ satisfies a two-sided estimate

$$C_2 e^{-b|x|^p} \le a(x) \le C_1 e^{-b|x|^p}, \quad p > 1,$$

then the following asymptotics for the rate function $\Phi(\zeta)$ holds:

$$\Phi(\zeta) = \frac{p}{p-1}\big(b(p-1)\big)^{1/p}|\zeta|\big(\ln|\zeta|\big)^{\frac{p-1}{p}}\big(1 + o(1)\big) \quad \text{as } |\zeta| \to \infty. \tag{19}$$

Relation (19) has an important consequence that will be used in the following sections. Namely, under condition (4), there exists a constant $c_0 = c_0(C, p, d)$ such that for all sufficiently large ζ the inequality

$$\Phi(\zeta) \ge c_0|\zeta|\big(\ln|\zeta|\big)^{\frac{p-1}{p}} \tag{20}$$

holds true.

Finally, we turn to the sample path large deviations results. Denote by $\mathbf{P}^\varepsilon$ the distribution of paths of the process $\xi_\Lambda^\varepsilon(t), 0 \le t \le T$, in the space $\mathbf{D}([0,T]; \mathbb{R}^d)$. This space is equipped with the metric

$$\operatorname{dist}(f, g) = \inf_{\pi(\cdot)} \max\left\{\sup_{0\le s<t\le T}\left|\log\left(\frac{\pi(t) - \pi(s)}{t - s}\right)\right|, \sup_{0\le t\le T}\big|f(t) - g\big(\pi(t)\big)\big|\right\},$$

where the infimum is taken over all continuous strictly monotone functions π such that $\pi(0) = 0$ and $\pi(T) = T$. In what follows this set of functions is denoted by $\mathcal{K}$, and $\ell(\pi) = \sup_{0\le s<t\le T}|\log(\frac{\pi(t)-\pi(s)}{t-s})|$.

In the case of the studied process with independent increments the large deviation principle (LDP) is valid for the family of probability measures $\{\mathbf{P}^\varepsilon\}$ in the Skorokhod space equipped with topology generated by the above introduced metric, the rate function being given by

$$I_\Lambda(\gamma(\cdot)) = \begin{cases} \int_0^T \Lambda L\left(\frac{1}{\Lambda}\dot{\gamma}(t)\right) dt & \text{if } \gamma(\cdot) \text{is absolutely continuous,} \\ +\infty & \text{otherwise,} \end{cases}$$

with $L(\cdot)$ defined in (15). This means that

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \mathbf{P}^\varepsilon(C) \le - \inf_{\gamma\in C}[I_\Lambda(\gamma)] \tag{21}$$

for every closed set C in $\mathbf{D}([0,T];\mathbb{R}^d)$, and

$$\liminf_{\varepsilon\to 0} \varepsilon \ln \mathbf{P}^\varepsilon(O) \ge - \inf_{\gamma\in O}[I_\Lambda(\gamma)] \tag{22}$$

for every open set O in $\mathbf{D}([0,T];\mathbb{R}^d)$.

As a consequence, for a small neighbourhood U of a curve γ we have

$$\varepsilon \ln \mathbf{P}^\varepsilon(U) \sim -I_\Lambda(\gamma) \quad \text{as } \varepsilon \to 0. \tag{23}$$

In the one-dimensional case this result was proved, under slightly weaker assumptions, by A. Mogulskii in [16, 17], and then in the multidimensional case by A. Pukhalskii in [19].

4. Environment with periodic microstructure $\boldsymbol{\Lambda(\frac{x}{\varepsilon}, \frac{y}{\varepsilon})}$.

In this section we consider the process with generator given by (2) with $\Lambda = \Lambda(\frac{x}{\varepsilon}, \frac{y}{\varepsilon})$, where $\Lambda(\eta,\zeta)$ is a measurable periodic function satisfying the lower and upper bounds in (9).

4.1. *Skewed generator.*

Consider an operator

$$A_0 u(x) = \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)u(y)\,dy - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy\,u(x), \tag{24}$$

where $\Lambda(x,y)$ is a periodic function satisfying bound (9), and $u \in L^2(\mathbb{R}^d)$. Denote by $S(t) = e^{tA_0}$ the Markov semigroup with generator A_0, and let $\xi_x(t)$ be the corresponding continuous time jump Markov process starting at x. Then

$$(S(t)f)(x) = e^{tA_0} f(x) = \mathbb{E} f(\xi_x(t)). \tag{25}$$

LEMMA 4.1. *For any $\lambda \in \mathbb{R}^d$ and $x \in \mathbb{R}^d$*

$$\mathbb{E} e^{\lambda \xi_x(t)} = e^{\lambda x} e^{tA_\lambda} 1, \tag{26}$$

where A_λ is the operator acting in the space of periodic functions $L^2(\mathbb{T}^d)$ and defined by

$$A_\lambda v(x) = \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda(y-x)}v(y)\,dy - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy\,v(x). \tag{27}$$

PROOF. Substitute $f(z) = e^{\lambda z}$ in (25) and denote $u(x,t) = \mathbb{E}e^{\lambda\xi_x(t)}$. Under our standing assumptions on $a(\cdot)$ the function $u(\cdot)$ is well defined. Indeed, denoting

$$\tilde{p}_x^t(y) = e^{-t\Lambda^-}\delta_x(y) + e^{-t\Lambda^-}\sum_{n=1}^\infty \frac{(\Lambda^+)^n t^n}{n!} a^{\star n}(x-y)$$

with Λ^- and Λ^+ defined in (9) we have

$$p_x^t(y) \le \tilde{p}_x^t(y),$$

where $p_x^t(\cdot) = e^{tA_0}\delta_x(\cdot)$ is the distribution of the process $\xi_x(t)$. Considering (20), in the same way as in [21], one can show that $\tilde p_x^t(y)$ does not exceed $e^{-c\frac{|x-y|}{t}(\ln\frac{|x-y|}{t})^{\frac{p-1}{p}}}$ for some $c>0$ and for all y such that $|x-y|\ge (1\vee t)$. Consequently, the integral $\int_{\mathbb{R}^d} e^{\lambda y} p_x^t(y)\,dy$ converges for any $t>0$ and $\lambda\in\mathbb{R}^d$, and the function

$$u(x,t) = \int_{\mathbb{R}^d} e^{\lambda y} p_x^t(y)\,dy, \tag{28}$$

is well defined. Moreover, due to periodicity of $\Lambda(x,y)$,

$$v(x,t) = e^{-\lambda x}u(x,t) = \int_{\mathbb{R}^d} e^{\lambda(y-x)} p_x^t(y)\,dy = B_\lambda^{-1} e^{tA_0} B_\lambda 1 \tag{29}$$

is a periodic function of x, that is, $v(\cdot,t)\in L^2(\mathbb{T}^d)$ for any $t>0$; here $B_\lambda g(x) = e^{\lambda x} g(x)$. In fact, under our assumptions $v(\cdot,t)\in L^\infty(\mathbb{T}^d)$. Since $A_\lambda = B_\lambda^{-1} A_0 B_\lambda$, where A_λ is defined by (27), we have $B_\lambda^{-1} e^{tA_0} B_\lambda = e^{tA_\lambda}$. This yields (26). $\square$

Consequently, for any $t>0$, we have

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{E} e^{\frac{\lambda}{\varepsilon}\xi_0^\varepsilon(t)} = \lim_{\varepsilon\to 0} \varepsilon \ln\big([e^{tA^\varepsilon_{\lambda/\varepsilon}}1](0)\big) = \lim_{s\to+\infty} \frac{1}{s}\ln\big([e^{tsA_\lambda}1](0)\big), \tag{30}$$

where $A^\varepsilon_{\lambda/\varepsilon} = B^{-1}_{\lambda/\varepsilon} A^\varepsilon B_{\lambda/\varepsilon}$. It is straightforward to check that for any $\lambda\in\mathbb{R}^d$ the skewed operator A_λ is bounded in $L^2(\mathbb{T}^d)$. Denote by $\sigma(A_\lambda)$ the spectrum of this operator in $L^2(\mathbb{T}^d)$, and by $\mathrm{s}(A_\lambda)$ the maximum of the real parts of the elements of $\sigma(A_\lambda)$:

$$\mathrm{s}(A_\lambda) = \max\{\mathcal{R}e(\theta) : \theta\in\sigma(A_\lambda)\}. \tag{31}$$

In the next subsection we will show that the limit on the right-hand side of (30) exists and is equal to $\mathrm{s}(A_\lambda)$ multiplied by t. Our goal is to study the properties of $\mathrm{s}(A_\lambda)$ as a function of λ.

4.2. *The spectral properties of the operator A_λ.* The operator A_λ defined by (27) has a continuous spectrum

$$\sigma_{\mathrm{cont}} = [-g_{\max}, -g_{\min}] := \mathrm{Im}\{-G(x)\}, \quad x\in\mathbb{T}^d,$$

if the function

$$G(x) = \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy$$

is not a constant. Letting

$$g_{\max} = \max_{x\in\mathbb{T}^d} G(x), \qquad g_{\min} = \min_{x\in\mathbb{T}^d} G(x),$$

we have $0 < g_{\min} \le g_{\max} < \infty$. The continuous spectrum, if it exists, does not depend on λ. In addition, depending on the value of λ, A_λ might have a discrete spectrum $\sigma_{\mathrm{disc}}(\lambda)$.

Adding to both sides of the spectral problem $A_\lambda v = \theta v$ the constant $g_{\max}$ we obtain an equivalent spectral problem that reads $(A_\lambda + g_{\max})v = (\theta + g_{\max})v$. We denote the new spectral parameter $(\theta+g_{\max})$ by ϑ. The operator on the left-hand side of the latter spectral problem is positive, its essential spectrum coincides with its continuous spectrum and is equal to the real interval $[0, g_{\max}-g_{\min}]$. According to [4], Theorem 1, there are only two options. Namely, either for any $\vartheta\in\sigma(A_\lambda+g_{\max})$ we have $|\vartheta|\le g_{\max}-g_{\min}$, or there exists a real positive eigenvalue $\vartheta(\lambda)$ of $A_\lambda+g_{\max}$ such that $\vartheta(\lambda) > |\tilde\vartheta|$ for any $\tilde\vartheta\in\sigma(A_\lambda+g_{\max})\setminus\vartheta(\lambda)$. In particular, in the latter case, $\vartheta(\lambda) > g_{\max}-g_{\min}$. Furthermore, there is a positive eigenfunction u_λ that corresponds to $\vartheta(\lambda)$.

As a consequence, either the element of $\sigma(A_\lambda)$ with the largest real part coincides with $-g_{\min}$, or it is equal to $\vartheta(\lambda) - g_{\max}$. The latter case takes place if and only if $\theta(\lambda) := \vartheta(\lambda) - g_{\max} > -g_{\min}$, in this case the real part of $\tilde{\theta}$ is less than $\theta(\lambda)$ for any $\tilde{\theta} \in \sigma(A_\lambda) \setminus \theta(\lambda)$. The set of $\lambda \in \mathbb{R}^d$ such that $\theta(\lambda) > -g_{\min}$ is denoted by Γ, and $\theta(\lambda)$ is called the principal eigenvalue of A_λ.

REMARK. Notice that $\theta(0) = 0$, that is, $\theta(0) > -g_{\min}$. Furthermore, $\theta(\lambda) \to \infty$ as $|\lambda| \to \infty$. Thus, $0 \in \Gamma$, and $\mathbb{R}^d \setminus \Gamma$ is a bounded set.

Assume that $\lambda \in \Gamma$. The spectral problem for A_λ reads

$$\begin{aligned}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda(y-x)}u_\lambda(y)\,dy - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy\,u_\lambda(x)\\ = \theta(\lambda)u_\lambda(x),\end{aligned} \tag{32}$$

where $u_\lambda(x)$ is the principal eigenfunction. Denote by $u_\lambda^\star(x)$ the principal eigenfunction of the adjoint operator $A_\lambda^\star$. For $\theta(\lambda) > -g_{\min}$, the spectral problem (32) is equivalent to the following problem:

$$D_\lambda u(x) = \big(G(x)+\theta(\lambda)\big)^{-1}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda(y-x)}u_\lambda(y)\,dy = u_\lambda(x)$$

for the compact positive operator D_λ in $L^2(\mathbb{T}^d)$. Since $\theta(\lambda)$ is an eigenvalue for A_λ, 1 is an eigenvalue for D_λ.

For an arbitrary $N \in \mathbb{Z}^+$ denote by $\beta_N(x,y)$ the kernel of the operator D_λ^N:

$$D_\lambda^N v(x) = \int_{\mathbb{T}^d} \beta_N(x,y)v(y)\,dy. \tag{33}$$

Then

$$\beta_1(x,y) = \big(G(x)+\theta(\lambda)\big)^{-1}\sum_{j\in\mathbb{Z}^d} a(x-y+j)\Lambda(x,y)e^{\lambda(y-x-j)},$$

and

$$\beta_{N+1}(x,y) = \int_{\mathbb{T}^d} \beta_1(x,z)\beta_N(z,y)\,dz.$$

We claim that there exist $N \in \mathbb{Z}^+$ and constants $\beta^- > 0$ and β^+ such that

$$\beta^- \le \beta_N(x,y) \le \beta^+ \quad \text{for all } x,y \in \mathbb{T}^d. \tag{34}$$

The upper bound is evident, it is a direct consequence of our assumptions and holds for any $N > 0$; of course, β^+ might depend on N. The lower bound is less evident. The function $\sum_{j\in\mathbb{Z}^d} a(x-y+j)$ might be equal to zero for all x from a set of positive measure on the torus. Thus, the lower bound in (33) need not hold for $N = 1$. However, as was proved for instance, in [18], Lemma 4.1, there exists $N \in \mathbb{Z}^+$ such that the Nth convolution of a denoted by a^{*N} satisfies the estimate $a^{*N} \ge \beta^-$ for some $\beta^- > 0$. Since under our assumptions $\beta_N(x,y) \ge c_N a^{*N}(x-y)$ with $c_N > 0$, the desired lower bound follows.

Recalling that u_λ is positive, by the Krein–Rutman theorem (see, e.g., [14], Section 6, Proposition β‘), 1 is the principal eigenvalue of D_λ, and this eigenvalue is simple. Then $\theta(\lambda)$ is also simple.

From (34) it readily follows that both for $u_\lambda(x)$ and for $u_\lambda^\star(x)$ the following bounds hold:

$$c^- \le u_\lambda(x) \le c^+ \quad \text{and} \quad c^- \le u_\lambda^\star(x) \le c^+ \quad \text{for all } x \in \mathbb{T}^d \tag{35}$$

for some constants $c^- > 0$ and c^+. In what follows we assume the following normalization conditions to hold:

$$\int_{\mathbb{T}^d} u_\lambda(x)\,dx = 1, \qquad \int_{\mathbb{T}^d} u_\lambda(x)u_\lambda^\star(x)\,dx = 1. \tag{36}$$

We now turn to relation (30).

LEMMA 4.2. *The limit on the right-hand side of* (30) *exists and is equal to* $t\,\mathrm{s}(A_\lambda)$.

PROOF. According to [5], Corollary IV.2.4, the following relation holds:

$$\lim_{s\to+\infty} \frac{1}{s} \ln\|e^{tsA_\lambda}\|_{\mathcal{L}(L^2(\mathbb{T}^d),L^2(\mathbb{T}^d))} = t\,\mathrm{s}(A_\lambda).$$

This readily yields an upper bound

$$\limsup_{s\to+\infty} \frac{1}{s} \ln([e^{tsA_\lambda}1](0)) \le t\,\mathrm{s}(A_\lambda).$$

To obtain the lower bound we consider separately the cases $\lambda \in \Gamma$ and $\lambda \in \mathbb{R}^d \setminus \Gamma$. If $\lambda \in \Gamma$, then $\mathrm{s}(A_\lambda) = \theta(\lambda)$, and the inequality

$$\liminf_{s\to+\infty} \frac{1}{s} \ln([e^{tsA_\lambda}1](0)) \ge t\,\mathrm{s}(A_\lambda)$$

follows from the fact that u_λ is positive and that e^{tsA_λ} is a positive operator.

If $\lambda \in \mathbb{R}^d \setminus \Gamma$ then $\mathrm{s}(A_\lambda) = -g_{\min}$. Consider an auxiliary semigroup with the generator $(\mathcal{G}u)(x) = -G(x)u(x)$. It is straightforward to check that

$$\lim_{s\to+\infty} \frac{1}{s} \ln([e^{ts\mathcal{G}}1](0)) = -tg_{\min}.$$

Since the operator $A_\lambda - \mathcal{G} = (A_\lambda + g_{\max}) - (\mathcal{G} + g_{\max})$ is positive, the operator $e^{stA_\lambda} - e^{st\mathcal{G}}$ is also positive, and we conclude that

$$\liminf_{s\to+\infty} \frac{1}{s} \ln([e^{tsA_\lambda}1](0)) \ge -tg_{\min}.$$

This completes the proof. □

Our next statement describes the behaviour of $\theta(\lambda)$ at infinity.

LEMMA 4.3. *There exists* $R_0 > 0$ *such that* $\mathrm{s}(A_\lambda) > -g_{\min}$ *for all* λ *with* $|\lambda| \ge R_0$. *Moreover, there exist constants* $c_e > 0$, $c_a > 0$ *and* C_s *such that*

$$\theta(\lambda) \ge c_a e^{c_e|\lambda|} - C_s$$

for all $\lambda \in \{\lambda \in \mathbb{R}^d : |\lambda| \ge R_0\}$.

PROOF. It follows from (3) and (6) that for any $\alpha \in S^{d-1}$ there exist a ball $Q^\alpha \subset \Pi_\alpha$ such that

$$c_1^\alpha := \mathrm{dist}(Q^\alpha, \partial\Pi_\alpha) > 0 \quad \text{and} \quad c_2^\alpha := \int_{Q^\alpha} a(-z)\,dz > 0.$$

Then, for $\lambda = r\alpha$ with $r > 0$ we have

$$\int_{\mathbb{R}^d} a(x-y)e^{\lambda\cdot(y-x)}\Lambda(x,y)\,dy \ge \Lambda^- c_2^\alpha e^{c_1^\alpha r} = \Lambda^- c_2^\alpha e^{c_1^\alpha|\lambda|}.$$

By the continuity argument,

$$\int_{\mathbb{R}^d} a(x-y)e^{\lambda\cdot(y-x)}\Lambda(x,y)\,dy \ge \Lambda^- c_2^\alpha e^{\frac{1}{2}c_1^\alpha|\lambda|}$$

if $\frac{\lambda}{|\lambda|}$ belongs to a sufficiently small neighbourhood of α. Due to the compactness of S^{d-1} this implies that for some $c_a > 0$ and $c_e > 0$ the inequality

$$\int_{\mathbb{R}^d} a(x-y)e^{\lambda\cdot(y-x)}\Lambda(x,y)\,dy \ge c_a e^{c_e|\lambda|}$$

holds for all $\lambda \in \mathbb{R}^d$. Therefore, $[(A_\lambda + g_{\max})1](x) \ge c_a e^{c_e|\lambda|}$. Since the operator $A_\lambda + g_{\max}$ is positive, this yields $[(A_\lambda + g_{\max})^n 1](x) \ge c_a^n e^{nc_e|\lambda|}$ for any $n \in \mathbb{Z}^+$, and we conclude that $\vartheta(\lambda) \ge c_a e^{c_e|\lambda|}$, and $\theta(\lambda) \ge c_a e^{c_e|\lambda|} - g_{\max}$. □

4.3. *Strict convexity of the principal eigenvalue $\theta(\lambda)$ of the operator A_λ.*

THEOREM 4.4. *The function $\theta(\lambda)$ is strictly convex on Γ, that is, $\frac{\partial^2\theta}{\partial\lambda_i\,\partial\lambda_j}(\lambda)$ is a positive definite matrix for all $\lambda \in \Gamma$.*

PROOF. We are going to show that the matrix $\nabla\nabla\theta(\lambda_0)$ coincides with an effective diffusion matrix for a family of convolution type operators with periodic coefficients.

Let us start with the case $\lambda_0 = 0$. Then $\theta(0) = 0$, and the principal eigenfunction $u_0(x) \equiv 1$. Differentiating equality (32) in λ_i, $i = 1, \dots, d$, yields

$$\begin{aligned}(37)\qquad &\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i - x_i)e^{\lambda(y-x)}u_\lambda(y)\,dy\\ &\quad + \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda(y-x)}\partial_{\lambda_i}u_\lambda(y)\,dy\\ &\quad - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy\,\partial_{\lambda_i}u_\lambda(x) = (\partial_{\lambda_i}\theta(\lambda))u_\lambda(x) + \theta(\lambda)(\partial_{\lambda_i}u_\lambda(x)).\end{aligned}$$

Relation (37) can be rearranged as follows:

$$\begin{aligned}(38)\qquad &\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda(y-x)}\partial_{\lambda_i}u_\lambda(y)\,dy\\ &\quad - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy\,\partial_{\lambda_i}u_\lambda(x) - \theta(\lambda)\partial_{\lambda_i}u_\lambda(x)\\ &\quad = -\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i - x_i)e^{\lambda(y-x)}u_\lambda(y)\,dy + (\partial_{\lambda_i}\theta(\lambda))u_\lambda(x).\end{aligned}$$

The solvability condition for (38) reads

$$\begin{aligned}(39)\qquad &\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i - x_i)e^{\lambda(y-x)}u_\lambda(y)u_\lambda^\star(x)\,dy\,dx\\ &\quad = \partial_{\lambda_i}\theta(\lambda)\int_{\mathbb{T}^d} u_\lambda(x)u_\lambda^\star(x)\,dx = \partial_{\lambda_i}\theta(\lambda).\end{aligned}$$

Differentiating (37) one more time in λ_j yields

$$
\begin{aligned}
&\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)(y_j-x_j)e^{\lambda(y-x)}u_\lambda(y)\,dy \\
&\quad+\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)e^{\lambda(y-x)}\partial_{\lambda_j}u_\lambda(y)\,dy \\
&\quad+\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_j-x_j)e^{\lambda(y-x)}\partial_{\lambda_i}u_\lambda(y)\,dy \\
&\quad+\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda(y-x)}\partial_{\lambda_i}\partial_{\lambda_j}u_\lambda(y)\,dy \\
&\quad-\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy\,\partial_{\lambda_i}\partial_{\lambda_j}u_\lambda(x)=\partial_{\lambda_i}\partial_{\lambda_j}\theta(\lambda)u_\lambda(x) \\
&\quad+\partial_{\lambda_i}\theta(\lambda)\partial_{\lambda_j}u_\lambda(x)+\partial_{\lambda_j}\theta(\lambda)\partial_{\lambda_i}u_\lambda(x)+\theta(\lambda)\partial_{\lambda_i}\partial_{\lambda_j}u_\lambda(x).
\end{aligned}
\tag{40}
$$

After rearranging (40) in the same way as (38) the solvability condition for (40) reads

$$
\begin{aligned}
&\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)(y_j-x_j)e^{\lambda(y-x)}u_\lambda(y)u_\lambda^\star(x)\,dy\,dx \\
&\quad+\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)e^{\lambda(y-x)}\partial_{\lambda_j}u_\lambda(y)u_\lambda^\star(x)\,dy\,dx \\
&\quad+\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_j-x_j)e^{\lambda(y-x)}\partial_{\lambda_i}u_\lambda(y)u_\lambda^\star(x)\,dy\,dx \\
&\quad-\partial_{\lambda_i}\theta(\lambda)\int_{\mathbb{T}^d}\partial_{\lambda_j}u_\lambda(x)u_\lambda^\star(x)\,dx-\partial_{\lambda_j}\theta(\lambda)\int_{\mathbb{T}^d}\partial_{\lambda_i}u_\lambda(x)u_\lambda^\star(x)\,dx \\
&=\partial_{\lambda_i}\partial_{\lambda_j}\theta(\lambda).
\end{aligned}
\tag{41}
$$

At $\lambda=0$ relation (41) takes the form

$$
\begin{aligned}
\partial_{\lambda_i}\partial_{\lambda_j}\theta(0)&=\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)(y_j-x_j)u_0^\star(x)\,dy\,dx \\
&\quad+\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)\partial_{\lambda_j}u_0(y)u_0^\star(x)\,dy\,dx \\
&\quad+\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_j-x_j)\partial_{\lambda_i}u_0(y)u_0^\star(x)\,dy\,dx \\
&\quad-\partial_{\lambda_i}\theta(0)\int_{\mathbb{T}^d}\partial_{\lambda_j}u_0(x)u_0^\star(x)\,dx-\partial_{\lambda_j}\theta(0)\int_{\mathbb{T}^d}\partial_{\lambda_i}u_0(x)u_0^\star(x)\,dx.
\end{aligned}
\tag{42}
$$

LEMMA 4.5. *The matrix $\nabla\nabla\theta(0)$ is positive definite.*

PROOF. Notice that the matrix defined on the right-hand side of (42) coincides with the symmetric part of the effective diffusion matrix

$$
\begin{aligned}
\Theta^{ij}&=\frac{1}{2}\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)(y_j-x_j)u_0^\star(x)\,dy\,dx \\
&\quad-\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(x_i-y_i)\varkappa_j(y)u_0^\star(x)\,dy\,dx+b_i\int_{\mathbb{T}^d}\varkappa_j(x)u_0^\star(x)\,dx,
\end{aligned}
\tag{43}
$$

that was constructed in [18] for the convolution type operator A_0.

Indeed, at $\lambda=0$ relation (39) takes the form

$$
\partial_{\lambda_i}\theta(0)=\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)(y_i-x_i)u_0^\star(x)\,dy\,dx,
\tag{44}
$$

where $u_0^\star$ is the eigenfunction of the adjoint operator $A_0^\star$ corresponding to the principal eigenvalue $\theta(0)=0$. Observe that the expression on the right-hand side of (44) taken with the negative sign, coincides with that for the ith coordinate of the effective drift b_i of the operator A_0; see [18]. That is,

$$\partial_{\lambda_i}\theta(0) = -b_i. \tag{45}$$

Letting $\lambda=0$ in (38), substituting (45) into (38), considering the relation $u_0(x)\equiv 1$ and recalling the equation for the corrector $\varkappa$ (see [18]), we conclude that

$$\partial_{\lambda_i} u_\lambda(x)|_{\lambda=0} = \varkappa_i(x). \tag{46}$$

Finally, by (45) and (46) we obtain $\partial_{\lambda_i}\partial_{\lambda_j}\theta(0)=\Theta^{ij}+\Theta^{ji}$. Then positive definiteness of the matrix $\nabla\nabla\theta(0)$ follows from [18], Proposition 6.1. $\square$

We turn to the case $\lambda=\lambda_0+r$ with $\lambda_0\neq 0$, $\lambda_0\in\Gamma$ and r belonging to a small neighbourhood of the origin. Then

$$\begin{aligned} A_\lambda u(x) = {} & \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda_0(y-x)}e^{r(y-x)}u(y)\,dy \\ & - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy u(x). \end{aligned} \tag{47}$$

Let us consider the operator $\tilde{A}_\lambda = R_{\lambda_0}^{-1}A_\lambda R_{\lambda_0}$, where $R_{\lambda_0}f(x)=u_{\theta(\lambda_0)}(x)f(x)$ is the operator of multiplication by the principal eigenfunction $u_{\theta(\lambda_0)}$ of the operator A_{λ_0}.

The operators A_λ and $\tilde{A}_\lambda$ are similar, thus they have the same spectrum. In particular, the spectral problem for $\tilde{A}_\lambda$ reads

$$\begin{aligned} & \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)u^{-1}_{\theta(\lambda_0)}(x)u_{\theta(\lambda_0)}(y)e^{\lambda_0(y-x)}e^{r(y-x)}v(y)\,dy \\ & \quad - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy v(x) = \theta(\lambda)v(x), \end{aligned} \tag{48}$$

where $\theta(\lambda)$ is the principal eigenvalue of A_λ. Denote

$$\theta_{\lambda_0}(r) = \theta(\lambda)-\theta(\lambda_0) \quad \text{with } r=\lambda-\lambda_0. \tag{49}$$

For $\lambda=\lambda_0$ we have, from (47),

$$\begin{aligned} A_{\lambda_0}u_{\theta(\lambda_0)}(x) = {} & \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)e^{\lambda_0(y-x)}u_{\theta(\lambda_0)}(y)\,dy \\ & - \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy u_{\theta(\lambda_0)}(x) = \theta(\lambda_0)u_{\theta(\lambda_0)}(x). \end{aligned} \tag{50}$$

Dividing this equation by $u_{\theta(\lambda_0)}(x)$ we get

$$\begin{aligned} & \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)u^{-1}_{\theta(\lambda_0)}(x)u_{\theta(\lambda_0)}(y)e^{\lambda_0(y-x)}\,dy \\ & \quad = \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy + \theta(\lambda_0). \end{aligned} \tag{51}$$

Thus (48), (49) and (51) imply

$$\begin{aligned} & \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)u^{-1}_{\theta(\lambda_0)}(x)u_{\theta(\lambda_0)}(y)e^{\lambda_0(y-x)}e^{r(y-x)}v(y)\,dy \\ & \quad = \left[\int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)\,dy + \theta(\lambda_0)\right]v(x) + \theta_{\lambda_0}(r)v(x) \\ & \quad = \int_{\mathbb{R}^d} a(x-y)\Lambda(x,y)u^{-1}_{\theta(\lambda_0)}(x)u_{\theta(\lambda_0)}(y)e^{\lambda_0(y-x)}\,dy v(x) + \theta_{\lambda_0}(r)v(x). \end{aligned} \tag{52}$$

This spectral problem is similar to that in (27), if we replace the kernel $a(x-y)\Lambda(x,y)$ with the kernel

$$a^{(\lambda_0)}(x-y)\Lambda^{(\lambda_0)}(x,y) = a(x-y)e^{\lambda_0(y-x)}\Lambda(x,y)u^{-1}_{\theta(\lambda_0)}(x)u_{\theta(\lambda_0)}(y).$$

According to (49),

$$\frac{\partial^2\theta(\lambda_0)}{\partial\lambda_i\,\partial\lambda_j} = \frac{\partial^2\theta_{\lambda_0}(0)}{\partial r_i\,\partial r_j},$$

and the desired positive definiteness follows. □

REMARK. The structure of the set $\Gamma = \{\lambda \in \mathbb{R}^d : \theta(\lambda) > -g_{\min}\}$ depends on the kernel $a(x-y)\Lambda(x,y)$ of the operator A_0. For example, if $a(-z) = a(z)$ and $\Lambda(x,y)$ is a symmetric periodic function, then $\theta(-\lambda) = \theta(\lambda)$ and $\theta(0) = 0$ is the minimum of $\theta(\lambda)$ (as a function of λ). Consequently, in this case $\Gamma = \mathbb{R}^d$ and $\theta(\lambda) \ge 0$ for all λ.

Also, $\Gamma = \mathbb{R}^d$ if $\Lambda(x,y) = \Lambda(x-y)$. In this case $g_{\min} = g_{\max}$, and the spectrum of A_λ coincides with the spectrum of the operator

$$u \mapsto \int_{\mathbb{R}^d} a(x-y)\Lambda(x-y)e^{\lambda(y-x)}u(y)\,dy, \quad u \in L^2(\mathbb{T}^d),$$

shifted by $-g_{\min}$. The latter operator is compact and maps positive functions to positive functions. One can show that the Krein–Rutman theorem applies, and thus this operator has a positive real eigenvalue. This implies the claim.

The following example illustrates that in general the set Γ need not coincide with $\mathbb{R}^d$.

EXAMPLE 1. Take $a(z) = \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]^d}$ equal to the characteristic function of the period, and $\Lambda(x,y) = b(x)\Lambda_0(x-y)$. We assume that $\Lambda_0(z)$ is a smooth periodic function, $0 < \alpha_1 \le \Lambda_0(z) \le \alpha_2 < \infty$, and Λ_0 has the form of a single peak:

$$\Lambda_0(z) = \begin{cases} \alpha_2, & |z-z_0| < \dfrac{c}{2}, \\ \alpha_1, & |z-z_0| > c. \end{cases}$$

Here $z_0 \neq 0, z \in \mathbb{T}^d$, and we choose sufficiently small constants α_1 and c and sufficiently large constant α_2 so that the following normalization condition holds:

$$\int_{\mathbb{R}^d} a(z)\Lambda_0(z)\,dz = \int_{\mathbb{T}^d} a(z)\Lambda_0(z)\,dz = 1.$$

Then the spectral problem (32) for A_λ reads

$$\begin{aligned} b(x)&\int_{\mathbb{T}^d} a(x-y)\Lambda_0(x-y)e^{\lambda(y-x)}u_\lambda(y)\,dy \\ &= b(x)\int_{\mathbb{T}^d} a(x-y)\Lambda_0(x-y)\,dy u_\lambda(x) + \theta(\lambda)u_\lambda(x), \end{aligned}$$

and, after straightforward rearrangements,

$$\frac{b(x)}{b(x)+\theta(\lambda)}\int_{\mathbb{T}^d} a(x-y)\Lambda_0(x-y)e^{\lambda(y-x)}u_\lambda(y)\,dy = u_\lambda(x), \tag{53}$$

where $u_\lambda > 0$ is the principal eigenfunction.

We now take a periodic continuous positive function $b(x)$, $0 < b_{\min} = \min_{x\in\mathbb{T}^d} b(x) \le b(x) \le 1$, such that for some $\varkappa \in (0, \frac{1}{2})$ it holds

$$\left\| \frac{b(x)}{b(x)-b_{\min}} \right\|_{L^2(\mathbb{T}^d)} < 1 + \varkappa. \tag{54}$$

Obviously, inequality (54) remains valid for $\frac{b(x)}{b(x)+\theta(\lambda)}$ with any $\theta(\lambda) > -b_{\min}$. Then the operator on the left-hand side of equation (53) is positive and compact in $L^2(\mathbb{T}^d)$. Observe that in this example $g_{\min} = b_{\min}$.

Assuming that α_1 is small enough we conclude that there exists λ_0 such that $\lambda_0 z_0 > 0$ and

$$0 < a(z)\Lambda_0(z)e^{-\lambda_0 z} < \frac{1}{2} \quad \text{for all } z \in \mathbb{T}^d. \tag{55}$$

Then from (54), (55) it follows that the $L^2(\mathbb{T}^d)$ norm of the left-hand side in (53) is strictly less than $\|u_{\lambda_0}\|_{L^2(\mathbb{T}^d)}$. Therefore, equation (53) has no positive solution $u(x) \in L^2(\mathbb{T}^d)$, and there are no points of the discrete spectrum of A_{λ_0} located above the continuous spectrum, that is,

$$\sigma_{\mathrm{disc}}(A_{\lambda_0}) \cap (-g_{\min}, +\infty) = \varnothing.$$

Observe that in this example equation (53) has no positive solutions for all λ situated in a sufficiently small neighbourhood of λ_0, thus λ_0 is an interior point of $\mathbb{R}^d \setminus \Gamma$.

REMARK. In the above example in dimensions 3 and more one can choose C^2 smooth functions $b(\cdot)$ and $\Lambda_0(\cdot)$ so that $\Lambda(x, y) = b(x)\Lambda_0(x - y)$ is also a C^2 function. Clearly, a small C^2 perturbation of the kernel $\Lambda(\cdot)$ does not change the structure of the Hamiltonian. Thus, in the class of C^2 coefficients Λ, the presence of a nonempty set $\mathbb{R}^d \setminus \Gamma$ (the flat area of the Hamiltonian) is one of the cases of the general position.

It is also important to note that the gradient of $\theta(\lambda)$ need not be equal to zero at the boundary $\partial(\mathbb{R}^d \setminus \Gamma)$. This is demonstrated in the second example.

EXAMPLE 2. In dimension 1 we consider the following functions:

$$a(z) = \begin{cases} 1 & \text{if } z \in [-1, \delta], \\ 0 & \text{otherwise,} \end{cases},$$

$$\Lambda_0(z) = \begin{cases} \dfrac{3}{1+3\delta} & \text{if } z \in \left[-1+\delta, -\dfrac{3}{4}+\delta\right], \\ \dfrac{1}{3(1+3\delta)} & \text{if } z \in \left(-\dfrac{3}{4}+\delta, \delta\right), \end{cases} \qquad \Lambda_0(\cdot) \text{is 1-periodic,}$$

$$b(x) = \min\left\{1, \left(\delta^{-1}\left|x + \frac{1}{2}\right|^{0.1} + \frac{1}{10}\right)\right\},$$

and, as in the previous example, define $\Lambda(x, y) = \Lambda_0(x - y)b(x)$; here $\delta > 0$ is a small parameter that will be chosen later on. Notice that $g_{\min} = 0.1$.

For $\lambda = -2$ and for sufficiently small $\delta > 0$ straightforward computations yield

$$\sum_{j\in\mathbb{Z}} a(x - y - j)\Lambda_0(x - y)e^{2(j+y-x)} < 0.7.$$

Using the same arguments as in the previous example we conclude that for sufficiently small $\delta > 0$ the point $\lambda = -2$ does not belong to Γ. We recall that the function $\mathtt{s}(A_\lambda)$ introduced in (31) is convex. Since $0 \in \Gamma$, then $\mathtt{s}(A_\lambda)$ is an increasing function of λ on the segment $[-2, 0]$. Consequently, there exists $\lambda_0 \in (-2, 0)$ such that $\mathtt{s}(A_\lambda) = \theta(\lambda) > -\frac{1}{10}$ if $\lambda > \lambda_0$, and $\mathtt{s}(A_\lambda) = -\frac{1}{10}$ if $-2 \le \lambda \le \lambda_0$.

For any $\lambda \in [0, \lambda_0]$ we have

$$\frac{1}{3}e^{-2} < \sum_{j\in\mathbb{Z}} a(x - y - j)\Lambda_0(x - y)e^{\lambda(j+y-x)} < 6. \tag{56}$$

Recalling (53) and the normalization $\|u_\lambda\|_{L^1(\mathbb{T}^1)} = 1$ we obtain $\frac{1}{3}e^{-2} < u_\lambda(x) < \frac{6}{b(x)-0.1}$.

We turn to u^*_λ. It is straightforward to check that the function $U^*_\lambda(x) = b(x)u^*_\lambda(x)$ satisfies the equation

$$\frac{b(x)}{b(x)+\theta(\lambda)} \int_{\mathbb{R}} a(y-x)\Lambda_0(y-x)e^{\lambda(x-y)}U^*_\lambda(y)\,dy = U^*_\lambda(x).$$

Then, by the same arguments as above we obtain

$$\frac{1}{6}e^{-2} \le \frac{U^*_\lambda(x)}{\|U^*_\lambda\|_{L^1(\mathbb{T}^1)}} \le \frac{6}{b(x)-0.1}.$$

Therefore,

$$\frac{1}{6}e^{-2} < \frac{1}{\|U^*_\lambda\|_{L^1(\mathbb{T}^1)}} \int_{-1}^0 u_\lambda(x)U^*_\lambda(x)\,dx \le \frac{1}{\|U^*_\lambda\|_{L^1(\mathbb{T}^1)}} \int_{-1}^0 u_\lambda(x)u^*_\lambda(x)\,dx = \frac{1}{\|U^*_\lambda\|_{L^1(\mathbb{T}^1)}}$$

and for sufficiently small δ

$$\frac{1}{\|U^*_\lambda\|_{L^1(\mathbb{T}^1)}} \le \frac{10}{\|U^*_\lambda\|_{L^1(\mathbb{T}^1)}} \int_{-1}^0 u_\lambda(x)U^*_\lambda(x)\,dx < \int_{-1}^0 \frac{360\,dx}{(b(x)-0.1)^2} < 400.$$

Thus, $\frac{1}{400} < \|U^*_\lambda\|_{L^1(\mathbb{T}^1)} < 6e^2$, and $\frac{1}{2400}e^{-2} < U^*_\lambda(x) < \frac{36e^2}{b(x)-0.1}$. This, in turn, implies the estimate $\frac{1}{2400}e^{-2} < u^*_\lambda(x) < \frac{360e^2}{b(x)-0.1}$. According to (39) we have

$$\begin{aligned}\frac{d}{d\lambda}\theta(\lambda) &= \int_{\mathbb{T}^1}\int_{\mathbb{R}^1} a(x-y)\Lambda_0(x-y)b(x)(y-x)e^{\lambda(y-x)}u_\lambda(y)u^\star_\lambda(x)\,dy\,dx \\ &> \frac{1}{22{,}000}e^{-6} - 1500e^4\delta\end{aligned}$$

for all $\lambda \in [\lambda_0, 0)$. If δ is small enough then $\frac{d}{d\lambda}\theta(\lambda) > 10^{-5}e^{-6}$, and the derivative $\frac{d}{d\lambda}\mathrm{s}(A_\lambda)$ has a jump at λ_0.

Then the Legendre transform of $\mathrm{s}(A_\lambda)$ is a linear function on an interval $[0, \frac{d}{d\lambda}(\lambda_0)]$.

4.4. *Properties of the Hamiltonian.* Denote

$$H(\lambda) := \mathrm{s}(A_\lambda) = \begin{cases} \theta(\lambda) & \lambda \in \Gamma, \\ -g_{\min} & \text{otherwise.}\end{cases} \tag{57}$$

As a consequence of Theorem 4.4 we have

PROPOSITION 4.6. *The function $H(\cdot)$ is convex. It is strictly convex on the set Γ. Moreover,*

$$\frac{H(\lambda)}{|\lambda|} \to +\infty \quad \textit{as } |\lambda| \to +\infty. \tag{58}$$

PROOF. The convexity and the strict convexity on Γ have been proved in Theorem 4.4. The relation in (58) follows from Lemma 4.3. □

By Lemma 4.2 we have

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{E}e^{\frac{\lambda}{\varepsilon}\xi_0^\varepsilon(t)} = tH(\lambda), \tag{59}$$

with $H(\lambda)$ defined in (57).

Concluding this subsection we summarize the properties of the function $H(\lambda)$:

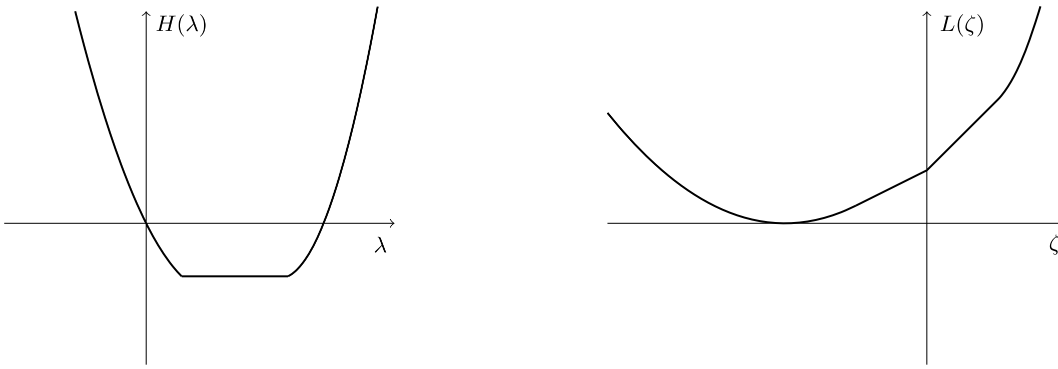

FIG. 1. *Possible graphs of the functions $H(\cdot)$ and $L(\cdot)$.*

(1) $H(\lambda)$ is convex, it is strictly convex for $\lambda \in \Gamma$,
(2) $H(0) = 0$ and $H(\lambda)$ is strictly convex at $\lambda = 0$,
(3) $\frac{H(\lambda)}{|\lambda|} \to +\infty$ as $|\lambda| \to +\infty$,
(4) the function $H(\lambda)$ equals a constant on the set $\lambda \in \Upsilon = \mathbb{R}^d \setminus \Gamma$:

$$H(\lambda) = -g_{\min}, \qquad \lambda \in \Upsilon = \mathbb{R}^d \setminus \Gamma;$$

if the set Υ is not empty, then it is bounded and convex. If the interior of Υ is not empty, then the boundary $\partial\Upsilon$ is Lipschitz continuous. A typical example of one-dimensional Hamiltonian $H(\cdot)$ with a nontrivial Υ and the corresponding Lagrangian are shown on Figure 1.

4.5. *The Legendre transform of $H(\lambda)$ and the Gärtner–Ellis theorem.* Let L and L_t be the Legendre transform of $H(\cdot)$ and $H_t := tH$, respectively, that is,

$$(60) \quad L(\zeta) = \sup_{\lambda}\bigl(\lambda\zeta - H(\lambda)\bigr), \qquad L_t(\zeta) = \sup_{\lambda}\bigl(\lambda\zeta - tH(\lambda)\bigr) = tL\left(\frac{\zeta}{t}\right), \quad \zeta \in \mathbb{R}^d.$$

We recall (see, for instance, [3]) that $\zeta' \in \mathbb{R}^d$ is an exposed point of L if for some $\theta \in \mathbb{R}^d$ and all $\zeta \neq \zeta'$,

$$\theta \cdot \zeta - L(\zeta) > \theta \cdot \zeta' - L(\zeta').$$

The properties of $H(\lambda)$ imply the following properties of $L(\zeta)$:

(1) $L(\zeta)$ is a convex function, $L(\zeta) < +\infty$ for any $\zeta \in \mathbb{R}^d$. It is strictly convex in the neighbourhood of infinity, that is, there exists R_0 such that $L(\zeta)$ is strictly convex for all ζ such that $|\zeta| \geq R_0$,
(2) $L(\zeta)$ is nonnegative: $L(\zeta) \geq 0$,
(3) $\min L(\zeta) = L(\zeta^*) = 0$ and L is strictly convex at ζ^*,
(4) $\frac{L(\zeta)}{|\zeta|} \to +\infty$ as $|\zeta| \to +\infty$, in particular, $L(\zeta)$ has compact sub-level sets,
(5) The complement to the set of exposed points of L, if not empty, consists of segments of bounded length with one end at 0, the restriction of L on each such segment is a linear function. The part of the graph of L over the set of points where L is not strictly convex is a bounded subset of a conical hypersurface centered at the origin.

Denote the set of exposed points of L by Ω. It should be emphasized that the origin need not be an exposed point of $L(\cdot)$. In particular, the restriction of L on two segments going from the origin in the opposite directions can form the same linear function. However, if $\mathbb{R}^d \setminus \Gamma$ has a nontrivial interior, then $0 \in \Omega$. This can be justified by the convex analysis arguments if we take into account the properties of $H(\cdot)$.

THEOREM 4.7. *For any $t > 0$ and any $x^0 \in \mathbb{R}^d$ the random vector $\xi^{\varepsilon}_{x^0}(t) - x_0$ satisfies the large deviation principle with the rate function $L_t(x) = tL(\frac{x}{t})$.*

PROOF. As an immediate consequence of formula (59) we obtain

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{E} e^{\frac{\lambda}{\varepsilon}(\xi^{\varepsilon}_{x^0}(t)-x^0)} = tH(\lambda). \tag{61}$$

Then the upper large deviation bound follows from the Gärtner–Ellis theorem. We have

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0} \varepsilon \log\bigl[\mathbb{P}\bigl\{\bigl|\bigl(\xi^{\varepsilon}_{x^0}(t) - x^0\bigr) - x\bigr| \le \delta\bigr\}\bigr] \le -L_t(x).$$

The lower bound is slightly more tricky. By the Gärtner–Ellis theorem for any $t > 0$ and any $x \in \mathbb{R}^d$ such that $\frac{x}{t}$ is an exposed point of $L(\cdot)$ the inequality

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0} \varepsilon \log\bigl[\mathbb{P}\bigl\{\bigl|\bigl(\xi^{\varepsilon}_{x^0}(t) - x^0\bigr) - x\bigr| \le \delta\bigr\}\bigr] \ge -L_t(x)$$

holds. Without loss of generality we assume that $x^0 = 0$. We first assume that $0 \in \Omega$. Consider $x \in \mathbb{R}^d$ which is a nonexposed point of $L_t(\cdot)$ and represent it as $x = r\phi$ with $\phi \in S^{d-1}$ and $r > 0$. Since $\xi(\cdot)$ is a Markov process, for any $\kappa \in (0, 1)$ and for any $\delta > 0$ we have

$$\begin{aligned}
&\mathbb{P}\bigl\{\bigl|\xi^{\varepsilon}_0(t) - x\bigr| \le 2\delta\bigr\}\\
&\quad = \mathbb{P}\bigl\{\bigl|\xi^{\varepsilon}_0(t) - r\phi\bigr| \le 2\delta\bigr\}\\
&\quad \ge \mathbb{P}\bigl\{\bigl\{\bigl|\xi^{\varepsilon}_0(\kappa t)\bigr| \le \delta\bigr\} \cap \bigl\{\bigl|\xi^{\varepsilon}_0(t) - \xi^{\varepsilon}_0(\kappa t) - r\phi\bigr| \le \delta\bigr\}\bigr\}\\
&\quad \ge \mathbb{P}\bigl\{\bigl\{\bigl|\xi^{\varepsilon}_0(\kappa t)\bigr| \le \delta\bigr\} \min_{|y|\le\delta} \mathbb{P}\bigl\{\bigl|\xi^{\varepsilon}_y\bigl((1-\kappa)t\bigr) - y - r\phi\bigr| \le 2\delta\bigr\}.
\end{aligned}\tag{62}$$

Denote by R the length of the segment $(0, R\phi) = (\mathbb{R}^d \setminus \Omega) \cap \{(0, s\phi) : s > 0\}$. Then, for any $h_0 > 0$, the point $(R + h_0)\phi$ is exposed for L_t. Therefore, choosing κ in (62) so that $\frac{r}{1-\kappa} = R + h_0$, that is, $\kappa = \frac{R+h_0-r}{R+h_0}$, and applying the Gärtner–Ellis theorem, we arrive for all sufficiently small $\delta > 0$ and $h_0 > 0$ at the following lower bound:

$$\begin{aligned}
&\mathbb{P}\bigl\{\bigl|\xi^{\varepsilon}_0(t) - x\bigr| \le 2\delta\bigr\}\\
&\quad \ge \exp\biggl[-\biggl(\frac{R+h_0-r}{R+h_0} L_t(0) - \psi(\delta)\biggr)\bigl(1+o(1)\bigr)\biggr]\\
&\qquad \times \exp\biggl[-\biggl(\frac{r}{R+h_0} L_t\bigl((R+h_0)\phi\bigr) - \psi(\delta)\biggr)\bigl(1+o(1)\bigr)\biggr]\\
&\quad \ge \exp\biggl[-\biggl(\frac{R-r}{R} L_t(0) + \frac{r}{R} L_t(R\phi) - C_L h_0 - 2\psi(\delta)\biggr)\bigl(1+o(1)\bigr)\biggr]\\
&\quad = \exp\bigl[-\bigl(L_t(r\phi) - C_L h_0 - 2\psi(\delta)\bigr)\bigl(1+o(1)\bigr)\bigr],
\end{aligned}$$

where $o(1)$ tends to zero as $\varepsilon \to 0$, $\psi(\delta) \to 0$ as $\delta \to 0$, and C_L is a constant which only depends on $L(\cdot)$; we have used here the fact that $L_t(\cdot)$ is linear on the segment $[0, R\phi]$. This implies the desired lower bound.

If 0 is not an exposed point then there is a segment that passes through 0, such that L_t is linear on this segment, and there are exposed points of L_t in the intersections of any neighbourhoods of the end points of this segment with the straight line that contains the segment. In this case in the same way as above one can show that

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0} \varepsilon \log\bigl[\mathbb{P}\bigl\{\bigl\{\bigl|\xi^{\varepsilon}_0(t)\bigr| \le \delta\bigr\}\bigr] \ge -L_t(0).$$

It remains to use one more time the same arguments as in the previous case to obtain the required lower bound for any $x \in \mathbb{R}^d$. This completes the proof of Theorem. $\square$

4.6. *Large deviation principle in the paths space.* The goal of this section is to show that the process $\xi_x^\varepsilon(\cdot)$ satisfies on any time interval $[0,T]$ the large deviation principle in the paths space $\mathbf{D}([0,T];\mathbb{R}^d)$ with the rate function defined by

$$I(\gamma(\cdot)) = \begin{cases} \int_0^T L(\dot\gamma(t))\,dt & \text{if } \gamma \text{ is absolutely continuous and } \gamma(0)=x, \\ +\infty & \text{otherwise,} \end{cases} \tag{63}$$

where $L(\cdot)$ is introduced in (60). An important property of $I(\cdot)$ is the compactness of its sublevel sets in the topology of uniform convergence in $C([0,T];\mathbb{R}^d)$.

LEMMA 4.8. *The set $\{\gamma \in C([0,T];\mathbb{R}^d) : I(\gamma) \le s, \gamma(0)=x\}$ is compact in $C([0,T];\mathbb{R}^d)$ for any $s \in \mathbb{R}$ and any $x \in \mathbb{R}^d$.*

PROOF. This statement is an immediate consequence of the Arzelà–Ascoli theorem and the relation $\lim_{|\zeta|\to\infty} \frac{L(\zeta)}{|\zeta|} = \infty$. $\square$

The next statement is also important for the further analysis.

PROPOSITION 4.9. *Let ξ_x^ε be a Markov process with the generator A^ε that satisfies conditions (3)–(9), and assume that $\gamma(\cdot)$ is an absolutely continuous function, $\gamma(0)=x$. Then for any $M>0$ there exists a function $\delta_0(\delta)$, $\delta_0 : (0,1] \mapsto \mathbb{R}^+$ such that $\delta_0(\delta) \to 0$ as $\delta \to 0$, and for any $\pi \in \mathcal{K}$ with $\ell(\pi) \le \delta$ we have*

$$\mathbb{P}\left\{\left\{\sup_{0\le t\le T} \left|\xi_x^\varepsilon(t) - \gamma(\pi(t))\right| \ge \delta_0\right\} \cap \left\{\left|\xi_x^\varepsilon(j\delta) - \gamma(\pi(j\delta))\right| \le \delta, j=0,\dots,\frac{T}{\delta}\right\}\right\} \\ \le \exp\left\{-\frac{M}{\varepsilon}\right\}$$

for all sufficiently small $\varepsilon>0$. Moreover, for any $s>0$ and for all sufficiently small $\varepsilon>0$,

$$\sup_{\gamma\in\Phi(s)} \mathbb{P}\left\{\left\{\sup_{0\le t\le T} \left|\xi_x^\varepsilon(t) - \gamma(\pi(t))\right| \ge \delta_0\right\} \cap \left\{\left|\xi_x^\varepsilon(j\delta) - \gamma(\pi(j\delta))\right| \le \delta, j=0,\dots,\frac{T}{\delta}\right\}\right\} \\ \le \exp\left\{-\frac{M}{\varepsilon}\right\},$$

where $\Phi(s) = \{\gamma \in \mathbf{D}([0,T],\mathbb{R}^d) : I(\gamma) \le s, \gamma(0)=x\}$.

PROOF. Consider an auxiliary process $\eta^\varepsilon(\cdot)$ with generator

$$A^\varepsilon_{\mathrm{sym}} v(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} a_s\left(\frac{x-y}{\varepsilon}\right)\left(v(y) - v(x)\right) dy,$$

where

$$a_s(z) = \mathrm{C}_0 e^{-k|z|^p}, \quad \mathrm{C}_0 = \Lambda^+ \mathrm{C},$$

with the same p, k, C and Λ^+ as those in (4) and (9). For the transition densities of the processes $\xi_x^\varepsilon(\cdot)$ and $\eta_x^\varepsilon(\cdot)$ we use the notation $q^\varepsilon(x,y,t)$ and $q_s^\varepsilon(x,y,t)$, respectively. We also define a function $q_+^\varepsilon(x,y,t)$ as the solution of the following problem:

$$\partial_t q(x,y,t) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} a_s\left(\frac{y-z}{\varepsilon}\right) q(x,z,t)\,dz, \qquad q(0,x,y) = \delta(y-x).$$

By the maximum principle we have

$$q^\varepsilon(x,y,t) \le q_+^\varepsilon(x,y,t) \quad \text{for all } x, y \in \mathbb{R}^d \text{ and all } t \ge 0. \tag{64}$$

It is also clear that

$$q_+^\varepsilon(x,y,t) = \exp\left(\frac{\mathrm{C}_1 t}{\varepsilon}\right) q_s^\varepsilon(x,y,t) \quad \text{with } \mathrm{C}_1 = \int_{\mathbb{R}^d} \mathrm{C}_0 \exp(-k|z|^p)\,dz.$$

The Hamiltonian and the Lagrangian that correspond to the process η^ε are defined in the same way as in the previous section. Namely,

$$H^s(\lambda) = \int_{\mathbb{R}^d} \mathrm{C}_0 \exp(-\lambda\cdot z - k|z|^p)\,dz - \mathrm{C}_1, \qquad L^s(\zeta) = \max_{\lambda\in\mathbb{R}^d}(\zeta\cdot\lambda - H^s(\lambda)).$$

One can easily check that both H^s and L^s are smooth strictly convex functions and, moreover, $\frac{L^s(\zeta)}{|\zeta|} \to +\infty$ as $|\zeta| \to \infty$.

Considering the continuity of $\gamma(\cdot)$ we can construct a function $\delta_0(\delta)$ such that:

- $\delta_0(\delta) \to 0$ as $\delta \to 0$.
- $|\gamma(t') - \gamma(t'')| \le \frac{1}{4}\delta_0$ if $|t' - t''| \le 3\delta$.
- $\min_{\phi\in S^{d-1}}\{\delta L^s(\frac{\delta_0\phi}{2\delta})\} \to +\infty$ as $\delta \to 0$.

LEMMA 4.10. *For any δ_0 and any $\tau > 0$ we have*

$$\mathbb{P}\Big\{\sup_{0\le t\le\tau} |\eta_x^\varepsilon(t) - x| \ge \delta_0\Big\} \le 2\mathbb{P}\{|\eta_x^\varepsilon(\tau) - x| \ge \delta_0\}.$$

PROOF. Denote by $\mathcal{E}_0$ and $\mathcal{E}_1$ the events

$$\mathcal{E}_0 = \Big\{\sup_{0\le t\le\tau} |\eta_x^\varepsilon(t) - x| \ge \delta_0\Big\}, \mathcal{E}_1 = \{|\eta_x^\varepsilon(\tau) - x| \le \delta_0\}.$$

Both $\mathcal{E}_0$ and $\mathcal{E}_1$ depend on ε, however, we do not indicate this dependence explicitly. Due to the symmetry of $a_s(\cdot)$ by the Markov property we have

$$\mathbb{P}(\mathcal{E}_0 \cap \mathcal{E}_1) = \mathbb{P}(\mathcal{E}_1|\mathcal{E}_0)\mathbb{P}(\mathcal{E}_0) < \frac{1}{2}\mathbb{P}(\mathcal{E}_0).$$

Therefore,

$$\mathbb{P}(\mathcal{E}_1^c) = \mathbb{P}(\mathcal{E}_0 \cap \mathcal{E}_1^c) > \frac{1}{2}\mathbb{P}(\mathcal{E}_0),$$

and the desired statement follows. $\square$

By the Gärtner–Ellis theorem for all sufficiently small $\varepsilon > 0$ we have

$$\mathbb{P}\{|\eta_x^\varepsilon(\delta) - x| \ge \delta_0\} \le \exp\left(-\frac{\delta}{\varepsilon}\min_{\phi\in S^{d-1}} L^s\left(\frac{\delta_0\phi}{\delta}\right)\right).$$

For arbitrary $M > 0$ we choose small enough $\delta > 0$ such that $\min_{\phi\in S^{d-1}} \delta L^s(\frac{\delta_0(\delta)\phi}{2\delta}) \ge 2M$. Then, for sufficiently small $\varepsilon > 0$ and for any $\pi \in \mathcal{K}$ with $\ell(\pi) \le \delta$,

$$\begin{aligned}
&\mathbb{P}\Big\{\Big\{\sup_{0\le t\le T} |\eta_x^\varepsilon(t) - \gamma(\pi(t))| \ge \delta_0(\delta)\Big\} \cap \Big\{|\eta_x^\varepsilon(j\delta) - \gamma(\pi(j\delta))| \le \delta, j = 0,\dots,\frac{T}{\delta}\Big\}\Big\} \\
&\quad \le \mathbb{P}\Big\{\sup_{0\le t\le\delta} |\eta_x^\varepsilon(t + j\delta) - \eta_x^\varepsilon(j\delta)| \ge \frac{\delta_0(\delta)}{2} \text{ for some } j \le \frac{T}{\delta}\Big\} \qquad (65)\\
&\quad \le \frac{T}{\delta}\exp\left(-\frac{\delta}{\varepsilon}\min_{\phi\in S^{d-1}} L^s\left(\frac{\delta_0\phi}{2\delta}\right)\right) \le \frac{T}{\delta}\exp\left\{-\frac{2M}{\varepsilon}\right\} \le \exp\left\{-\frac{M}{\varepsilon}\right\}.
\end{aligned}$$

Next, for any partition of the interval $[0, T]$, $0 \le t_1 \le \cdots \le t_{N_1} \le T$, and for any collection of domains $\mathcal{B}_1, \dots, \mathcal{B}_{N_1}$ the following inequality holds:

$$\begin{aligned}
&\mathbb{P}\Big\{\bigcap_{j=1}^{N_1}\{\xi_x^\varepsilon(t_j) \in \mathcal{B}_j\}\Big\} \\
&\quad = \int_{\mathcal{B}_1} q^\varepsilon(x, y^1, t_1)\, dy^1 \int_{\mathcal{B}_2} q^\varepsilon(y^1, y^2, t_2 - t_1)\, dy^2 \dots \\
&\qquad \int_{\mathcal{B}_{N_1}} q^\varepsilon(y^{N_1 - 1}, y^{N_1}, t_{N_1} - t_{N_1 - 1})\, dy^{N_1} \\
&\quad \le \int_{\mathcal{B}_1} q_+^\varepsilon(x, y^1, t_1)\, dy^1 \int_{\mathcal{B}_2} q_+^\varepsilon(y^1, y^2, t_2 - t_1)\, dy^2 \dots \\
&\qquad \int_{\mathcal{B}_{N_1}} q_+^\varepsilon(y^{N_1 - 1}, y^{N_1}, t_{N_1} - t_{N_1 - 1})\, dy^{N_1} \\
&\quad \le \exp\Big(\frac{\mathrm{C}_1 T}{\varepsilon}\Big) \\
&\qquad \times \int_{\mathcal{B}_1} q_s^\varepsilon(x, y^1, t_1)\, dy^1 \int_{\mathcal{B}_2} q_s^\varepsilon(y^1, y^2, t_2 - t_1)\, dy^2 \dots \\
&\qquad \int_{\mathcal{B}_{N_1}} q_s^\varepsilon(y^{N_1 - 1}, y^{N_1}, t_{N_1} - t_{N_1 - 1})\, dy^{N_1} \\
&\quad = \exp\Big(\frac{\mathrm{C}_1 T}{\varepsilon}\Big)\mathbb{P}\{\bigcap_{j=1}^{N_1}\{\eta_x^\varepsilon(t_j) \in \mathcal{B}_j\}\}.
\end{aligned}$$

Combining this inequality with (65) yields the first inequality stated in the Proposition.

In order to prove the second one it suffices to observe that, due to the compactness of the set $\Phi(s)$ in $C([0, T]; \mathbb{R}^d)$, the function $\delta_0(\delta)$ can be chosen in such a way that $|\gamma(t') - \gamma(t'')| \le \frac{1}{4}\delta_0$ if $|t' - t''| \le 3\delta$ for all $\gamma \in \Phi(s)$. $\square$

PROPOSITION 4.11. *For any $\gamma \in \mathbf{D}([0, T]; \mathbb{R}^d)$, $\gamma(0) = x$, that is not absolutely continuous we have*

$$\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \varepsilon \log\big(\mathbb{P}\{\mathrm{dist}(\xi_x^\varepsilon(\cdot), \gamma(\cdot)) \le \delta\}\big) = -\infty.$$

PROOF. Consider auxiliary operators defined by

$$A^{\mathrm{u}} v(x) = \int_{\mathbb{R}^d} \Lambda^+ a(x - y) v(y)\, dy - \Lambda^- v(x) \int_{\mathbb{R}^d} a(x - y)\, dy$$

and

$$A^+ v(x) = \int_{\mathbb{R}^d} \Lambda^+ a(x - y) v(y)\, dy - \Lambda^+ v(x) \int_{\mathbb{R}^d} a(x - y)\, dy,$$

and the corresponding scaled operators

$$A^{\mathrm{u},\varepsilon} v(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} \Lambda^+ a\Big(\frac{x - y}{\varepsilon}\Big) v(y)\, dy - \frac{1}{\varepsilon^{d+1}} \Lambda^- v(x) \int_{\mathbb{R}^d} a\Big(\frac{x - y}{\varepsilon}\Big)\, dy$$

and

$$A^{+,\varepsilon} v(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} \Lambda^+ a\Big(\frac{x - y}{\varepsilon}\Big) v(y)\, dy - \frac{1}{\varepsilon^{d+1}} \Lambda^+ v(x) \int_{\mathbb{R}^d} a\Big(\frac{x - y}{\varepsilon}\Big)\, dy.$$

Denote by $q^{\mathrm{u},\varepsilon}(x,y,t)$, $q^{+,\varepsilon}(x,y,t)$ and $q^{\varepsilon}(x,y,t)$ the solutions of the equations

$$\partial_t q = A^{\mathrm{u},\varepsilon} q, \qquad \partial_t q = A^{+,\varepsilon} q \quad \text{and} \quad \partial_t q = A^{\varepsilon} q,$$

respectively, with the common initial condition $q(x,y,0)=\delta(y-x)$.

Since $\Lambda^+ \ge \Lambda(x,y)$ and $\Lambda^- \le \Lambda(x,y)$ for all x and y from $\mathbb{R}^d$, by the maximum principle we have

$$q^{\varepsilon}(x,y,t) \le q^{\mathrm{u},\varepsilon}(x,y,t) \quad \text{for all } x,y \in \mathbb{R}^d \text{ and } t>0. \tag{66}$$

It is also clear that

$$q^{\mathrm{u},\varepsilon}(x,y,t) = \exp\left(\frac{(\Lambda^+ - \Lambda^-)t}{\varepsilon}\right) q^{+,\varepsilon}(x,y,t).$$

For an arbitrary partition $0 \le t_1 < t_2 < \cdots < t_N \le T$ of the interval $[0,T]$, an arbitrary set $x^1,\dots,x^N$, $x^j \in \mathbb{R}^d$ and any $\delta > 0$ we have

$$\begin{aligned}
&\mathbb{P}\left\{\bigcap_{j=1}^{N}\{|\xi^{\varepsilon}(t_j)-x_j| \le \delta\}\right\} \\
&\quad = \int_{Q_\delta(x^1)} q^{\varepsilon}(0,y^1,t_1)\,dy^1 \int_{Q_\delta(x^2)} q^{\varepsilon}(y^1,y^2,t_2-t_1)\,dy^2 \dots \\
&\qquad \int_{Q_\delta(x^N)} q^{\varepsilon}(y^{N-1},y^N,t_N-t_{N-1})\,dy^N \\
&\quad \le \int_{Q_\delta(x^1)} q^{\mathrm{u},\varepsilon}(0,y^1,t_1)\,dy^1 \int_{Q_\delta(x^2)} q^{\mathrm{u},\varepsilon}(y^1,y^2,t_2-t_1)\,dy^2 \dots \\
&\qquad \int_{Q_\delta(x^N)} q^{\mathrm{u},\varepsilon}(y^{N-1},y^N,t_N-t_{N-1})\,dy^N \\
&\quad = \exp\left(\frac{(\Lambda^+ - \Lambda^-)T}{\varepsilon}\right) \times \\
&\qquad \times \int_{Q_\delta(x^1)} q^{+,\varepsilon}(0,y^1,t_1)\,dy^1 \int_{Q_\delta(x^2)} q^{+,\varepsilon}(y^1,y^2,t_2-t_1)\,dy^2 \dots \\
&\qquad \int_{Q_\delta(x^N)} q^{+,\varepsilon}(y^{N-1},y^N,t_N-t_{N-1})\,dy^N.
\end{aligned}$$

Let γ be an arbitrary curve in $\mathbf{D}([0,T];\mathbb{R}^d)$ which is not absolutely continuous. Setting $x^j = \gamma(t_j)$, taking uniform partitions of the interval $[0,T]$ and sending N to infinity, from the last relation we deduce

$$\mathbb{P}\Big\{\sup_{0\le t\le T}|\xi^{\varepsilon}(t)-\gamma(\pi(t))| \le \delta\Big\} \le \exp\left(\frac{(\Lambda^+ - \Lambda^-)T}{\varepsilon}\right)\mathbb{P}\Big\{\sup_{0\le t\le T}|\xi^{+,\varepsilon}(t)-\gamma(\pi(t))| \le \delta\Big\};$$

here $\xi^{+,\varepsilon}(t)$ is a process with independent increments whose generator is $A^{+,\varepsilon}$.

Due to [19], for any γ that is not absolutely continuous this yields

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\mathbb{P}\{\operatorname{dist}(\xi^{\varepsilon}(\cdot),\gamma(\cdot)) \le \delta\} = -\infty = -I_\Lambda(\gamma). \tag{67}$$

This implies the desired statement. $\square$

The main result of this section reads as follows.

THEOREM 4.12. *Let $\Lambda(x,y,\xi,\eta)=\Lambda(\xi,\eta)$, and assume that $\Lambda(\xi,\eta)$ is a measurable function for which conditions* (3)–(7) *and* (9) *are fulfilled. Then the process $\xi_x^\varepsilon(t)$, $0\le t\le T$, satisfies in $\mathbf{D}([0,T];\mathbb{R}^d)$ the large deviation principle with the rate function $I(\cdot)$ introduced in* (63).

In particular, for any $\gamma\in\mathbf{D}([0,T];\mathbb{R}^d)$, $\gamma(0)=x$, the following relation holds:

$$\lim_{\delta\to 0}\lim_{\varepsilon\to 0}\varepsilon\log\mathbb{P}\{\operatorname{dist}(\xi_x^\varepsilon(\cdot),\gamma(\cdot))\le\delta\}=-I(\gamma). \tag{68}$$

PROOF. For any $\gamma(\cdot)$ that is not absolutely continuous the relation

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\varepsilon\log[\mathbb{P}\{\{\operatorname{dist}(\xi_x^\varepsilon(\cdot),\gamma(\cdot))\le\delta\}]=-\infty$$

follows from Proposition 4.11.

Assume that $\gamma(\cdot)$ is absolutely continuous, and $\int_0^T L(\dot\gamma)\,dt<+\infty$. We consider a piecewise linear approximation of γ defined by

$$\gamma_N(t)=\begin{cases}\gamma(t) & \text{if } t=0,\frac{1}{N},\frac{2}{N},\dots,T,\\ \gamma(t_j)+(\gamma(t_{j+1})-\gamma(t_j))\dfrac{t-t_j}{t_{j+1}-t_j} & \text{if } t\in(t_j,t_{j+1}).\end{cases}$$

For any $\varkappa>0$ there exists $N_0=N_0(\varkappa)$ such that for any $N\ge N_0$

$$0\le\int_0^T L(\dot\gamma(t))\,dt-\int_0^T L(\dot\gamma_N(t))\,dt\le\varkappa.$$

Denote $\delta=\frac{1}{N}$. Then, by Proposition 4.9 there exists a function $\delta_0(\delta)$, $\delta_0:(0,1]\mapsto\mathbb{R}^+$, such that $\delta_0(\delta)\to 0$ as $\delta\to 0$, and for any $\pi\in\mathcal{K}$ with $\ell(\pi)\le\delta$

$$\begin{aligned}&\mathbb{P}\Bigl\{\{|\xi_x^\varepsilon(t_j)-\gamma(\pi(t_j))|\le\delta,\ j=0,\dots,N\}\cap\Bigl\{\sup_{0\le t\le T}|\xi_x^\varepsilon(t)-\gamma(\pi(t))|\ge\delta_0\Bigr\}\Bigr\}\\ &\quad\le\exp\Bigl(-\frac{M}{\varepsilon}(1+o(1))\Bigr),\end{aligned}\tag{69}$$

where $M=I(\gamma)+1$ and $o(1)\to 0$ as $\varepsilon\to 0$.

In order to achieve the upper bound we fix $N\ge N_0$ and choose $\delta_1>0$ in such a way that for any $\pi\in\mathcal{K}$ with $\ell(\pi)\le\delta_1$

$$\begin{aligned}&\mathbb{P}\{|\xi_y^\varepsilon(t_{j+1}-t_j)-(\gamma(\pi(t_{j+1}))-\gamma(\pi(t_j)))|\le\delta_1\}\\ &\quad\le\exp\Bigl[-\frac{t_{j+1}-t_j}{\varepsilon}\Bigl\{L\Bigl(\frac{\gamma(t_{j+1})-\gamma(t_j)}{t_{j+1}-t_j}\Bigr)-\varkappa\Bigr\}\Bigr]\\ &\quad=\exp\Bigl(-\frac{t_{j+1}-t_j}{\varepsilon}\{L(\dot\gamma_N(t))_{t\in(t_j,t_{j+1})}-\varkappa\}\Bigr)\end{aligned}\tag{70}$$

for all y such that $|y-\gamma(\pi(t_j))|\le\delta_1$ and for all sufficiently small ε. This choice is possible due to Theorem 4.7. Considering the Markov property of the process $\xi^\varepsilon(t)$ we deduce from (70) that for all sufficiently small $\varepsilon>0$ the following inequalities hold:

$$\begin{aligned}&\mathbb{P}\Bigl\{\sup_{0\le t\le T}|\xi_x^\varepsilon(t)-\gamma(\pi(t))|\le\delta_1\Bigr\}\\ &\quad\le\mathbb{P}\{|\xi_x^\varepsilon(t_j)-\gamma(\pi(t_j))|\le\delta_1,\ j=0,\dots,N\}\\ &\quad\le\prod_{j=0}^{N-1}\exp\Bigl(-\frac{t_{j+1}-t_j}{\varepsilon}\{L(\dot\gamma_N(t))_{t\in(t_j,t_{j+1})}-\varkappa\}\Bigr)\end{aligned}$$

$$
\begin{aligned}
&= \exp\Bigl(-\frac{1}{\varepsilon}\Bigl\{\int_0^T L(\dot\gamma_N(t))\,dt - T\varkappa\Bigr\}\Bigr)\\
&\le \exp\Bigl(-\frac{1}{\varepsilon}\Bigl\{\int_0^T L(\dot\gamma(t))\,dt - (T+1)\varkappa\Bigr\}\Bigr).
\end{aligned}
$$

This yields the desired upper bound in (68).

The lower bound can be obtained in a similar way. It suffices to combine the statement of Theorem 4.7 with (69) and use the Markov property of $\xi^\varepsilon(\cdot)$. Indeed, for any $\delta_0 > 0$ and $\varkappa > 0$ we choose the corresponding $\delta > 0$ and $\delta_1 > 0$ so that (69) holds and

$$
\begin{aligned}
&\mathbb{P}\bigl\{\bigl|\xi_y^\varepsilon(t_{j+1}-t_j) - \bigl(\gamma(t_{j+1}) - \gamma(t_j)\bigr)\bigr| \le \delta_1\bigr\}\\
&\quad\ge \exp\Bigl[-\frac{t_{j+1}-t_j}{\varepsilon}\Bigl\{L\Bigl(\frac{\gamma(t_{j+1})-\gamma(t_j)}{t_{j+1}-t_j}\Bigr) + \varkappa\Bigr\}\Bigr]\\
&\quad= \exp\Bigl(-\frac{t_{j+1}-t_j}{\varepsilon}\bigl\{L\bigl(\dot\gamma_N(t)\bigr)_{t\in(t_j,t_{j+1})} + \varkappa\bigr\}\Bigr)
\end{aligned}
\tag{71}
$$

for all y such that $|y-\gamma(t_j)| \le \delta_1$ and all sufficiently small $\varepsilon > 0$. Then considering the statement of Proposition 4.9 we have

$$
\begin{aligned}
&\mathbb{P}\Bigl\{\sup_{0\le t\le T}\bigl|\xi_x^\varepsilon(t) - \gamma(t)\bigr| \le \delta_0\Bigr\}\\
&\quad\ge \mathbb{P}\Bigl\{\bigl\{\bigl|\xi_x^\varepsilon(t_j) - \gamma(t_j)\bigr| \le \delta_1, j = 0,\dots,N\bigr\} \cap \Bigl\{\sup_{0\le t\le T}\bigl|\xi_x^\varepsilon(t) - \gamma(t)\bigr| \le \delta_0\Bigr\}\Bigr\}\\
&\quad\ge \mathbb{P}\bigl\{\bigl\{\bigl|\xi_x^\varepsilon(t_j) - \gamma(t_j)\bigr| \le \delta_1, j = 0,\dots,N\bigr\} - \exp\Bigl(-\frac{M}{\varepsilon}\Bigr)\\
&\quad\ge \prod_{j=0}^{N-1}\exp\Bigl(-\frac{t_{j+1}-t_j}{\varepsilon}\bigl\{L\bigl(\dot\gamma_N(t)\bigr)_{t\in(t_j,t_{j+1})} + \varkappa\bigr\}\Bigr) - \exp\Bigl(-\frac{M}{\varepsilon}\Bigr)\\
&\quad\ge \exp\Bigl(-\frac{1}{\varepsilon}\Bigl\{\int_0^T L\bigl(\dot\gamma_N(t)\bigr)\,dt + T\varkappa\Bigr\}\Bigr) \ge \exp\Bigl(-\frac{1}{\varepsilon}\Bigl\{\int_0^T L\bigl(\dot\gamma(t)\bigr)\,dt + T\varkappa\Bigr\}\Bigr);
\end{aligned}
\tag{72}
$$

here we have also used that fact that $M = I(\gamma) + 1$. This completes the proof of the lower bound in (68).

In order to justify the large deviation principle we need one more estimate. Recall that for any $s \in \mathbb{R}$ the symbol $\Phi(s)$ denotes $\Phi(s) = \{\gamma(\cdot) \in \mathbf{D}([0,T];\mathbb{R}^d) : I(\gamma) \le s, \gamma(0) = x\}$. Observe that the set $\Phi(s)$ consists of absolutely continuous curves and, according to Lemma 4.8, this set is compact.

LEMMA 4.13. *For any $s \in \mathbb{R}$, any $\varkappa > 0$ and any $\delta_0 > 0$ for all sufficiently small $\varepsilon > 0$ the following inequality holds*:

$$
\mathbb{P}\bigl\{\operatorname{dist}\bigl(\xi_x^\varepsilon(\cdot), \Phi(s)\bigr) > \delta_0\bigr\} \le \exp\Bigl\{-\frac{s-\varkappa}{\varepsilon}\Bigr\}.
\tag{73}
$$

PROOF. For any trajectory $\xi_x^\varepsilon(\cdot)$ and any $\delta = \frac{T}{N}$, $N \in \mathbb{Z}^+$ denote by $\gamma_{\delta,\omega}^\varepsilon(t)$ a piece-wise linear function such that

$$
\gamma_{\delta,\omega}^\varepsilon(j\delta) = \xi_x^\varepsilon(j\delta), \qquad j = 0, 1\dots, N;
$$

the argument ω indicates that $\gamma_\delta^\varepsilon(\cdot)$ is a random function, in what follows the dependence on ω is not indicated explicitly. We choose $\delta > 0$ such that

$$
\bigl|\gamma(t') - \gamma(t'')\bigr| \le \frac{1}{4}\delta_0 \quad \text{if } \bigl|t' - t''\bigr| \le \delta \text{ and } I\bigl(\gamma(\cdot)\bigr) \le s,
$$

and

$$\min_{\phi\in S^{d-1}}\left\{\delta L\left(\frac{\delta_0\phi}{2\delta}\right)\right\}\geq s+1.$$

Denote by $\mathcal{E}_-$ and $\mathcal{E}_+$ the events

$$\mathcal{E}_-=\{\xi_x^\varepsilon(\cdot)\notin\Phi_{\delta_0}(s),I(\gamma_\delta^\varepsilon)<s\},\qquad \mathcal{E}_+=\{\xi_x^\varepsilon(\cdot)\notin\Phi_{\delta_0}(s),I(\gamma_\delta^\varepsilon)\geq s\},$$

where $\Phi_{\delta_0}(s)=\{\gamma(\cdot)\in\mathbf{D}([0,T];\mathbb{R}^d):\operatorname{dist}(\gamma,\Phi(s))\leq\delta_0\}$. By Proposition 4.9 for all sufficiently small $\varepsilon>0$ we have

$$\mathbb{P}(\mathcal{E}_-)\leq\mathbb{P}\Big\{\sup_{0\leq t\leq T}\left|\xi_x^\varepsilon(t)-\gamma_\delta^\varepsilon(t)\right|\geq\delta_0\Big\}\leq\exp\left(-\frac{s+1}{\varepsilon}\right).\tag{74}$$

Consider a (Nd)-dimensional vector $\{\xi_x^\varepsilon((j+1)\delta)-\xi_x^\varepsilon(j\delta)\}_{j=1}^{N-1}$. By Theorem 4.7 taking into account Markov property of $\xi^\varepsilon(\cdot)$ we deduce that the family of random vectors $\{\xi_x^\varepsilon((j+1)\delta)-\xi_x^\varepsilon(j\delta)\}_{j=0}^{N-1}$ satisfies for any $\lambda_0,\dots,\lambda_{N-1}\in\mathbb{R}^d$ the following relation:

$$\lim_{\varepsilon\to0}\varepsilon\log\mathbb{E}\left\{\exp\left[\sum_{j=0}^{N-1}\lambda_j\cdot\left(\xi_x^\varepsilon((j+1)\delta)-\xi_x^\varepsilon(j\delta)\right)\right]\right\}=\sum_{j=0}^{N-1}\delta h(\lambda_j).$$

By the Gärtner–Ellis theorem this implies the upper large deviation bound with the rate function

$$L_\delta(p_1)+L_\delta(p_2)+\cdots+L_\delta(p_N),\qquad p_j\in\mathbb{R}^d,$$

where $L_\delta(p)=\delta L(\frac{p}{\delta})$, as was defined in (60). For an arbitrary piece-wise linear function γ corresponding to the partition $\{j\delta\}_{j=0}^N$ we have

$$I(\gamma)=\sum_{j=0}^{N-1}L_\delta\left(\gamma((j+1)\delta)-\gamma(j\delta)\right).$$

Therefore, by the Gärtner–Ellis theorem, for sufficiently small $\varepsilon>0$ we have

$$\begin{aligned}\mathbb{P}(\mathcal{E}_+)&\leq\mathbb{P}\{I(\gamma_\delta^\varepsilon)\geq s\}\\&=\mathbb{P}\left\{\sum_{j=0}^{N-1}L_\delta\left(\xi_x^\varepsilon((j+1)\delta)-\xi_x^\varepsilon(j\delta)\right)\geq s\right\}\leq\exp\left(-\frac{s-\varkappa}{\varepsilon}\right).\end{aligned}$$

Combining this estimate with (74) yields the desired statement. □

From the proof of Lemma 4.13 it follows that for any $s_0>0$ inequality (73) holds uniformly in $s\in[0,s_0]$, that is, for any $\varkappa>0$ and $\delta_0>0$ there exists $\varepsilon_0>0$ such that (73) holds for all $\varepsilon\leq\varepsilon_0$ and all $s\leq s_0$.

It is then well known (see, for instance, [11]) that the lower bound in (72) and Lemma 4.13 imply the large deviation principle stated in Theorem. □

5. Environments with slowly varying characteristics. In this section we consider the case of environments whose characteristics $\Lambda(x,y)$ do not depend on the fast variables that is, Λ is a continuous function on $\mathbb{R}^{2d}$ for which condition (9) is fulfilled. Our approach in this section is somehow inspired by the small perturbations arguments used in the previous works, in particular in the Wentzell–Freidlin theory; see [11]. However, the results from these works do not apply directly to the operators considered in the present paper and require some adaptation.

Under the assumptions of this section the generator of $\xi^\varepsilon(t)$ takes the form

$$A^\varepsilon u(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} a\left(\frac{x-y}{\varepsilon}\right) \Lambda(x,y)\big(u(y)-u(x)\big)\,dy, \quad u \in L^2(\mathbb{R}^d), \tag{75}$$

$\varepsilon > 0$ is a small parameter and the convolution kernel $a(z)$ in (75) satisfies conditions (3)–(5) introduced in the previous section.

REMARK. According to Corollary 5.2 below the Markov jump process $\xi^\varepsilon(t)$ is a small random perturbation of a deterministic trajectory determined by an ordinary differential equation $\dot{x} = b(x)$ with

$$b(x) = -\Lambda(x,x) \int a(z) z\,dz.$$

5.1. *Markov process with slow variables*. We turn now to the case of nonconstant $\Lambda(x,y)$ that does not depend on the fast variables and recall that the function $\Lambda(x,y)$ is continuous in both variables and satisfies condition (9). Since Λ does not depend on the fast variables, condition (8) can be replaced with the following continuity condition:

$$\Lambda(x,y) \quad \text{is continuous on } \mathbb{R}^d \times \mathbb{R}^d. \tag{76}$$

After changing variables $\tilde{x} = \frac{x}{\varepsilon}$ the operator A^ε in (75) takes the form

$$\tilde{A}^\varepsilon u(\tilde{x}) = \frac{1}{\varepsilon} \int_{\mathbb{R}^d} a(\tilde{x}-\tilde{y}) \Lambda(\varepsilon\tilde{x}, \varepsilon\tilde{y})\big(u(\tilde{y}) - u(\tilde{x})\big)\,d\tilde{y}. \tag{77}$$

The Hamiltonian $H(x,\lambda)$ and the Lagrangian $L(x,\zeta)$ are introduced in this case as follows:

$$H(x,\lambda) = \Lambda(x,x)\left(\int a(z) e^{-\lambda z}\,dz - 1\right) = \Lambda(x,x) H(\lambda), \tag{78}$$

$$L(x,\zeta) = \sup_\lambda \{\lambda\zeta - \Lambda(x,x) H(\lambda)\} = \Lambda(x,x) L\left(\frac{\zeta}{\Lambda(x,x)}\right). \tag{79}$$

Observe that the function $L(x,\zeta)$ is continuous and nonnegative on $\mathbb{R}^d \times \mathbb{R}^d$. Moreover, it is smooth and strictly convex in $\zeta \in \mathbb{R}^d$. The corresponding *rate function* I_Λ is defined by

$$I_\Lambda\big(\gamma(\cdot)\big) = \begin{cases} \int_0^T L\big(\gamma(t), \dot{\gamma}(t)\big)\,dt & \text{if } \gamma \text{ is absolutely continuous,} \\ +\infty & \text{otherwise.} \end{cases}$$

THEOREM 5.1. *Under assumptions* (3)–(6), (9) *and* (76) *the family of processes* $\{\xi^\varepsilon(t), 0 \le t \le T\}$ *satisfies, as* $\varepsilon \to 0$, *the large deviation principle in the Skorokhod space* $\mathbf{D}([0,T];\mathbb{R}^d)$ *with the rate function* $I_\Lambda(\cdot)$.

The proof of this Theorem is based on the large deviations results obtained for processes with independent increments and the arguments of perturbation theory. Since, in contrast with the case of diffusion processes, in our case the coefficient $\Lambda(x,y)$ depends on two variables, x and y, the results of the previous works do not apply directly and require some adaptation. The detailed proof of Theorem 5.1 is provided in the Appendix.

Denote

$$\zeta(x) = \arg\min_p L_\Lambda(x,p).$$

It is straightforward to check that

$$\zeta(x) = -\Lambda(x,x) \int_{\mathbb{R}^d} a(z) z\,dz = \Lambda(x,x) \nabla H(\lambda)|_{\lambda=0}. \tag{80}$$

Letting $\gamma_x^0(t)$ be the solution of the ODE

$$\dot\gamma(t) = \zeta\big(\gamma(t)\big), \quad \gamma(0) = x,$$

one can deduce from the last theorem the following.

COROLLARY 5.2. *For any* $x \in \mathbb{R}^d$

$$\lim_{\varepsilon\to 0} \mathbb{E}\Big(\sup_{0\le t\le T} \big|\xi_x^\varepsilon(t) - \gamma_x^0(t)\big|\Big) = 0.$$

6. The general case of locally periodic environment $\Lambda(x, y, \frac{x}{\varepsilon}, \frac{y}{\varepsilon})$. In this section we consider the case of the most general locally periodic media. Here we assume that $\Lambda^\varepsilon(x, y) = \Lambda(x, y, \frac{x}{\varepsilon}, \frac{y}{\varepsilon})$, where $\Lambda(x, y, \xi, \eta)$ satisfies conditions (7)–(9).

Here for each $x \in \mathbb{R}^d$ we introduce a Hamiltonian $H = H(x, \lambda)$ in the same way as in (57), x being a parameter. Namely, we set

$$H(x, \lambda) := \mathrm{s}(A_{x,\lambda}) = \begin{cases} \theta(x, \lambda), & \lambda \in \Gamma(x), \\ -g_{\min}(x), & \text{otherwise}; \end{cases} \tag{81}$$

here

$$A_{x,\lambda}u(z) = \int_{\mathbb{R}^d} \Lambda(x, x, z, y) a(z - y) e^{\lambda\cdot(y-z)} u(y)\, dy - \int_{\mathbb{R}^d} \Lambda(x, x, z, y) a(z - y)\, dy u(z),$$

and, for each x, we define $\Gamma(x)$, $\theta(x, \lambda)$ and $g_{\min}(x)$ in the same way as in Section 4. Then we introduce the corresponding Lagrangian $L(x, \zeta)$. The main result of this section reads:

THEOREM 6.1. *Let conditions* (3)–(9) *be fulfilled. Then the family of processes* $\xi_x^\varepsilon(\cdot)$ *with the generators* A^ε *defined in* (2) *satisfies, as* $\varepsilon \to 0$, *the large deviation principle in the path space* $\mathbf{D}([0, T]; \mathbb{R}^d)$; *the corresponding rate function is given by*

$$I\big(\gamma(\cdot)\big) = \begin{cases} \displaystyle\int_0^T L\big(\gamma(t), \dot\gamma(t)\big)\, dt & \textit{if } \gamma \textit{ is absolutely continuous and } \gamma(0) = x, \\ +\infty & \textit{otherwise}. \end{cases}$$

PROOF. The proof relies on combining the statement of Theorem 4.7 and the arguments used in the proof of Theorem 5.1. We leave the details to the reader. □

It is interesting to observe that for small $\varepsilon > 0$ the process $\xi_x^\varepsilon(\cdot)$ can be interpreted as a small random perturbation of a deterministic dynamical system defined by the ODE

$$\dot\gamma(t) = \nabla_\lambda H\big(\gamma(t), 0\big), \quad \gamma(0) = x. \tag{82}$$

COROLLARY 6.2. *For any* $x \in \mathbb{R}^d$

$$\lim_{\varepsilon\to 0} \mathbb{E}\Big\{\sup_{0\le t\le T} \big|\xi_x^\varepsilon(t) - \gamma_x(t)\big|\Big\} = 0,$$

where $\gamma_x(\cdot)$ *is a solution of* (82).

APPENDIX

PROOF OF THEOREM 5.1. Consider an absolutely continuous curve $\gamma(\cdot)$ such that

$$I_\Lambda(\gamma) = \int_0^T L_{\Lambda(\gamma(t),\gamma(t))}(\dot\gamma(t))\,dt < +\infty.$$

We first justify the upper bound. For any $N \geq 2$ denote by γ_N a piece-wise linear interpolation of γ such that $\gamma_N(t_j) = \gamma(t_j)$ with $t_j = j\frac{T}{N}$, $j = 0, 1, \dots, N$, and by $\widehat\gamma_N$ the corresponding piece-wise constant interpolation, $\widehat\gamma_N(t) = \gamma(t_j)$ for $t \in [t_j, t_j + \frac{1}{N})$. For any $\varkappa > 0$ there exists $\delta > 0$ such that for all $N \geq \delta^{-1}$ we have

$$\int_0^T L_{\Lambda(\widehat\gamma_N(t),\widehat\gamma_N(t))}(\dot\gamma_N(t))\,dt > I_\Lambda(\gamma) - \varkappa.$$

Denote by $\nu(s)$ the modulus of continuity of $L(x, y)$ in 1-neighbourhood of the curve γ. Since $\min_{\phi\in S^{d-1}} r^{-1}L(r\phi)$ tends to infinity as $r \to \infty$, there exists a function $\delta_0(\delta) > 0$ such that $\delta_0(\delta) \to 0$ as $\delta \to 0$, and $\min_{\phi\in S^{d-1}}\{\delta L(\frac{r\phi}{\delta}) : r \geq \delta_0\} \to \infty$. It is then clear that for any sufficiently small $\delta > 0$ there exists $\delta_1(\delta) > 0$ such that

$$\left|\sum_{j=0}^{N-1} \delta L_{(\Lambda(\gamma(t_j),\gamma(t_j))+\nu(\delta_0))}\left(\frac{x_{j+1} - x_j}{\delta}\right) - \int_0^T L_{\Lambda(\widehat\gamma_N(t),\widehat\gamma_N(t))}(\dot\gamma_N(t))\,dt\right| \leq \varkappa,$$

if $|x_j - \gamma(t_j)| \leq \delta_1$, $j = 0, \dots, N$.

Consider a Markov process $\xi_x^{\varepsilon,N}(t)$, $0 \leq t \leq T$, whose generator on the interval $[t_j, t_j + \frac{1}{N})$ is

$$A_{t_j}^\varepsilon v(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} \big(\Lambda(\gamma(t_j), \gamma(t_j)) + \nu(\delta_0)\big) a\left(\frac{x-y}{\varepsilon}\right)\big(v(y) - v(x)\big)\,dy,$$
$$j = 0, 1, \dots, N.$$

We also define a Markov process $\widetilde\xi_x^{\varepsilon,N}(t)$, $0 \leq t \leq T$ such that its generator on the interval $[t_j, t_j + \frac{1}{N})$ reads

$$\widetilde A_{t_j}^\varepsilon v(x) = \frac{1}{\varepsilon^{d+1}} \int_{\mathbb{R}^d} (\Lambda_{t_j}(x, y) a\left(\frac{x-y}{\varepsilon}\right)\big(v(y) - v(x)\big)\,dy,$$
$$j = 0, 1, \dots, N,$$

with $\Lambda_{t_j}(x, y) = \Lambda(\gamma(t_j), \gamma(t_j)) + \nu(\delta_0)$ if $|x| + |y| \leq \delta_0$, and $\Lambda_{t_j}(x, y) = \Lambda(x, y)$ otherwise.

Using the inequality similar to that in (66) we conclude that for any x such that $|x - \gamma(t_j)| \leq \delta_1$ the density $\widetilde q^{\varepsilon,N}(t, x, y)$ of the process $\widetilde\xi_x^{\varepsilon,N}(t)$, $\widetilde\xi_x^{\varepsilon,N}(t_j) = x$, on the set $\{(y, t) : t_j \leq t \leq t_j + \delta, |y - \gamma(t_j)| > 2\delta_0\}$ does not exceed $\exp(-\frac{\delta}{\varepsilon}L_{\Lambda^+}(\frac{\delta_0}{\delta}))$. Denote by $q^{\varepsilon,N}(t, x, y)$ the density of the process $\xi_x^{\varepsilon,N}(t)$, $\xi_x^{\varepsilon,N}(t_j) = x$. Straightforward computations show that the difference $\Xi^\varepsilon(t, x, y) = \widetilde q^{\varepsilon,N}(t, x, y) - q^{\varepsilon,N}(t, x, y)$ satisfies on the interval $(t_j, t_j + \delta)$ the equation

$$\partial_t \Xi^\varepsilon = \int_{\mathbb{R}^d} [\Lambda(\gamma(t_j), \gamma(t_j)) + \nu(\delta_0)] a\left(\frac{y-z}{\varepsilon}\right)(\Xi^\varepsilon(t, x, z) - (\Xi^\varepsilon(t, x, y))\,dz + R^\varepsilon(t, x, y)$$

with

$$|R^\varepsilon(t, x, y)| \leq \exp\left(-\frac{\delta}{2\varepsilon}L_{\Lambda^+}\left(\frac{\delta_0}{\delta}\right)\right)$$

and $\Xi^\varepsilon(t_j, x, y) = 0$.

With the help of the standard a priori estimates this yields

$$\|\widetilde{q}^{\varepsilon,N}(t,x,y)-q^{\varepsilon,N}(t,x,y)\|_{L^2(\mathbb{R}^d)}\le\exp\biggl(-\frac{\delta}{4\varepsilon}L_{\Lambda^+}\biggl(\frac{\delta_0}{\delta}\biggr)\biggr).$$

We choose $\delta>0$ in such a way that $\frac{\delta}{4}L_{\Lambda^+}(\frac{\delta_0}{\delta})>M$ with $M=I_\Lambda(\gamma)+1$. Combining the above estimates we obtain

$$\begin{aligned}
&\mathbb{P}\{\operatorname{dist}(\xi^\varepsilon(\cdot),\gamma(\cdot))\le\delta_1\}\\
&\quad\le\exp\biggl(\frac{2\nu(\delta_0)T}{\varepsilon}\biggr)\mathbb{P}\{\operatorname{dist}(\widetilde{\xi}^{\varepsilon,N}(\cdot),\gamma(\cdot))\le\delta_1\}\\
&\quad\le\exp\biggl(\frac{2\nu(\delta_0)T}{\varepsilon}\biggr)\biggl[\mathbb{P}\{\{|\xi^{\varepsilon,N}(t_j)-\gamma(\pi(t_j))|\le\delta_1,\ j=0,\dots,N\}\}+\exp\biggl(-\frac{M}{\varepsilon}\biggr)\biggr]\qquad(83)\\
&\quad\le\exp\biggl(\frac{2\nu(\delta_0)T}{\varepsilon}\biggr)\exp\biggl[\frac{(1+o(1))}{\varepsilon}\biggl(-\int_0^T L_{\Lambda(\widehat{\gamma}_N(t),\widehat{\gamma}_N(t))}(\dot{\gamma}_N(t))\,dt+\varkappa\biggr)\biggr]\\
&\quad\le\exp\biggl(\frac{2\nu(\delta_0)T}{\varepsilon}\biggr)\exp\biggl[\frac{(1+o(1))}{\varepsilon}(-I_\Lambda(\gamma)+2\varkappa)\biggr],
\end{aligned}$$

where $o(1)$ tends to zero as $\varepsilon\to0$. This implies the desired upper bound.

We turn to the lower bound. Here we introduce $\delta_1=\delta_1(\delta)$ and $\delta_0=\delta_0(\delta)$ in such a way that

$$\biggl|\sum_{j=0}^{N-1}\delta L_{(\Lambda(\gamma(t_j),\gamma(t_j))-\nu(\delta_0))}\biggl(\frac{x_{j+1}-x_j}{\delta}\biggr)-\int_0^T L_{\Lambda(\widehat{\gamma}_N(t),\widehat{\gamma}_N(t))}(\dot{\gamma}_N(t))\,dt\biggr|\le\varkappa.$$

Define Markov processes $\xi^{\varepsilon,N}_{-,x}(t)$ and $\widetilde{\xi}^{\varepsilon,N}_{-,x}(t)$, $0\le t\le T$, whose generators on the interval $[t_j,t_j+\frac{1}{N})$ read, respectively,

$$\begin{aligned}
A^\varepsilon_{-,t_j}v(x)&=\frac{1}{\varepsilon^{d+1}}\int_{\mathbb{R}^d}[\Lambda(\gamma(t_j),\gamma(t_j))-\nu(\delta_0)]a\biggl(\frac{x-y}{\varepsilon}\biggr)(v(y)-v(x))\,dy,\\
&j=0,1,\dots,N
\end{aligned}$$

and

$$\widetilde{A}^\varepsilon_{-,t_j}v(x)=\frac{1}{\varepsilon^{d+1}}\int_{\mathbb{R}^d}(\Lambda_{-,t_j}(x,y)a\biggl(\frac{x-y}{\varepsilon}\biggr)(v(y)-v(x))\,dy,\qquad j=0,1,\dots,N,$$

with $\Lambda_{-,t_j}(x,y)=\Lambda(\gamma(t_j),\gamma(t_j))-\nu(\delta_0)$ if $|x|+|y|\le\delta_0$, and $\Lambda_{t_j}(x,y)=\Lambda(x,y)$ otherwise.

By comparison with the process $\xi^{+,\varepsilon}(t)$ one can show that for any $M>0$ for sufficiently small $\delta>0$ we have

$$\mathbb{P}\Bigl\{\sup_{t_j\le t\le t_j+\delta}|\xi^\varepsilon(t)-\xi^\varepsilon(t_j)|\ge\delta_0\Bigr\}\le\exp\biggl(-\frac{M}{\varepsilon}\biggr).$$

Using this inequality and choosing $M=I_\Lambda(\gamma)+2$, in the same way as in the proof of the upper bound we obtain

$$\begin{aligned}
&\mathbb{P}\Bigl\{\sup_{0\le t\le T}|\xi^\varepsilon(t)-\gamma(t)|\le\delta_0\Bigr\}\\
&\quad\ge\mathbb{P}\Bigl\{\max_j|\xi^\varepsilon(t_j)-\gamma(t_j)|\le\delta_1\Bigr\}-\exp\biggl(-\frac{M-1}{\varepsilon}\biggr)\\
&\quad\ge\exp\biggl(-\frac{2\nu(\delta_0)T}{\varepsilon}\biggr)\mathbb{P}\Bigl\{\max_j|\widetilde{\xi}^{\varepsilon,N}(t_j)-\gamma(t_j)|\le\delta_1\Bigr\}-\exp\biggl(-\frac{M-1}{\varepsilon}\biggr)
\end{aligned}$$

$$
\begin{aligned}
&\geq \exp\Big(-\frac{2\nu(\delta_0)T}{\varepsilon}\Big)\mathbb{P}\Big\{\max_j\big|\xi^{\varepsilon,N}(t_j)-\gamma(t_j)\big|\leq\delta_1\Big\}-\exp\Big(-\frac{M-1}{\varepsilon}\Big)\\
&\geq \exp\Big(-\frac{2\nu(\delta_0)T}{\varepsilon}\Big)\exp\Big[\frac{(1+o(1))}{\varepsilon}\Big(-\int_0^T L_{\Lambda(\widehat{\gamma}_N(t),\widehat{\gamma}_N(t))}\big(\dot{\gamma}_N(t)\big)\,dt-\varkappa\Big)\Big]\\
&\geq \exp\Big(-\frac{2\nu(\delta_0)T}{\varepsilon}\Big)\exp\Big[\frac{(1+o(1))}{\varepsilon}\big(-I_\Lambda(\gamma)-2\varkappa\big)\Big],
\end{aligned}
$$

where $o(1)$ tends to zero as $\varepsilon\to 0$. This yields the lower bound.

We should also show that for any $s\geq 0$, any $\delta_0>0$ and any $\varkappa>0$

$$
\mathbb{P}\big\{\operatorname{dist}\big(\xi_x^\varepsilon(\cdot),\Phi(s)\big)>\delta_0\big\}\leq\exp\Big\{-\frac{s-\varkappa}{\varepsilon}\Big\} \tag{84}
$$

for all sufficiently small ε. The proof of this inequality relies on the arguments from the proof of Lemma 4.13 and that of inequality (83). One should combine these arguments in a straightforward way. We skip the details. □

REFERENCES

[1] BALDI, P. (1991). Large deviations for diffusion processes with homogenization and applications. *Ann. Probab.* **19** 509–524. MR1106273

[2] BOROVKOV, A. A. (1967). Boundary value problems for random walks and large deviations in function spaces. *Theory Probab. Appl.* **12** 635–654. MR0243585

[3] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. Springer, New York. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4

[4] EDMUNDS, D. E., POTTER, A. J. B. and STUART, C. A. (1972). Non-compact positive operators. *Proc. R. Soc. Lond. Ser. A*, *Math. Phys. Sci.* **328** 67–81. MR0637097 https://doi.org/10.1098/rspa.1972.0069

[5] ENGEL, K.-J. and NAGEL, R. (2000). *One-Parameter Semigroups for Linear Evolution Equations*. *Graduate Texts in Mathematics* **194**. Springer, New York. MR1721989

[6] FENG, J. and KURTZ, T. G. (2006). *Large Deviations for Stochastic Processes*. *Mathematical Surveys and Monographs* **131**. Amer. Math. Soc., Providence, RI. MR2260560 https://doi.org/10.1090/surv/131

[7] FERNANDO, K. and HEBBAR, P. (2021). Higher order asymptotics for large deviations—Part I. *Asymptot. Anal.* **121** 219–257. MR4212539 https://doi.org/10.3233/asy-201602

[8] FERNANDO, K. and HEBBAR, P. (2021). Higher order asymptotics for large deviations—Part II. *Stoch. Dyn.* **21** Paper No. 2150025. MR4297948 https://doi.org/10.1142/S0219493721500258

[9] FREIDLIN, M. I. and SOWERS, R. B. (1999). A comparison of homogenization and large deviations, with applications to wavefront propagation. *Stochastic Process. Appl.* **82** 23–52. MR1695068 https://doi.org/10.1016/S0304-4149(99)00003-4

[10] FREIDLIN, M. I. and WENTZELL, A. D. (1970). On small random perturbations of dynamical systems. *Russian Math. Surveys* **25** 1–55.

[11] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **260**. Springer, New York. MR0722136 https://doi.org/10.1007/978-1-4684-0176-9

[12] GRIGOR'YAN, A., KONDRATIEV, Y., PIATNITSKI, A. and ZHIZHINA, E. (2018). Pointwise estimates for heat kernels of convolution-type operators. *Proc. Lond. Math. Soc.* (3) **117** 849–880. MR3873136 https://doi.org/10.1112/plms.12144

[13] HENNION, H. and HERVÉ, L. (2001). *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness*. *Lecture Notes in Math.* **1766**. Springer, Berlin. MR1862393 https://doi.org/10.1007/b87874

[14] KRASNOSEL'SKIJ, M. A., LIFSHITS, E. A. and SOBOLEV, A. V. (1989). *Positive Linear Systems*. *The Method of Positive Operators*. *Sigma Series in Applied Mathematics* **5**. Heldermann Verlag, Berlin. MR1038527

[15] LYNCH, J. and SETHURAMAN, J. (1987). Large deviations for processes with independent increments. *Ann. Probab.* **15** 610–627. MR0885133

[16] MOGUL'SKIĬ, A. A. (1993). Large deviations for processes with independent increments. *Ann. Probab.* **21** 202–215. MR1207223

[17] MOGUL'SKIĬ, A. A. (2016). The large deviations principle for a generalized Poisson process. *Mat*. *Tr.* **19** 119–157. MR3643017

[18] PIATNITSKI, A. and ZHIZHINA, E. (2019). Homogenization of biased convolution type operators. *Asymptot*. *Anal*. **115** 241–262. MR4033742 https://doi.org/10.3233/asy-191533

[19] PUHALSKII, A. (1994). The method of stochastic exponentials for large deviations. *Stochastic Process*. *Appl*. **54** 45–70. MR1302694 https://doi.org/10.1016/0304-4149(94)00004-2

[20] WENTZELL, A. D. (1990). *Limit Theorems on Large Deviations for Markov Stochastic Processes*. *Mathematics and Its Applications* (*Soviet Series*) **38**. Kluwer Academic, Dordrecht. MR1135113 https://doi.org/10.1007/978-94-009-1852-8

[21] ZHIZHINA, E. A. and PIATNITSKI, A. L. (2019). On the behaviour at infinity of solutions to nonlocal parabolic type problems. *Izv. Irkutsk*. *Gos*. *Univ. Ser. Mat*. **30** 99–113. MR4053915 https://doi.org/10.26516/1997-7670.2019.30.99