

Pointwise estimates for heat kernels of convolution-type operators

Alexander Grigor'yan, Yuri Kondratiev, Andrey Piatnitski and Elena Zhizhina

ABSTRACT

We study the large-time behaviour of the fundamental solution of parabolic equations with an elliptic part being non-local convolution-type operator. We assume that this operator is a generator of a Markov jump process, and that its convolution kernel decays at least exponentially at infinity. The fundamental solution shows rather different asymptotic behaviour depending on whether $|x| \lesssim \sqrt{t}$, or $\sqrt{t} \ll |x| \ll t$, or $|x| \sim t$, or $|x| \gg t$. In each of these regions we obtain sharp pointwise estimates for the fundamental solution.

1. Introduction

In this paper we are concerned with estimates of the heat kernel (=fundamental solution) of certain evolution equations with non-local elliptic part. The heat kernel of the classical heat equation

$$\partial_t u - \Delta u = 0,$$

where Δ is the Laplace operator in \mathbb{R}^d , is given by the Gauss–Weierstrass function

$$p_t(x, y) = \frac{1}{(4\pi t)^{d/2}} \exp\left(-\frac{|x-y|^2}{4t}\right). \quad (1.1)$$

For a more general parabolic equation

$$\partial_t u - Lu = 0,$$

where L is a uniformly elliptic second-order operator in divergence form, Aronson [2] proved the following Gaussian estimates for its heat kernel:

$$p_t(x, y) \asymp \frac{C}{t^{d/2}} \exp\left(-\frac{|x-y|^2}{ct}\right),$$

where the sign \asymp means both \leq and \geq but with different values of positive constants C, c .

A simplest heat equation with non-local elliptic part is

$$\partial_t u + (-\Delta)^{\alpha/2} u = 0, \quad (1.2)$$

where $0 < \alpha < 2$. Applying the subordination techniques of [20] to the Gauss–Weierstrass function, one obtains that the heat kernel of (1.2) satisfies the following estimates:

$$p_t(x, y) \asymp \frac{C}{t^{d/\alpha}} \left(1 + \frac{|x-y|}{t^{1/\alpha}}\right)^{-(d+\alpha)} \quad (1.3)$$

(see also [3]). Note that $(-\Delta)^{\alpha/2}$ is an integro-differential operator of the form

$$(-\Delta)^{\alpha/2} f(x) = c_{d,\alpha} \text{p.v.} \int_{\mathbb{R}^d} \frac{f(x) - f(y)}{|x-y|^{d+\alpha}} dy. \quad (1.4)$$

Received 20 September 2017; revised 11 February 2018; published online 16 April 2018.

2010 *Mathematics Subject Classification* 47D06, 60J35 (primary), 60F05, 60J75 (secondary).

The work was partially supported by SFB1283 of German Research Council. The third and the fourth authors were supported by the Russian Science Foundation, Project 14-50-00150.

The heavy tail of the heat kernel in the estimate (1.3) is a consequence of the heavy integral kernel in (1.4). Similar estimates hold also for non-local heat kernels on fractals [12].

A natural class of non-local operators arises on graphs. Let Γ be a countable, locally finite, connected graph. Let $d(x, y)$ be the graph distance on Γ . The discrete Laplace operator Δ on Γ acts on functions $f : \Gamma \rightarrow \mathbb{R}$ as follows:

$$\Delta f(x) = \frac{1}{\deg(x)} \sum_{\{y \in \Gamma: y \sim x\}} (f(y) - f(x)) = \sum_{y \in \Gamma} (f(y) - f(x)) J(x, y),$$

where

$$J(x, y) = \frac{1}{\deg(x)} \mathbf{1}_{\{d(x,y)=1\}}.$$

Davies has obtained in [8] the upper bounds of the heat kernel $p_t(x, y)$ of the heat equation $\partial_t u - \Delta u = 0$ on Γ that in the case of uniformly bounded degree $\deg(x)$ of vertices amounts to

$$p_t(x, y) \leq \exp\left(-ct\Phi\left(\frac{d(x, y)}{ct}\right)\right), \tag{1.5}$$

where

$$\Phi(\xi) = \sup_{\lambda > 0} \{\xi\lambda - \cosh \lambda\} = \xi \ln\left(\xi + \sqrt{\xi^2 + 1}\right) - \sqrt{1 + \xi^2}.$$

Since

$$\Phi(\xi) \sim \frac{\xi^2}{2} \text{ as } \xi \rightarrow 0 \quad \text{and} \quad \Phi(\xi) \sim \xi \ln \xi \text{ as } \xi \rightarrow \infty, \tag{1.6}$$

the estimate (1.5) implies for small $\frac{d(x,y)}{t}$ the Gaussian estimate

$$p_t(x, y) \leq \exp\left(-\frac{d^2(x, y)}{ct}\right),$$

and for large $(d(x, y))/t$

$$p_t(x, y) \leq \exp\left(-cd(x, y) \ln \frac{d(x, y)}{ct}\right).$$

Estimate (1.5) gives a rather sharp upper bound of the tail of the heat kernel on an arbitrary graph because on $\Gamma = \mathbb{Z}$ the heat kernel admits the following two-sided estimate (see [19])

$$p_t(x, y) \asymp \frac{C}{(t + d(x, y))^{1/2}} \exp\left(-2t\Phi\left(\frac{d(x, y)}{2t}\right)\right).$$

In this paper we consider the non-local operator A on functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ given by

$$Af = a * f - f, \tag{1.7}$$

where the convolution kernel a is such that

$$a(x) \geq 0; \quad a(x) = a(-x); \quad a(x) \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d), \tag{1.8}$$

$$\int_{\mathbb{R}^d} a(x) dx = 1, \quad \int_{\mathbb{R}^d} |x|^2 a(x) dx < \infty. \tag{1.9}$$

In particular, under condition (1.9) there exists a positive definite matrix $\sigma = \{\sigma_{ij}\}$ with $\sigma_{ij} = \int_{\mathbb{R}^d} x_i x_j a(x) dx$. The third condition in (1.8) implies that $a(x) \in L^2(\mathbb{R}^d)$, and for the Fourier transform $\hat{a}(p)$ we have

$$\hat{a}(p) \in C_b(\mathbb{R}^d) \cap L^2(\mathbb{R}^d), \quad \max_{\mathbb{R}^d} \hat{a}(p) = \hat{a}(0) = 1, \quad \hat{a}(p) \rightarrow 0 \text{ as } |p| \rightarrow \infty. \tag{1.10}$$

The operator A takes a form of an integro-differential operator as follows:

$$Af(x) = \int_{\mathbb{R}^d} (f(y) - f(x)) a(x-y) dy.$$

An essential difference from the operator (1.4) is that the integral kernel $a(x-y)$ of A is bounded and integrable. Surprisingly, these assumptions do not make the task of estimating of the heat kernel easier.

Since A is a bounded operator in $L^2(\mathbb{R}^d)$, its heat semigroup e^{tA} can be easily computed by using the exponential series that leads to

$$e^{tA} = e^{-t} e^{ta^*} = e^{-t} \sum_{k=0}^{\infty} t^k \frac{a^{*k}}{k!} = e^{-t} \text{Id} + e^{-t} \sum_{k=1}^{\infty} t^k \frac{a^{*k}}{k!}.$$

By removing the singular part $e^{-t} \text{Id}$ of the heat semigroup, we obtain the *regularized* heat kernel

$$v(x, t) = e^{-t} \sum_{k=1}^{\infty} t^k \frac{a^{*k}(x)}{k!} \quad (1.11)$$

with the source at the origin. In other words, for any $f \in L^2(\mathbb{R}^d)$, a solution to the non-local Cauchy problem

$$\partial_t u - Au = 0, \quad u|_{t=0} = f, \quad (1.12)$$

has the form $u(x, t) = e^{-t} f(x) + (v * f)(x, t)$ with v given by (1.11). In particular, the fundamental solution of the problem (1.12) is

$$u(x, t) = e^{-t} \delta(x) + v(x, t).$$

The function v is the main subject of this paper.

A probabilistic interpretation of the function $v(x, t)$ is of great interest. Under conditions (1.8), (1.9) the operator A defined in (1.7) is a generator of a continuous-time Markov jump process. If this process starts at zero, its transition probability has a regular part and a singularity at zero, and $v(x, t)$ is the density of the regular part. The results of this work allow us to describe the large-time behaviour of this Markov process in different regions of the space. In particular we obtain the local moderate and large deviations results for this Markov process.

In the recent years there is an essential progress in studying the large-time behaviour of solutions to evolution problems in \mathbb{R}^d for convolution-type operators with integrable kernels, see, for instance, [1, 6, 7], and the references therein. One of the key questions of interest here is obtaining pointwise estimates for the corresponding non-local heat kernels and solutions. To our best knowledge there are just few papers devoted to this topic. In [5] the asymptotic behaviour of fundamental solution for evolution equations with a convolution kernel has been considered. For Gaussian and compactly supported kernels that are radially symmetric, two-sided estimates have been obtained. Since [5] mostly deals with problems with unbounded initial conditions, the authors focus on the behaviour of heat kernel in the region of extra large $|x| \gg t$, and their estimates are rather loose in other regions. The kernels showing sub-exponential decay at infinity have been studied in [10], this work deals with the asymptotic behaviour of the fundamental solution in the region $|x| \gg t$.

Papers [18] deals with the large deviations asymptotics for a symmetric random walk on a multidimensional lattice under the assumptions that the transition intensities have super-light tails. The large-time asymptotics for the transition probabilities has been obtained in the regions of moderate and large deviations. Interesting results on extended large deviation principle for generalized Poisson processes can be found in the recent work [17]

Stationary problems for convolution-type operators are also of essential interest. A number of results on pointwise estimates for a resolvent kernel of non-local convolution-type operators have been obtained in the recent works [13, 15]. In [13] both polynomially and exponentially decaying kernels were considered, while [15] focused on the case of very light tails. Resolvent bounds for discrete operators have been used in [11] to show the intermittency phenomenon for branching walk with heavy tails.

The spectral analysis of the corresponding non-local Schrödinger operator has been done in [14, 15]. In particular, in these papers the question of existence of the ground state has been addressed. It should be noted that the pointwise estimates obtained in [13, 15] allow to derive pointwise bounds for the principal eigenfunction of non-local Schrödinger operator.

In the present paper we deal with convolution kernels $a(x)$ that decay at infinity at least exponentially and admit an estimate from above by a radially symmetric function: $a(x) \leq ce^{-b|x|^p}$ with $b > 0$ and $p \geq 1$.

The large-time behaviour of the studied heat kernel depends crucially on the relation between $|x|$ and t . We consider separately four different regions in (x, t) space, namely,

- (i) $|x| = O(t^{1/2})$, (ii) $t^{\frac{1}{2}} \ll |x| \ll t$, (iii) $|x| \sim t$, (iv) $|x| \gg t$.

In particular, it will be shown that in the region (iv) the function $-\ln v(x, t)$ behaves like $|x|(\ln \frac{|x|}{t})^{\frac{p-1}{p}}$ for $a(x) \sim e^{-b|x|^p}$ with $p \geq 1$, and like $|x| \ln(|x|/t)$ for $a(x)$ with a finite support.

Remark that for the corresponding Markov jump process with the generator defined in (1.7) the region (i) corresponds to the standard deviations where the local central limit theorem applies, (ii) is the region of the moderate deviations, (iii) is the region of large deviation, and (iv) should probably be called the ‘extra-large’ deviation region.

Before considering the case of generic convolution kernels with a light tail we first study the Gaussian kernels for which the k th convolution admits an explicit formula. This allows us to find the asymptotics of the corresponding heat kernel in all the regions mentioned above, see Theorem 2.1.

The Gaussian asymptotics of a generic non-local heat kernel in the region (i) is a consequence of the (local) central limit theorem. It is interesting to observe that in the region (ii) the logarithmic asymptotics of the non-local heat kernel still remains the same as for the classical heat kernel with the covariance matrix σ , see Theorem 3.1. The transition between Gaussian and non-Gaussian behaviour occurs in the region $x = rt$. For small r the behaviour is still close to Gaussian, while as $r \rightarrow \infty$ the asymptotics of the non-local heat kernel does not look like Gaussian at all, as shown in Theorems 3.4 and 3.8. The difference is getting even more drastic in the region $|x| \gg t$, see Theorem 3.2.

2. Gaussian convolution kernel

We consider in this section the case of a Gaussian convolution kernel:

$$a(x) = \frac{1}{(4\pi)^{d/2}} e^{-\frac{x^2}{4}}, \quad \hat{a}(p) = e^{-p^2}. \quad (2.1)$$

In this case the convolutions $a^{*k}(x)$ admit explicit formulae for all $k \geq 1$ which essentially simplify our analysis. The large-time asymptotics (or log asymptotics) of the fundamental solution depends essentially on the relation between x and t . We consider separately four different regions in (x, t) -space, namely, $|x| = O(t^{\frac{1}{2}})$ and $|x| \sim t^{\frac{1+\delta}{2}}$ with $0 < \delta < 1$, or $\delta = 1$, or $\delta > 1$.

Denote $\Phi_G(r) = 1 + 2\xi_r \ln \xi_r - \xi_r$, where ξ_r is a solution to the equation $\xi^2 \ln \xi = \frac{r^2}{4}$.

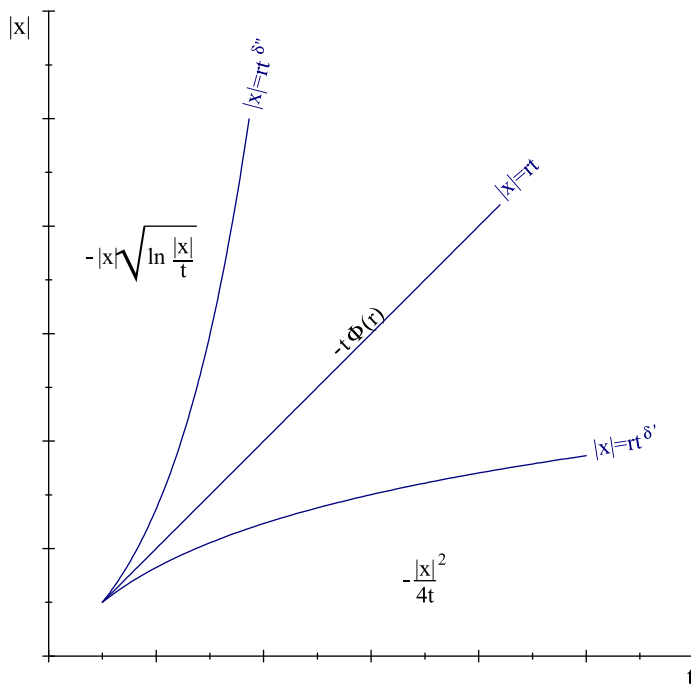


FIGURE 2.1 (colour online). The large-time behaviour of $\ln v(x, t)$ depends crucially on whether $|x| \ll t$ (under the lower curve), or $|x| \sim t$ (the middle curve), or $|x| \gg t$ (over the upper curve). Here $\delta' < 1$ and $\delta'' > 1$.

THEOREM 2.1 (Gaussian kernel). Let the convolution kernel $a(x)$ be defined by (2.1). Then for the function $v(x, t)$ defined by (1.11) the following asymptotics holds as $t \rightarrow \infty$ (see Figure 2.1).

(1) For any $r > 0$, if $|x| \leq rt^{\frac{1}{2}}$, then

$$v(x, t) = \frac{1}{(4\pi t)^{d/2}} e^{-\frac{x^2}{4t}} (1 + o(t^{-\frac{1}{4}})). \tag{2.2}$$

(2) For any $r > 0$, if $|x| = rt^{\frac{1+\delta}{2}}$ with $0 < \delta < 1$, then

$$\frac{\ln v(x, t)}{\frac{x^2}{4t}} \rightarrow -1. \tag{2.3}$$

In particular, if $r_1 t^{\frac{1+\delta}{2}} \leq |x| \leq r_2 t^{\frac{1+\delta}{2}}$ with some $0 < r_1 < r_2$ and $0 < \delta < 1$, then

$$e^{-\frac{r_2^2}{4} t^\delta (1+o(1))} \leq v(x, t) \leq e^{-\frac{r_1^2}{4} t^\delta (1+o(1))}.$$

(3) For any $r > 0$, if $|x| = rt$, then

$$\frac{\ln v(x, t)}{t} \rightarrow -\Phi_G(r). \tag{2.4}$$

Furthermore, the function $\Phi_G(r)$ possesses the following properties:

$$\begin{aligned} 0 < \Phi_G(r) < r^2/4 & \quad \text{for all } r \neq 0, \\ \Phi_G(r) = \frac{r^2}{4}(1 + o(1)) & \quad \text{as } r \rightarrow 0+ \\ \Phi_G(r) = r\sqrt{\ln r}(1 + o(1)) & \quad \text{as } r \rightarrow \infty \end{aligned}$$

(4) If $|x| > t^{\frac{1+\delta}{2}}$ with $\delta > 1$, then

$$\frac{\ln v(x, t)}{|x| \sqrt{\ln \frac{|x|}{t}}} \rightarrow -1. \tag{2.5}$$

COROLLARY 2.2. For any $r > 0$, if $|x| = rt^{\frac{1+\delta}{2}}$ with $\delta > 1$, then it follows from (2.5) that

$$\frac{\ln v(x, t)}{t^{\frac{\delta+1}{2}} \sqrt{\ln t}} \rightarrow -\tilde{c}(\delta, r), \quad \text{with} \quad \tilde{c}(\delta, r) = r \sqrt{\frac{\delta-1}{2}}.$$

REMARK 2.3. In the case $\delta > 1$ the function $\frac{x^2}{4t} = \frac{r^2}{4} t^\delta$ exhibits the faster polynomial growth at infinity than the function $t^{\frac{\delta+1}{2}} \sqrt{\ln t}$. Consequently, in this region the non-local heat kernel $v(x, t)$ has a more ‘fat’ tail $v(x, t) \sim e^{-\tilde{c}(\delta, r)t^{\frac{\delta+1}{2}} \sqrt{\ln t}}$ than the classical heat kernel $w(x, t) = \frac{1}{(4\pi)^{d/2} t^{d/2}} e^{-\frac{x^2}{4t}} \sim e^{-\frac{r^2}{4} t^\delta}$.

In the next sections we prove all statements of Theorem 2.1.

2.1. Asymptotics in the case $|x| \leq rt^{\frac{1}{2}}$

The asymptotics (2.2) follows from the local limit theorem for a general probability distribution that satisfies (1.8) and (1.9). To justify the estimate for the remainder in (2.2) we give a short analytic proof based on the following representation for $v(x, t)$

$$v(x, t) = \int_{\mathbb{R}^d} e^{ixp} \left(e^{-t(1-\hat{a}(p))} - e^{-t} \right) dp.$$

This integral can be rewritten as the following sum:

$$\begin{aligned} v(x, t) &= \int_{|p| < \sqrt{\ln 2t}} e^{ixp} e^{-t(1-\hat{a}(p))} dp - e^{-t} \int_{|p| < \sqrt{\ln 2t}} e^{ixp} dp \\ &+ \int_{|p| > \sqrt{\ln 2t}} e^{ixp} \left(e^{-t(1-\hat{a}(p))} - e^{-t} \right) dp. \end{aligned} \tag{2.6}$$

The second and the third integral in (2.6) can be estimated from above by $O(e^{-t}(\ln t)^{\frac{d}{2}})$ and $o(e^{-t}t)$ correspondingly. Denoting $1 - \hat{a}(p) = p^2 - p^4 f(p)$ and taking into account the relation $p^4 f(p) = O(-t^{\frac{d}{3}})$ valid for $|p| < t^{-\frac{1}{3}}$, for the first integral in (2.6) we get

$$\begin{aligned} \int_{|p| < \sqrt{\ln 2t}} e^{ixp} e^{-t(1-\hat{a}(p))} dp &= \int_{|p| < t^{-\frac{1}{3}}} e^{ixp} e^{-t(p^2 - p^4 f(p))} dp \\ &+ \int_{t^{-\frac{1}{3}} < |p| < \sqrt{\ln 2t}} e^{ixp} e^{-t(1-\hat{a}(p))} dp = \frac{1}{(4\pi)^{\frac{d}{2}} t^{\frac{d}{2}}} e^{-\frac{x^2}{4t}} \left(1 + o(t^{-\frac{1}{4}}) \right) + o(e^{-t\frac{1}{4}}). \end{aligned} \tag{2.7}$$

This yields (2.2).

2.2. The case $|x| = rt^{\frac{1+\delta}{2}}$, $0 < \delta \leq 1$

In this region we exploit the first representation for $v(x, t)$ in (1.11). Since

$$a^{*k}(x) = c_k e^{-\frac{x^2}{4k}} \quad \text{with } c_k = \frac{\tilde{c}}{k^{d/2}}, \tilde{c} > 0, \tag{2.8}$$

then using Stirling’s approximation we get

$$\frac{t^k a^{*k}(x)}{k!} = \exp \left\{ k \ln t - k \ln k + k - \frac{x^2}{4k} - c(d) \ln k + r_k \right\} \tag{2.9}$$

with a constant $c(d) = \frac{d+1}{2}$ and $|r_k| \leq C$. Let us estimate the maximal term in the sum

$$\sum_{k=1}^{\infty} \frac{t^k a^{*k}(x)}{k!}. \tag{2.10}$$

To this end we introduce a function

$$S(z, t) = \left(z \ln t - z \ln z + z - \frac{x^2}{4z} - c(d) \ln z \right) \Big|_{x^2=rt^{1+\delta}}, \quad z > 0, \tag{2.11}$$

and locate $\max_{z>0} S(z, t)$ in z for each $t > 0$. Since for each positive t the function $S(z, t)$ tends to $-\infty$ both as $z \rightarrow 0$ and as $z \rightarrow \infty$, it attains its maximum on $(0, +\infty)$. Denote

$$\hat{z} = \hat{z}(t) = \operatorname{argmax}_{z>0} S(z, t).$$

PROPOSITION 2.4. Let $\delta > 0$, then $\hat{z}(t) = t\hat{\xi}(t)$, where $\hat{\xi} = \hat{\xi}(t)$ is the solution of equation

$$\frac{4}{r^2} \xi^2 \ln \xi + \frac{4c(d)}{r^2 t} \xi = t^{\delta-1}. \tag{2.12}$$

Moreover, $\hat{\xi}(r, t) = \xi_r(t)(1 + o(1))$, $t \rightarrow \infty$, where $\xi_r(t)$ is the solution of equation

$$\frac{4}{r^2} \xi^2 \ln \xi = t^{\delta-1}. \tag{2.13}$$

Proof. The maximum point of $S(z, t)$ is defined by the equation

$$\frac{\partial}{\partial z} S(z, t) = \ln t - \ln z + \frac{x^2}{4z^2} - \frac{c(d)}{z} = 0. \tag{2.14}$$

Making the change of variables $z = t\xi$, we rewrite (2.14) as (2.12). Denote a solution of this equation by $\hat{\xi}(r, t)$. In what follows if it does not lead to ambiguity we drop the arguments of the function $\hat{\xi}(r, t)$. Observe that $\hat{\xi} > 1$ for sufficiently large t . Indeed, for $\xi \in (0, 1]$ we have $\xi^2 \ln \xi \leq 0$, and $\frac{\xi}{t} = o(t^{\delta-1})$ as $t \rightarrow \infty$. This yields the required inequality.

Note that the function on the left-hand side of (2.12) is increasing as $\xi \in (1, +\infty)$ and therefore, equation (2.12) has a unique solution for large t . It is easy to see that

$$\frac{4c(d)}{r^2 t} \xi = o(\xi^2 \ln \xi) \quad \text{if } \xi \rightarrow \infty \text{ and } t \geq 1,$$

and

$$\frac{4c(d)}{r^2 t} \xi = o(t^{\delta-1}) \quad \text{as } t \rightarrow \infty \text{ and } \xi \text{ is bounded.}$$

Consequently, the solution $\hat{\xi}(r, t)$ of (2.12) can be approximated for large t by the solution $\xi_r(t)$ of equation (2.13). □

We consider separately the following cases: $0 < \delta < 1$ and $\delta = 1$.

In the case $0 < \delta < 1$ we have $t^{\delta-1} \rightarrow 0$ as $t \rightarrow \infty$ and, therefore, the solution of (2.13) converges to 1. The Taylor expansion of $\frac{4}{r^2} \xi^2 \ln \xi$ about 1 reads

$$\frac{4}{r^2} \xi^2 \ln \xi = \frac{4}{r^2} (\xi - 1) + \frac{6}{r^2} (\xi - 1)^2 + O((\xi - 1)^3), \quad \xi - 1 \rightarrow 0.$$

Combining this expansion with (2.12) we obtain

$$\frac{4}{r^2} (\xi - 1) + \frac{6}{r^2} (\xi - 1)^2 = t^{\delta-1} - \frac{4c(d)}{r^2 t} + O((\xi - 1)^3 + (\xi - 1)t^{-1}).$$

The straightforward computations yield

$$\hat{\xi}(r, t) = 1 + \frac{r^2}{4} t^{\delta-1} - \left(\frac{3r^2}{8} t^{2(\delta-1)} + \frac{c(d)}{t} \right) + o\left(\max\{t^{2(\delta-1)}, t^{-1}\}\right)$$

and

$$\hat{z} = t\hat{\xi} = t + \frac{r^2}{4} t^\delta - \left(\frac{3r^2}{8} t^{2\delta-1} + c(d) \right) + o\left(\max\{t^{2\delta-1}, 1\}\right).$$

Substituting this expression for \hat{z} in (2.11) and considering the relation $x^2 = r^2 t^{1+\delta}$ we get

$$S(\hat{z}, t) = t - \frac{r^2}{4} t^\delta + \frac{r^4}{16} t^{2\delta-1} - c(d) \ln t + o\left(\max\{t^{2\delta-1}, \ln t\}\right). \tag{2.15}$$

Now from (1.11) and (2.9), taking into account the fact that $a^{*k}(x) > 0$ for all k and x , we obtain the following estimate of $v(x, t)$ from below:

$$\begin{aligned} v(x, t) &= e^{-t} \sum_{k=1}^{\infty} \frac{t^k a^{*k}(x)}{k!} \geq e^{-t+S(\hat{z}, t)} \\ &= e^{-\frac{r^2}{4} t^\delta + \frac{r^4}{16} t^{2\delta-1} - c(d) \ln t + o(\max\{t^{2\delta-1}, \ln t\})}, \quad \text{as } t \rightarrow \infty. \end{aligned} \tag{2.16}$$

To get an upper bound on $v(x, t)$ we divide the sum in (1.11) into two parts, in the first sum the summation index varies from 1 to n_0 where n_0 is chosen in such a way that

$$\frac{t^{n+1} e^{-\frac{x^2}{n+1}}}{(n+1)!} \frac{n!}{t^n e^{-\frac{x^2}{n}}} = \frac{t}{n+1} e^{\frac{x^2}{n(n+1)}} < \frac{1}{2} \quad \text{for all } n \geq n_0. \tag{2.17}$$

Using the relation $x^2 = r^2 t^{1+\delta}$ and the fact that $f(u) = ue^{cu^2}$ is an increasing function for any $c > 0$, we have

$$\frac{t}{n+1} e^{\frac{x^2}{n(n+1)}} < \frac{1}{3} e^{\frac{r^2}{9} t^{\delta-1}} < \frac{1}{2} \quad \text{for all } n > n_0 = [3t] \text{ and } t \geq \left(\frac{r^2}{9 \ln(3/2)} \right)^{\frac{1}{1-\delta}}.$$

This implies that

$$\sum_{n>3t} \frac{t^n a^{*n}(x)}{n!} \leq e^{S(\hat{z}, t)} = e^{t - \frac{r^2}{4} t^\delta + \frac{r^4}{16} t^{2\delta-1} - c(d) \ln t + o(\max\{t^{2\delta-1}, \ln t\})}, \tag{2.18}$$

as $t \rightarrow \infty$. Due to (2.9) and (2.11) the upper bound for the sum $\sum_{n=1}^{[3t]} \frac{t^n a^{*n}(x)}{n!}$ reads

$$\sum_{n=1}^{[3t]} \frac{t^n a^{*n}(x)}{n!} \leq Cte^{S(\hat{z}, t)} = e^{t - \frac{r^2}{4} t^\delta + \frac{r^4}{16} t^{2\delta-1} - (c(d)-1) \ln t + o(\max\{t^{2\delta-1}, \ln t\})}. \tag{2.19}$$

From (2.18) and (2.19) we derive the estimate of $v(x, t)$ from above as $t \rightarrow \infty$:

$$v(x, t) = e^{-t} \sum_{k=1}^{[3t]} \frac{t^k a^{*k}(x)}{k!} + e^{-t} \sum_{k>3t} \frac{t^k a^{*k}(x)}{k!} \tag{2.20}$$

$$\leq e^{-\frac{r^2}{4}t^\delta + \frac{r^4}{16}t^{2\delta-1} - (c(d)-1) \ln t + o(\max\{t^{2\delta-1}, \ln t\})}.$$

Finally from (2.16) and (2.20) we get (2.3).

REMARK 2.5. Since $x^2/4t = r^2t^\delta/4$, $0 < \delta < 1$, the logarithmic asymptotics of $v(x, t)$ coincides with that for the classic heat kernel:

$$\frac{\ln v(x, t)}{\frac{x^2}{4t}} \rightarrow -1 \quad \text{as } t \rightarrow \infty. \tag{2.21}$$

Moreover, for $\delta \leq \frac{1}{2}$ estimates (2.16) and (2.20) take the form

$$C_1 t^{\frac{1-d}{2}} e^{-\frac{r^2}{4}t^\delta} \leq v(x, t) \leq C_2 t^{\frac{1-d}{2}} e^{-\frac{r^2}{4}t^\delta} \tag{2.22}$$

with $C_1, C_2 > 0$. For $\delta \in (\frac{1}{2}, 1)$ estimates (2.16) and (2.20) imply that

$$v(x, t) = e^{-\frac{r^2}{4}t^\delta + \frac{r^4}{16}t^{2\delta-1} + o(t^{2\delta-1})}. \tag{2.23}$$

We proceed with the case $\delta = 1$. In this case equation (2.13) reads

$$\xi^2 \ln \xi = \frac{r^2}{4}. \tag{2.24}$$

It is easy to check that equation (2.24) has a unique solution $\xi_r \in (1, \infty)$. Then for solution $\hat{\xi}(r, t)$ of (2.12) by the implicit function theorem it follows that $\hat{\xi}(r, t) = \xi_r + O(t^{-1})$. Therefore,

$$S(\hat{z}, t) = t\xi_r \ln t - t\xi_r(\ln t + \ln \xi_r) + t\xi_r - t\xi_r \ln \xi_r + O(\ln t)$$

$$= t(\xi_r - 2\xi_r \ln \xi_r) + O(\ln t), \quad \hat{z} = t\hat{\xi},$$

and using the same arguments as above we have

$$v(x, t) = e^{-t(1+2\xi_r \ln \xi_r - \xi_r) + O(\ln t)}, \quad t \rightarrow \infty, \tag{2.25}$$

where $\xi_r > 1$ is the solution of (2.24). Thus the logarithmic asymptotics of $v(x, t)$ is given by

$$\frac{\ln v(x, t)}{t} \rightarrow -\Phi_G(r), \quad \text{as } t \rightarrow \infty, \tag{2.26}$$

where $\Phi_G(r) = \Phi_G(\xi_r) = 1 + 2\xi_r \ln \xi_r - \xi_r$.

LEMMA 2.6. For any $r > 0$

$$0 < \Phi_G(r) < \frac{r^2}{4}. \tag{2.27}$$

Moreover,

$$\Phi_G(r) = \frac{1}{4}r^2(1 + o(1)), \quad \text{as } r \rightarrow 0, \tag{2.28}$$

$$\Phi_G(r) = r\sqrt{\ln r}(1 + o(1)) \quad \text{as } r \rightarrow \infty.$$

Proof. Let $\gamma(\xi) = 1 + 2\xi \ln \xi - \xi$. To prove the lower bound in (2.27) we note that $\gamma(1) = 0$ and $\frac{\partial}{\partial \xi} \gamma(\xi) = 1 + 2 \ln \xi \geq 1$ (for $\xi \geq 1$). To prove that

$$1 + 2\xi_r \ln \xi_r - \xi_r < \frac{r^2}{4} = \xi_r^2 \ln \xi_r,$$

we denote by $\varkappa(\xi) = \xi^2 \ln \xi$. Then $\gamma(1) = \varkappa(1) = 0$, and $\gamma'(\xi) < \varkappa'(\xi)$, $\xi > 1$. This yields the desired upper bound in (2.27).

The asymptotics (2.28) is a particular case of Theorem 3.8 describing the asymptotic behaviour of $\Phi(r)$ under the general assumptions on the kernel $a(x)$. In our case $\Phi_G(r) = \Phi(\xi_r^{-1})$, and if we take $p = 2, b = \frac{1}{4}$, then $I(s) = \frac{1}{4}s^2$, and we immediately obtain (2.28) from (3.22). \square

2.3. The case $|x| \geq t^{\frac{1+\delta}{2}}$, $\delta > 1$

In this subsection we consider a region in (x, t) -space of super-large $|x|$, where $|x| > t^{(1+\delta)/2}$ with $\delta > 1$. In this case we again begin with the description of $\max_z \mathcal{S}(z, t, x)$, where

$$\mathcal{S}(z, t, x) = \left(z \ln t - z \ln z + z - \frac{x^2}{4z} - c(d) \ln z \right), \quad z > 0.$$

Since $t \rightarrow \infty$, we can omit the last term in (2.11), and write as above the following equation on $\hat{z} = \hat{z}(t, x) = \operatorname{argmax} S(z, t, x)$:

$$\frac{\partial}{\partial z} \mathcal{S}(\hat{z}, t, x) = \ln t - \ln \hat{z} + \frac{x^2}{4\hat{z}^2} = 0, \tag{2.29}$$

or equivalently,

$$x^2 = 4\hat{z}^2(\ln \hat{z} - \ln t) = 4\hat{z}^2 \ln \frac{\hat{z}}{t}. \tag{2.30}$$

Taking the logarithm on both sides of equation (2.30) we obtain

$$\ln \hat{z} = \ln |x| (1 + o(1)),$$

consequently equality (2.30) can be rewritten as

$$x^2 = 4\hat{z}^2 \ln \frac{|x|}{t} (1 + o(1)), \quad t \rightarrow \infty.$$

Substituting $\frac{|x|}{2\sqrt{\ln \frac{|x|}{t}}}$ for \hat{z} in (2.11), we get

$$\begin{aligned} S(\hat{z}, t, x) &= \frac{|x|}{2\sqrt{\ln \frac{|x|}{t}}} \ln t - \frac{|x|}{2\sqrt{\ln \frac{|x|}{t}}} \left(\ln |x| - \ln 2 - \frac{1}{2} \ln \ln \frac{|x|}{t} \right) \\ &\quad + \frac{|x|}{2\sqrt{\ln \frac{|x|}{t}}} - \frac{x^2 \sqrt{\ln \frac{|x|}{t}}}{2|x|} + O(\ln |x|) = -|x| \sqrt{\ln \frac{|x|}{t}} (1 + o(1)), \quad t \rightarrow \infty. \end{aligned}$$

Since $|x| > t^{(1+\delta)/2}$ with $\delta > 1$, we can take $n_0 = |x|$ in (2.17) for large enough t . Then as above we get the following two-sided estimate on $v(x, t)$:

$$e^{-t} e^{-|x| \sqrt{\ln \frac{|x|}{t} (1+o(1))}} \leq v(x, t) \leq |x| e^{-t} e^{-|x| \sqrt{\ln \frac{|x|}{t} (1+o(1))}}.$$

Since $t = o(|x|)$, this yields

$$\frac{\ln v(x, t)}{|x| \sqrt{\ln \frac{|x|}{t}}} \rightarrow -1, \quad \text{as } t \rightarrow \infty, \quad |x| > t^{(1+\delta)/2}, \delta > 1. \tag{2.31}$$

Conclusions

(1) If $|x| \leq rt^{1/2}$, then the main term of the asymptotics of $v(x, t)$ coincides with the classical heat kernel $p_t(x, 0)$ defined by (1.1).

(2) If $|x| \leq rt^{\frac{1}{2} + \frac{\delta}{2}}$, $0 < \delta < 1$, then the main term of the logarithmic asymptotics of $v(x, t)$ coincides with that of the classical heat kernel.

(3) If $|x| = rt$, then the leading term of the logarithmic asymptotics of $v(x, t)$ is a linear function $\Phi_G(r)t$. The leading term of the logarithmic asymptotics of the classical heat kernel is also a linear function $(r^2/4)t$. However, the corresponding coefficient $\Phi_G(r)$ is strictly less than $r^2/4$ for all $r > 0$. This reflects the fact that in this range of x the non-local heat kernel has more heavy tail than the classical one. It should also be noted that the coefficient $\Phi_G(r)$ is close to $r^2/4$ for small r while $\Phi_G(r) \ll r^2$ for large r .

(4) If $|x| \geq rt^{1+\delta}$, $\delta > 1$, then the main term of the logarithmic asymptotics of $v(x, t)$ given by (2.31) differs essentially from the logarithmic asymptotics of the classical heat kernel, in particular $v(x, t)$ has more heavy tail, than the classical heat kernel.

3. Kernels with generic light tails

3.1. Main results

In this section we consider generic non-local operators with convolution kernels that have light tails at infinity. More precisely, we assume that, in addition to (1.8)–(1.9), the convolution kernel $a(x)$ satisfies for some $p \geq 1$ the following condition

$$0 \leq a(x) \leq C_1 e^{-b|x|^p}, \tag{3.1}$$

or even a more strong condition

$$a(x) \in C_0(\mathbb{R}^d), \quad \text{supp } a(x) \subset K_\mu = \{x \in \mathbb{R}^d : |x| \leq \mu\} \quad \text{for some } \mu > 0. \tag{3.2}$$

From now on we assume that μ is chosen in the optimal way, that is, $\mu = \min\{\tilde{\mu} > 0 : \text{supp } a \subset K_{\tilde{\mu}}\}$.

Since, in contrast with the Gaussian case, here a^{*k} do not admit an explicit formula for $k \geq 1$, we have to obtain sharp enough estimates for these higher order convolutions. To this end we first use the results on the asymptotic behaviour of distributions of the sums of i.i.d. random variables, such as the local central limit theorem and the large deviations principle, and then combine these results with analytic techniques in order to obtain the asymptotics for $v(x, t)$.

As in the previous section, for large t four different regions of x are considered:

- (1) $|x| \leq rt^{1/2}(1 + o(1))$ (standard deviations region);
- (2) $|x| = rt^{\frac{1+\delta}{2}}(1 + o(1))$, $0 < \delta < 1$ (moderate deviations region);
- (3) $|x| = rt(1 + o(1))$ ($\delta = 1$) (large deviations region);
- (4) $|x| = rt^{\frac{1+\delta}{2}}(1 + o(1))$, $\delta > 1$ ('extra-large' deviations region).

The next two theorems describe the asymptotic behaviour of $v(x, t)$ in the regions 1, 2, and 4.

THEOREM 3.1 (The regions of standard and moderate deviations). *Assume that $a(x)$ satisfies (1.8)–(1.9) and (3.1). Then for the function $v(x, t)$ the following asymptotic relations hold as $t \rightarrow \infty$:*

(1) *if $|x| \leq rt^{\frac{1}{2}}$ for some $r > 0$, then*

$$v(x, t) = \frac{c(\sigma)}{t^{\frac{d}{2}}} e^{-\frac{(\sigma^{-1}x, x)}{2t}} (1 + o(1)), \tag{3.3}$$

where $c(\sigma) = (2\pi)^{-\frac{d}{2}} |\det(\sigma)|^{-\frac{1}{2}}$, σ is the covariance matrix of the distribution $a(x)$;

(2) *if $x = rt^{\frac{1+\delta}{2}}(1 + o(1))$ with $0 < \delta < 1$ and $r \in R^d \setminus \{0\}$, then*

$$v(x, t) = e^{-\frac{(\sigma^{-1}x, x)}{2t}(1+o(1))} = e^{-\frac{1}{2}(\sigma^{-1}r, r) t^\delta(1+o(1))}. \tag{3.4}$$

THEOREM 3.2 (The regions of extra-large deviations). *Assume that $a(x)$ satisfies (1.8)–(1.9) and (3.1). Then for $|x| = rt^{\frac{1+\delta}{2}}(1 + o(1))$ with $\delta > 1$ and $r > 0$ the following asymptotic upper bound holds:*

$$v(x, t) \leq e^{-c_p t^{\frac{\delta+1}{2}} (\ln t)^{\frac{p-1}{p}} (1+o(1))}, \quad \text{as } t \rightarrow \infty, \tag{3.5}$$

where the constant $c_p = c_p(b, r)$ depends on b, r and p .

If in addition $a(x)$ satisfies (3.2), then for $|x| = rt^{\frac{\delta+1}{2}}(1 + o(1))$ with $\delta > 1$

$$v(x, t) \leq e^{-\tilde{c}(\mu) t^{\frac{\delta+1}{2}} \ln t(1+o(1))}, \quad \text{as } t \rightarrow \infty, \tag{3.6}$$

where $\tilde{c}(\mu) = \frac{(\delta-1)r}{2\mu}$.

In the region $x \sim t$, usually called large deviations region, our approach relies essentially on the properties of the rate function $I(r)$ of the sum of i.i.d. random variables. From now on S_k stands for the sum of i.i.d. random variables (vectors) X_1, \dots, X_k with common distribution $a(x)$. From (3.1) it follows that the random variables X_j have exponential moment $\Lambda(\gamma) = \mathbb{E}e^{\gamma X_1}$ for all γ from a neighbourhood of 0 (the so-called Cramer condition). Under this condition the large deviation principle holds for S_k with a rate function

$$I(r) = \sup_{\gamma} (\gamma \cdot r - L(\gamma)), \quad r, \gamma \in R^d, \tag{3.7}$$

where $I(r)$ is the Legendre transform of the cumulant generating function $L(\gamma) = \ln \Lambda(\gamma)$, and $\gamma \cdot r$ stands for the scalar product in R^d .

In order to formulate the main result of this section we denote by ξ_r a positive solution of the equation

$$\ln \xi = I(\xi r) - \xi r \cdot \nabla I(\xi r), \quad \xi \in R, \tag{3.8}$$

and introduce the function

$$\Phi(r) = 1 - \frac{1}{\xi_r} (1 + \ln \xi_r - I(\xi_r r)). \tag{3.9}$$

Equation (3.8) has a unique solution $\xi_r > 0$ for any $r \in R^d \setminus \{0\}$, moreover $0 < \xi_r < 1$, see Lemma 3.11.

We introduce now additional technical conditions on the kernel.

(A₁) in the case $p = 1$ for any $b_1 > b$ and any $\theta \in S^{d-1}$

$$\mathbb{E}e^{b_1 X \cdot \theta} = \infty, \tag{3.10}$$

where b is the same constant as in (3.1).

(\mathbf{A}_1^s) in the case $p = 1$ for any $\theta \in S^{d-1}$

$$\mathbb{E}|X|e^{bX \cdot \theta} = \infty.$$

(\mathbf{A}_p) in the case $p > 1$

$$L(\gamma) = \ln \mathbb{E}e^{\gamma \cdot X} = C(b, p)|\gamma|^{p/(p-1)}(1 + o(1)), \quad \text{as } |\gamma| \rightarrow \infty, \tag{3.11}$$

where $C(b, p) = \frac{p-1}{p}(bp)^{-1/(p-1)}$ is a constant appearing in the logarithmic asymptotics of the Laplace transform of $e^{-b|x|^p}$.

REMARK 3.3. Condition \mathbf{A}_p , $p \geq 1$, can be treated as a sort of soft lower bound for $a(x)$. In particular, it holds if $a(x)$ satisfies the following two-sided estimate

$$C_2e^{-b|x|^p} \leq a(x) \leq C_1e^{-b|x|^p}, \quad p \geq 1.$$

Observe also that under condition \mathbf{A}_p , $p \geq 1$, the function $a(x)$ cannot satisfy (3.2).

It should be emphasized that in the case $p = 1$ conditions $\mathbf{A}_1, \mathbf{A}_1^s$ are required for proving the main result on the asymptotics of the heat kernel, while in the case $p > 1$ condition \mathbf{A}_p is only used for determining the asymptotic behaviour of the function $\Phi(r)$ for large r .

THEOREM 3.4 (Asymptotic upper bounds). *Let conditions (1.8)–(1.9) and (3.1) be fulfilled, and assume additionally that in the case $p = 1$ condition \mathbf{A}_1 holds. Then for any $r \in \mathbb{R}^d \setminus \{0\}$ and for $x = rt(1 + o(1))$ the following asymptotic estimate holds as $t \rightarrow \infty$:*

$$v(x, t) \leq e^{-\Phi(r)t(1+o(1))}, \tag{3.12}$$

where the function $\Phi(r)$ is defined by (3.9).

Moreover, $\Phi(0) = 0$, $\Phi(r) > 0$, if $r \neq 0$, Φ is a convex function, and the following limit relations hold:

$$\Phi(r) = \frac{1}{2} \sigma^{-1} r \cdot r (1 + o(1)), \quad \text{as } r \rightarrow 0; \tag{3.13}$$

$$\Phi(r) \rightarrow \infty, \quad \text{as } r \rightarrow \infty. \tag{3.14}$$

If $p = 1$, then

$$\Phi(r) = b|r|(1 + o(1)), \quad \text{as } |r| \rightarrow \infty. \tag{3.15}$$

If $p > 1$ and condition \mathbf{A}_p holds, then

$$\Phi(r) = \frac{p}{p-1}(b(p-1))^{1/p}|r|(\ln|r|)^{\frac{p-1}{p}}(1 + o(1)), \quad \text{as } |r| \rightarrow \infty. \tag{3.16}$$

If condition (3.2) holds, then

$$\Phi(r) \geq \frac{1}{\mu} |r| \ln|r| \quad \text{as } |r| \rightarrow \infty. \tag{3.17}$$

REMARK 3.5. Note that under the assumptions of Theorem 3.4 the function $\Phi(r)$ need not be isotropic. This is illustrated by formula (3.13). However, the additional condition \mathbf{A}_p ensures that for large $|r|$ the principal term of the asymptotics of $\Phi(r)$ is radially symmetric, see (3.15)–(3.16).

COROLLARY 3.6 (Spherically symmetric kernels). *Let $a(x) = a(|x|)$, $x \in \mathbb{R}^d$, be a spherically symmetric kernel satisfying all the conditions of Theorem 3.4. Then for any $s > 0$ and for $|x| = st(1 + o(1))$ the following asymptotic estimate holds as $t \rightarrow \infty$:*

$$v(x, t) \leq e^{-\Phi(s)t(1+o(1))}, \tag{3.18}$$

where the function $\Phi(s)$ is defined in (3.9), and formulae (3.13)–(3.17) from Theorem 3.4 take an easier form, namely

$$\Phi(s) = \frac{s^2}{2\sigma}(1 + o(1)), \quad \text{as } s \rightarrow 0; \tag{3.19}$$

$$\Phi(s) \rightarrow \infty, \quad \text{as } s \rightarrow \infty \tag{3.20}$$

If $p = 1$, then

$$\Phi(s) = bs(1 + o(1)), \quad s \rightarrow \infty. \tag{3.21}$$

If $p > 1$ and in addition conditions \mathbf{A}_p hold, then

$$\Phi(s) = \frac{p}{p-1}(b(p-1))^{1/p} s(\ln s)^{\frac{p-1}{p}}(1 + o(1)), \quad s \rightarrow \infty. \tag{3.22}$$

If (3.2) holds, then

$$\Phi(s) \geq \frac{1}{\mu} s \ln s \quad \text{as } s \rightarrow \infty. \tag{3.23}$$

REMARK 3.7. Observe that relation (3.18) and the asymptotics in (3.19) and (3.23) of Φ coincide with the estimates of the heat kernel on graphs [8] (see (1.5)–(1.6)).

Using another approach that relies on some exponential transformation of the random variable with density $a(x)$ under slightly more strong condition (for $p = 1$) we can show that the upper bound obtained in Theorem 3.4 gives in fact the large-time asymptotics of the fundamental solution. The following statement holds.

THEOREM 3.8 (Large-time asymptotics). *Let conditions (1.8)–(1.9) and (3.1) be fulfilled, and assume additionally that in the case $p = 1$ condition \mathbf{A}_1^s holds. Then for any $r \in R^d \setminus \{0\}$ and $x = rt(1 + o(1))$*

$$v(x, t) = e^{-\Phi(r)t(1+o(1))} \quad \text{as } t \rightarrow \infty, \tag{3.24}$$

where the function $\Phi(r)$ is defined by (3.9) and possesses all the properties enumerated in Theorem 3.4.

3.2. Properties of $I(r)$ and $\Phi(r)$

We preface the proof of the theorems by a number of technical statements. We discuss in this section the asymptotic properties of the function $I(r)$ defined by (3.7) that will be used further in the analysis of the function $v(x, t)$ in the regions of moderate and large deviations. Due to the symmetry of $a(x)$ stated in (1.8) the functions $I(r)$ and $L(\gamma)$ are symmetric with respect to zero, that is $I(-r) = I(r)$ and $L(-\gamma) = L(\gamma)$. We denote by \mathcal{A} the convex hull of the support of $a(\cdot)$. From our conditions (1.8)–(1.9) it follows that \mathcal{A} contains a neighbourhood of zero. Note that the set \mathcal{A} is symmetric with respect to the origin.

First we consider the one-dimensional case. In this case, $\mathcal{A} = [\inf \text{supp } a, \sup \text{supp } a] = [-\mu, \mu]$.

PROPOSITION 3.9 (One-dimensional case).

(1) *For any distribution $a(x)$ satisfying (1.8)–(1.9) and (3.1) we have*

$$I(s) = \frac{s^2}{2\sigma}(1 + o(1)) \quad \text{as } s \rightarrow 0. \tag{3.25}$$

(2) If the distribution $a(x)$ in addition satisfies condition \mathbf{A}_p , $p \geq 1$, then $I(s)$ has the following asymptotics as $s \rightarrow \infty$:

$$I(s) = bs(1 + o(1)), \quad \text{if } p = 1; \quad I(s) = bs^p(1 + o(1)), \quad \text{if } p > 1, \quad (3.26)$$

where b is the same constant as in (3.1). If $p > 1$ then

$$\lim_{s \rightarrow \infty} \frac{I(s)}{|s|} = +\infty. \quad (3.27)$$

(3) If the distribution $a(x)$ in addition satisfies (3.2), then $I(s)$ is a smooth function on $(-\mu, \mu)$,

$$I(s) \rightarrow \infty \quad \text{as } s \rightarrow \mu - 0 \quad \text{or } s \rightarrow -\mu + 0,$$

and $I(s) = \infty$ if $|s| \geq \mu$.

Proof. (1) By the definition of $I(s)$ considering the smoothness of $L(\gamma)$ in the vicinity of zero we have $I(s) = s\gamma^* - L(\gamma^*)$, where $\gamma^* = \gamma^*(s)$ is the solution of equation $s = L'(\gamma)$. Using the Taylor decomposition for $L'(\gamma)$ about zero by the implicit function theorem we obtain $\gamma^* = \frac{s}{L''(0)}(1 + o(1)) = \frac{s}{\sigma}(1 + o(1))$ for small enough s . Consequently,

$$I(s) = s\gamma^* - L(\gamma^*) = \frac{s^2(1 + o(1))}{\sigma} - \frac{1}{2}\gamma^{*2}L''(0) = \frac{s^2}{2\sigma}(1 + o(1)).$$

(2) In the case $p = 1$ conditions (3.1) and (3.10) on the distribution $a(x)$ imply that $\Lambda(b_1) = \infty$ for any $b_1 > b$, and $\Lambda(b')$ is finite for all $0 < b' < b$. Therefore, for $s \geq 0$

$$I(s) < bs, \quad I'(s) \leq b, \quad \text{and} \quad \lim_{s \rightarrow \infty} I'(s) \leq b. \quad (3.28)$$

The last limit exists since $I'(s)$ is monotone and bounded. If we assume that $\lim_{s \rightarrow \infty} I'(s) = a < b$, then taking $\beta = \frac{a+b}{2}$ we get

$$I(s) = \sup_{\gamma < b} (\gamma s - L(\gamma)) > \beta s - L(\beta) = \beta s(1 + o(1)) \quad \text{as } s \rightarrow \infty. \quad (3.29)$$

On the other hand, if $\lim_{s \rightarrow \infty} I'(s) = a < \beta$, then $I(s) < as(1 + o(1))$ as $s \rightarrow \infty$, which contradicts (3.29). Thus $\lim_{s \rightarrow \infty} I'(s) = b$, and the first formula in (3.26) follows.

In the case $p > 1$ due to (3.11) the solution γ^* of equation $L'(\gamma) = s$ has the asymptotics $\gamma^* = bp s^{p-1}(1 + o(1))$ as $s \rightarrow \infty$, and thus

$$I(s) = \sup_{\gamma} (s\gamma - L(\gamma)) = s\gamma^* - L(\gamma^*) = bs^p(1 + o(1)).$$

Limit relation (3.27) follows from the fact that for $p > 1$ the function $L(\gamma)$ is finite for all $\gamma \in \mathbb{R}$. Then for any $N > 0$

$$I(x) = \sup_{\gamma} (s\gamma - L(\gamma)) \geq sN - L(N),$$

and thus $\liminf_{s \rightarrow \infty} \frac{I(s)}{s} \geq N$, which yields (3.27).

(3) Since $\text{supp } a \subset [-\mu, \mu]$, then, for $\gamma \geq 0$, $\Lambda(\gamma) = \int e^{\gamma x} a(x) dx \leq e^{\gamma \mu}$. Consequently, $L(\gamma) \leq \gamma \mu$, and for $s > \mu$ we have

$$I(s) = \sup_{\gamma} (\gamma s - L(\gamma)) \geq \sup_{\gamma} (s - \mu)\gamma = \infty.$$

On the other hand, since $\mu = \text{supp } a$, for $\gamma \geq 0$ and for any $\delta > 0$

$$\Lambda(\gamma) = \int e^{\gamma x} a(x) dx \geq c_{\delta} e^{\gamma(\mu - \delta)}$$

for some $c_\delta > 0$. Thus,

$$L(\gamma) \geq \ln c_\delta + \gamma(\mu - \delta) \quad \text{and} \quad I(s) \leq \sup_\gamma (s - (\mu - \delta))\gamma - \ln c_\delta < \infty,$$

if $s < \mu - \delta$. Since we take an arbitrary $\delta > 0$, then $I(s)$ is finite for all $s \in (-\mu, \mu)$. The smoothness of $I(s)$ follows from the standard convexity arguments.

It remains to prove that $I(s) \rightarrow \infty$ as $s \rightarrow \mu - 0$. Since $a(x) \leq C_1$, we have

$$\Lambda(\gamma) \leq C_1 \int_{-\mu}^{\mu} e^{\gamma x} dx = \frac{C_1}{\gamma} (e^{\mu\gamma} - e^{-\mu\gamma}) < \frac{C_1}{\gamma} e^{\mu\gamma}.$$

Then

$$L(\gamma) < -\ln \gamma + \gamma\mu + \ln C_1,$$

and

$$I(s) \geq \sup_\gamma ((s - \mu)\gamma + \ln \gamma) - \ln C_1 \geq (s - \mu)\gamma^*(s) + \ln \gamma^*(s) - \ln C_1,$$

where $\gamma^*(s) = \frac{1}{\mu - s}$ is the argmax of the function $(s - \mu)\gamma + \ln \gamma$. Since $\gamma^*(s) \rightarrow \infty$, as $s \rightarrow \mu - 0$, then

$$I(s) \geq \ln \gamma^*(s) - \tilde{C} \rightarrow +\infty, \quad \text{as } s \rightarrow \mu - 0.$$

The statement for negative s follows from the symmetry of a . □

Next we describe the properties of the rate function $I(r)$ in the multidimensional case.

PROPOSITION 3.10 (Multidimensional case).

(1) For any distribution $a(x)$ satisfying (1.8)–(1.9) and (3.1) we have

$$I(r) = \frac{1}{2} \sigma^{-1} r \cdot r (1 + o(1)) \quad \text{as } r \rightarrow 0. \tag{3.30}$$

(2) If $p = 1$, and in addition to the above conditions \mathbf{A}_1 is fulfilled, then $I(r)$ has the following asymptotics:

$$I(r) = b|r|(1 + o(1)), \quad \nabla I(r) = b \frac{r}{|r|} (1 + o(1)), \quad \text{as } |r| \rightarrow \infty, \tag{3.31}$$

where b is the same constant as in (3.1). Moreover, $|\nabla I(r)| \leq b$ for all $r \in R^d$. If $p > 1$, and in addition to (1.8)–(1.9) and (3.1) the function $a(x)$ satisfies condition \mathbf{A}_p , then

$$I(r) = b|r|^p(1 + o(1)), \quad \text{as } |r| \rightarrow \infty. \tag{3.32}$$

(3) If (3.2) holds, then $I(r)$ is a smooth function in the interior of the convex hull \mathcal{A} . Moreover, $I(r) \rightarrow \infty$ as $\text{dist}(r, \partial\mathcal{A}) \rightarrow 0$, and $I(r) = \infty$ for all $r \in R^d \setminus \mathcal{A}$.

Proof. The proof of this proposition is mostly based on the same arguments as the proof of Proposition 3.9.

(1) Using the Taylor decomposition for $L(\gamma)$ about zero we obtain as above $\gamma^* = (\nabla\nabla L(0))^{-1}r(1 + o(1))$ for small enough r . Consequently,

$$\begin{aligned} I(r) &= (\nabla\nabla L(0))^{-1}r \cdot r - \frac{1}{2}\nabla\nabla L(0)\gamma^* \cdot \gamma^* + o(r^2) \\ &= \frac{1}{2}(\nabla\nabla L(0))^{-1}r \cdot r + o(r^2) = \frac{1}{2}\sigma^{-1}r \cdot r + o(r^2), \end{aligned}$$

since $\sigma = \nabla\nabla L(0)$, and the asymptotics (3.30) follows.

(2) In the case $p = 1$ conditions (3.1) and (3.10) on the distribution $a(x)$ imply that for any $\theta \in S^{d-1}$

$$\Lambda(b_1\theta) = \mathbb{E}e^{b_1\theta \cdot X} = \infty, \quad \text{if } b_1 > b \quad \text{and} \quad \Lambda(b_1\theta) < \infty \quad \text{if } b_1 < b.$$

Therefore,

$$I(r) = \sup_{\gamma \in R^d} (r \cdot \gamma - L(\gamma)) = \sup_{|\gamma| \leq b} (r \cdot \gamma - L(\gamma)) < |r|b. \tag{3.33}$$

The function $I(s\theta)$ is a convex function of $s \in R^1$ for any $\theta \in S^{d-1}$. Consequently, (3.33) implies inequality

$$|\nabla I(r)| \leq b \quad \forall r \in R^d. \tag{3.34}$$

In the same way as in Proposition 3.9 using the convexity of $I(s\theta)$ we obtain

$$b|r|(1 + o(1)) \leq I(r) \leq b|r|, \quad \text{and} \quad \left(\nabla I(r) \cdot \frac{r}{|r|} \right) \rightarrow b \quad \text{as } |r| \rightarrow \infty.$$

Combining the last relation with (3.34) we obtain the second equality in (3.31).

In the case $p > 1$ considering the convexity of $L(\gamma)$ with the help of the implicit function theorem we get that the solution $\gamma^* \in R^d$ of equation $\nabla L(\gamma) = r$ has the asymptotics

$$\gamma^* = bp|r|^{p-2}r(1 + o(1)) \quad \text{as } r \rightarrow \infty.$$

This implies (3.32).

(3) Denote by $G(r)$ the following auxiliary function:

$$G(r) = \begin{cases} 0, & r \in \mathcal{A}, \\ +\infty, & r \notin \mathcal{A}. \end{cases}$$

Then the Legendre transform of G is equal to $G^*(\gamma) = \mu\left(\frac{\gamma}{|\gamma|}\right)|\gamma|$, where

$$\mu(\theta) = \sup_{r \in \mathcal{A}} r \cdot \theta = \sup_{r \in \text{supp } a} r \cdot \theta, \quad \theta \in S^{d-1}.$$

In the same way as in the proof of Proposition 3.9 one can show that

$$L(\gamma) = \ln \mathbb{E} e^{\gamma \cdot X} = \mu\left(\frac{\gamma}{|\gamma|}\right)|\gamma|(1 + o(1)), \quad |\gamma| \rightarrow \infty, \tag{3.35}$$

and moreover,

$$L(\gamma) \leq \mu\left(\frac{\gamma}{|\gamma|}\right)|\gamma| - \ln|\gamma| + C \tag{3.36}$$

for some constant C .

Since $G^{**}(r) = G(r)$, comparing (3.35) with $G^*(\gamma)$ we conclude that $I(r) = +\infty$ in $R^d \setminus \mathcal{A}$ and $I(r) < \infty$ for r in the interior of \mathcal{A} . The fact that $I(r) \rightarrow \infty$ as $\text{dist}(r, \partial\mathcal{A}) \rightarrow 0$ can be justified in the same way as in the proof of Proposition 3.9 using inequality (3.36). \square

LEMMA 3.11. *Let $a(x)$ satisfy (1.8)–(1.9) and (3.1). Then for any $r \in R^d \setminus \{0\}$ equation (3.8) has a unique solution ξ_r and $0 < \xi_r < 1$.*

Proof. If the convex hull \mathcal{A} of $\text{supp } a$ coincides with R^d , then differentiating the right-hand side of (3.8) in ξ we obtain

$$r \cdot \nabla I(\xi r) - r \cdot \nabla I(\xi r) - \xi r \cdot \nabla \nabla I(\xi r)r = -\xi r \cdot \nabla \nabla I(\xi r)r \leq 0$$

because of convexity of I ; here $\nabla \nabla$ denotes the Hessian. Moreover, for sufficiently small ξ we have $-r \cdot \nabla \nabla I(\xi r)r < 0$. Thus the function on the right-hand side of (3.8) is decreasing in ξ , and, since $I(0) = 0$, we immediately conclude that (3.8) has a unique solution and $0 < \xi_r < 1$.

If $\mathcal{A} \neq R^d$, but the ray $\{sr\}_{s \geq 0}$ lies inside \mathcal{A} , then we use the same arguments as above. If the ray $\{sr\}_{s \geq 0}$ intersects $\partial \mathcal{A}$ at a point s_r^* , then it follows from Proposition 3.10 that $I(sr) \rightarrow \infty$ as $s \rightarrow s_r^* - 0$. In addition, the convexity of $I(r)$ and the Newton–Leibniz formula imply

$$\lim_{s \rightarrow s_r^* - 0} \frac{\frac{d}{ds} I(sr)}{I(sr)} = \infty.$$

Consequently,

$$I(sr) - sr \cdot \nabla I(sr) = I(sr) - s \frac{d}{ds} I(sr) \rightarrow -\infty \quad \text{as } s \rightarrow s_r^* - 0,$$

and again we obtain the unique solution $0 < \xi_r < 1$ of equation (3.8). □

PROPOSITION 3.12. *The function $\Phi(r)$ is a convex function, $\Phi(0) = 0$, and $\Phi(r) > 0$ for any $r \in R^d \setminus \{0\}$. Moreover, if $a(x)$ satisfies (3.1) with $p \geq 1$ and, in the case $p = 1$, also condition \mathbf{A}_1^s , then Φ is strictly convex: $\nabla \nabla \Phi(r)r \cdot r > 0$.*

Proof. If $r = 0$, then (3.8) implies that $\xi_0 = 1$, and $\Phi(0) = 0$. Let us show that $\nabla \Phi(r) \cdot r > 0$ for any $r \in R^d \setminus \{0\}$. Indeed, $\Phi(r) = \Phi(\xi(r))$ with $\xi(r) = \xi_r$, then using (3.8) and considering the properties of $I(r)$ we have

$$\nabla \Phi(r) = \frac{\nabla \xi(r)}{\xi^2(r)} [\ln \xi(r) - I(\xi(r)r) + \xi(r)r \cdot \nabla I(\xi(r)r)] + \nabla I(\xi(r)r) = \nabla I(\xi(r)r).$$

Consequently, $\nabla \Phi(r) \cdot r = \nabla I(\xi(r)r) \cdot r > 0$ and $\Phi(r) > 0$ for any $r \in R^d \setminus \{0\}$.

To prove the convexity of Φ we differentiate equation (3.8) in r and obtain

$$\nabla \xi(r) = -\xi^2(r) \nabla \nabla I(\xi(r)r) r [\xi(r) + r \cdot \nabla \xi(r)].$$

The assumption $\nabla \xi(r) \cdot r > 0$ leads to a contradiction. Therefore, $\nabla \xi(r) \cdot r \leq 0$ and $\xi(r) + r \cdot \nabla \xi(r) \geq 0$ for all $r \in R^d$. This yields the inequality $\nabla \nabla \Phi(r)r \cdot r \geq 0$. Additionally, $\nabla \nabla I(r)r \cdot r > 0$ and $\xi(r) + r \cdot \nabla \xi(r) > 0$ in the case $p > 1$ or $p = 1$ under condition \mathbf{A}_1^s . This yields a strict convexity of Φ . □

PROPOSITION 3.13 (Skewed distribution).

(1) *Let distribution $a(x)$ satisfy (1.8)–(1.9) and (3.1). Assume also that in the case $p = 1$ condition \mathbf{A}_1^s is fulfilled and in the case $p > 1$ the following condition holds:*

$$\int_{x \cdot \theta > N} a(x) dx > 0 \quad \text{for any } N > 0 \text{ and any } \theta \in S^{d-1}. \tag{3.37}$$

Then for any $x^* \in R^d$ equation

$$\nabla L(\gamma) = x^* \tag{3.38}$$

has a unique solution $\gamma^* \in R^d$ and, furthermore, the following relations hold:

$$I(x^*) = x^* \cdot \gamma^* - L(\gamma^*) \tag{3.39}$$

$$\frac{1}{\Lambda(\gamma^*)} \int xa(x)e^{\gamma^* \cdot x} dx = x^*. \tag{3.40}$$

Moreover, denoting $a_\gamma(x) = \frac{a(x)e^{\gamma \cdot x}}{\Lambda(\gamma)}$ we get

$$a^{*k}(kx^*) = a_{\gamma^*}^{*k}(kx^*)e^{-I(x^*)k}. \tag{3.41}$$

(2) If distribution $a(x)$ satisfies (3.1) (or (3.2)), then for any small enough $x^* \in R^d$ equation (3.38) has a unique solution γ^* , and relations (3.39)–(3.41) hold.

Proof. (1) Assume first that $p > 1$. From the properties of the function L it follows that $\nabla L : R^d \rightarrow R^d$ is a semicontinuous strictly monotone operator, that is, $(\nabla L(\gamma_1) - \nabla L(\gamma_2)) \cdot (\gamma_1 - \gamma_2) > 0$. Condition (3.37) implies that $\Lambda(\gamma) > e^{N\|\gamma\|}c(N)$ with $c(N) > 0$ for any N , and consequently

$$L(\gamma) > \frac{1}{2}N\|\gamma\| \quad \text{for all large enough } \|\gamma\|. \tag{3.42}$$

Since L is a convex function, then $\frac{\nabla L(\gamma) \cdot \gamma}{\|\gamma\|}$ is monotonically increasing. This together with (3.42) and $L(0) = 0$ imply that

$$\lim_{\|\gamma\| \rightarrow \infty} \frac{\nabla L(\gamma) \cdot \gamma}{\|\gamma\|} = +\infty.$$

Then the unique solution of (3.38) exists by the solvability theorem for monotone operators, see, for example, [16]. If with $p = 1$, then under condition \mathbf{A}_1^s using the Lebesgue theorem we obtain that $\lim_{\|\gamma\| \rightarrow b-0} \frac{\nabla L(\gamma) \cdot \gamma}{\|\gamma\|} = +\infty$. Then we can repeat the similar arguments to prove the existence of the unique solution of (3.38).

Equality (3.39) follows from the definition (3.7) of the function $I(r)$. Equality (3.40) follows from (3.38). Equality (3.41) is a direct consequence of relation (3.39).

(2) If $x^* \in R^d$ is small enough, then for any distribution $a(x)$ satisfying (3.1) and (3.2), equation (3.38) can be solved using the implicit function theorem. \square

3.3. The regions of standard and moderate deviations: Proof of Theorem 3.1

Under our standing assumptions the local central limit theorem applies to the sum of independent random variables with a common distribution $a(x)$, see for instance [4, Theorem 19.1]. This implies the desired asymptotics (3.3) of $v(x, t)$ in the region $|x| \leq rt^{1/2}$ with an arbitrary $r > 0$.

In this subsection we show that, in the region $x = rt^{\frac{1+\delta}{2}}(1 + o(1))$ with $0 < \delta < 1$ and $r \in \mathbb{R}^d \setminus \{0\}$, the asymptotics (3.4) for $v(x, t)$ holds, as $t \rightarrow \infty$. First we obtain the asymptotics for the k th convolution power $a^{*k}(x)$ for large enough k . Our approach essentially relies on probabilistic arguments.

LEMMA 3.14. *Let conditions (1.8) and (1.9) be satisfied, and assume that (3.1) holds. Then*

$$a^{*k}(x) = e^{-\frac{1}{2} \frac{\sigma^{-1}x \cdot x}{k} (1+o(1))}, \quad \text{as } k \rightarrow \infty, \frac{|x|^2}{k} \rightarrow \infty, \text{ and } \frac{|x|}{k} \rightarrow 0, \tag{3.43}$$

where σ is the covariance matrix of the distribution $a(x)$, and $o(1) \rightarrow 0$ as $\frac{|x|^2}{k} \rightarrow \infty$.

Proof. Let $x^* = \frac{x}{k}$. Then $x^* \rightarrow 0$ as $|x| \rightarrow \infty$, $k \rightarrow \infty$, and using Proposition 3.13 we conclude that the equation $\nabla L(\gamma) = x^*$ has a unique solution $\gamma^* = \gamma^*(x, k)$, where $\gamma^* \rightarrow 0$ as $x^* \rightarrow 0$. Relation (3.41) implies

$$a^{*k}(x) = a^{*k}(kx^*) = a_{\gamma^*}^{*k}(kx^*)e^{-I(x^*)k} = a_{\gamma^*}^{*k}(kx^*)e^{-I(\frac{x}{k})k}. \tag{3.44}$$

It follows from (3.40) and the local limit theorem for the sum of i.i.d. random variables with the common distribution a_{γ^*} that

$$a_{\gamma^*}^{*k}(kx^*) = \frac{\det \sigma_{\gamma^*}^{-1}}{(2\pi k)^{d/2}}(1 + o(1)) = \frac{\det \sigma^{-1}}{(2\pi k)^{d/2}}(1 + o(1)), \tag{3.45}$$

as $x^* \rightarrow 0$ (and $\gamma^* \rightarrow 0$), where σ is the covariance matrix for $a(x)$.

Finally using the asymptotic relation (3.30) for $I(\frac{x}{k})$ we obtain (3.43) from (3.44) and (3.45). □

COROLLARY 3.15. *If $x = rt^{\frac{1+\delta}{2}}(1 + o(1))$ with $0 < \delta < 1$ and $k \sim t$, then $\frac{\sigma^{-1}x \cdot x}{k} \sim (\sigma^{-1}r \cdot r)t^\delta$ and*

$$a^{*k}(x) = e^{-\frac{1}{2}(\sigma^{-1}r \cdot r)\frac{t^{1+\delta}}{k}(1+o(1))}, \quad \text{as } k \rightarrow \infty. \tag{3.46}$$

Let us study now the asymptotic behaviour of the function $\frac{t^k a^{*k}(x)}{k!}$ as $t \rightarrow \infty$ and $x = rt^{\frac{1+\delta}{2}}(1 + o(1))$, $0 < \delta < 1$. By Lemma 3.14 and estimate (3.46) for any constants α_1 and α_2 such that $0 < \alpha_1 < 1 < \alpha_2 < \infty$, and for all k from the interval $\alpha_1 t < k < \alpha_2 t$ the following asymptotics holds:

$$\frac{t^k a^{*k}(x)}{k!} = \exp \left\{ k \ln t - k \ln k + k - \tilde{c} \frac{t^{1+\delta}}{k}(1 + o(1)) \right\} = \exp S(k, t), \quad t \rightarrow \infty,$$

where $\tilde{c} = \frac{1}{2} \sigma^{-1}r \cdot r$ and $S(z, t) = z \ln t - z \ln z + z - \tilde{c} \frac{t^{1+\delta}}{z}(1 + o(1))$. If α_1 is sufficiently small and α_2 is sufficiently large then the $\max_{\alpha_1 t < z < \alpha_2 t} S(z, t)$ is attained in an interior point of the interval $(\alpha_1 t, \alpha_2 t)$, and the corresponding necessary condition reads

$$\ln \frac{t}{z} + \tilde{c} \frac{t^{1+\delta}}{z^2} = 0.$$

Setting $z = \xi t$ we arrive at the following equation for ξ :

$$\xi^2 \ln \xi = \tilde{c} t^{\delta-1}.$$

Since $0 < \delta < 1$, the right-hand side in this equation vanishes as $t \rightarrow \infty$. Therefore, the solution $\hat{\xi}$ of this equation admits the representation $\hat{\xi} = 1 + \tilde{c} t^{\delta-1}(1 + o(1))$. Consequently,

$$\hat{z} = \hat{\xi}t = t + \tilde{c}t^\delta + o(t^\delta), \quad \max_{\alpha_1 t < z < \alpha_2 t} S(z, t) = S(\hat{z}, t) = t - \tilde{c}t^\delta + o(t^\delta),$$

and for any α_1 and α_2 such that $0 < \alpha_1 < 1 < \alpha_2 < \infty$ we have

$$\max_{\alpha_1 t < k < \alpha_2 t} \frac{t^k a^{*k}(x)}{k!} \leq e^{S(\hat{z}, t)} = e^{t - \frac{1}{2}(\sigma^{-1}r \cdot r)t^\delta + o(t^\delta)}, \quad \text{as } t \rightarrow \infty. \tag{3.47}$$

To estimate $v(x, t)$ in the region $x = rt^{\frac{1+\delta}{2}}(1 + o(1))$ we split the sum in (1.11) into three parts:

$$v(x, t) = e^{-t} \sum_{k=1}^{[t/2]} \frac{t^k a^{*k}(x)}{k!} + e^{-t} \sum_{k=[t/2]+1}^{[2t]} \frac{t^k a^{*k}(x)}{k!} + e^{-t} \sum_{k>2t} \frac{t^k a^{*k}(x)}{k!}. \tag{3.48}$$

Considering the inequalities $k! > k^k e^{-k}$ and $a^{*k}(x) \leq C_1$, one can estimate the first sum in (3.48) as follows:

$$e^{-t} \sum_{k=1}^{[t/2]} \frac{t^k a^{*k}(x)}{k!} \leq C_1 \frac{t}{2} e^{-t} (2e)^{t/2} \leq C_2 e^{-\beta t}, \quad t \rightarrow \infty \tag{3.49}$$

with an arbitrary $\beta \in (0, \frac{1-\ln 2}{2})$. For the third sum using the relation $\frac{t}{k} < \frac{1}{2}$ we get

$$e^{-t} \sum_{k>2t} \frac{t^k a^{*k}(x)}{k!} \leq C_1 e^{t-2t \ln 2}. \tag{3.50}$$

To estimate the second sum in (3.48) we use (3.47). This yields

$$e^{-t} \sum_{k=[t/2]+1}^{[2t]} \frac{t^k a^{*k}(x)}{k!} \leq 2te^{-t} e^{S(\hat{z},t)} \leq e^{-\frac{1}{2}(\sigma^{-1}r \cdot r) t^\delta + o(t^\delta)}, \quad t \rightarrow \infty. \tag{3.51}$$

Finally, from (3.48) and (3.51) we get the asymptotical upper bound in the region $x = rt^{\frac{1+\delta}{2}}(1 + o(1))$.

Taking in the sum (3.48) just one term that corresponds to $\max S(z, t)$ we obtain the lower bound

$$v(x, t) \geq e^{-t} e^{S(\hat{z},t)} \geq e^{-\frac{1}{2}(\sigma^{-1}r \cdot r) t^\delta (1+o(1))}.$$

This completes the proof of (3.4).

3.4. The region of extra-large deviations: Proof of Theorem 3.2

This section deals with the large-time behaviour of $v(x, t)$ in the region $x \gg t$ which is associated with the ‘extra-large’ deviations of the corresponding process. In this region we use the Markov inequality for estimating $\Pr(|S_k| > |x|)$.

LEMMA 3.16. *Let X_i be i.i.d. one-dimensional random variables with a common distribution $a(x)$, satisfying (1.8)–(1.9) and (3.1). Then there exist constants $\alpha_p = \alpha_p(b, p)$ and $\varkappa_p = \varkappa_p(b, p)$, such that for all $1 \leq k \leq \alpha_p x$ the following estimate holds:*

$$P\{S_k > x\} \leq e^{-\varkappa_p (\frac{x}{k})^p k}. \tag{3.52}$$

Proof. The cases $p = 1$ and $p > 1$ are considered in a slightly different way. If $p = 1$, the inequality

$$\mathbb{E}e^{mX_1} \leq e^{hm^2}$$

being valid for all $m \in (0, \frac{b}{2})$ with some constant $h > 0$. Then the Markov inequality yields

$$\begin{aligned} P\{S_k > x\} &\leq \min_{0 < m \leq \frac{b}{2}} \frac{(\mathbb{E}e^{mX_1})^k}{e^{mx}} = \exp \left\{ \min_{0 < m \leq \frac{b}{2}} (hm^2 k - mx) \right\} \\ &= e^{\frac{hb^2 k}{4} - \frac{bx}{2}} = e^{(\frac{hb^2 k}{4} - \frac{bx}{4}) - \frac{bx}{4}} \leq e^{-\frac{bx}{4}} \end{aligned} \tag{3.53}$$

for $k < \frac{x}{hb}$. Thus in the case $p = 1$ inequality (3.52) holds with $\varkappa_1 = \frac{b}{4}$ and $\alpha_1 = \frac{1}{2hb}$.

If $p > 1$ then applying the Markov inequality we get

$$P\{S_k > x\} \leq \min_{m>0} \frac{(\mathbb{E}e^{mX_1})^k}{e^{mx}} = \left(\min_{m>0} \frac{\mathbb{E}e^{mX_1}}{e^{\frac{mx}{k}}} \right)^k. \tag{3.54}$$

Let us estimate $\mathbb{E}e^{mX_1}$. Setting $\varphi(x) = mx - bx^p$, $x > 0$, we obtain

$$\max_x \varphi(x) = \varphi \left(\left(\frac{m}{bp} \right)^{\frac{1}{p-1}} \right) = c_2(b, p) m^{\frac{p}{p-1}}, \quad c_2(b, p) = \frac{p-1}{p} (bp)^{-\frac{1}{p-1}}, \tag{3.55}$$

and $\varphi(x) < 0$ as $x > x_1 = (\frac{m}{b})^{\frac{1}{p-1}}$. Since $mx - bx^p \leq \varphi'(x_1)(x - x_1) = m(1 - p)(x - x_1)$ for $x \geq x_1$, from (3.55) it follows that

$$\begin{aligned} \mathbb{E}e^{mX_1} &\leq \int_{|x| \leq x_1} e^{\varphi(x)} dx + \int_{|x| > x_1} e^{mx} a(x) dx \\ &\leq c_3(b, p) m^{\frac{1}{p-1}} e^{c_2(b, p) m^{\frac{p}{p-1}}} + C_1((p - 1)m)^{-1}. \end{aligned}$$

Then there exists a constant $c_4 = c_4(p, b)$ such that for all $m \geq 1$

$$\mathbb{E}e^{mX_1} \leq e^{c_4(b, p) m^{\frac{p}{p-1}}}. \tag{3.56}$$

Inserting (3.56) into (3.54) yields

$$P\{S_k > x\} \leq \left(\min_{m \geq 1} e^{c_4(b, p) m^{\frac{p}{p-1}} - m \frac{x}{k}} \right)^k = e^{k \min_{m \geq 1} f_{p, k}(m)}, \tag{3.57}$$

where

$$f_{p, k}(m) = c_4(b, p) m^{\frac{p}{p-1}} - m \frac{x}{k}, \quad m \geq 1.$$

If we take $k \leq \alpha_p x$ with $\alpha_p \leq \frac{p-1}{2pc_4}$, then $f_{p, k}(1) < 0$ and $f'_{p, k}(1) < 0$. Determining the minimum of $f_{p, k}$ we obtain

$$\min_{m \geq 1} f_{p, k}(m) = -\varkappa_p \left(\frac{x}{k} \right)^p$$

with some constant $\varkappa_p > 0$. Inequality (3.52) then follows from (3.57). □

COROLLARY 3.17. *In the multidimensional case estimate (3.52) takes the form*

$$P\{|S_k| > |x|\} \leq e^{-\varkappa_p \left(\frac{|x|}{k}\right)^p k(1+o(1))}, \quad \text{as } |x| \rightarrow \infty. \tag{3.58}$$

Proof. Given a sequence of i.i.d. random vectors $X_j, j = 1, 2, \dots$, with a common distribution density $a(\cdot)$, for any $\theta \in S^{d-1}$ we consider one-dimensional random variables $\theta \cdot X_j$. Denote the distribution density of $\theta \cdot X_j$ by $a_\theta(s)$. Then

$$a_\theta(s) \leq C(1 + s)^{d-1} e^{-b|s|^p}.$$

Therefore, by (3.52)

$$P\{S_j \cdot \theta > |x|\} \leq e^{-\varkappa_p \left(\frac{|x|}{k}\right)^p k(1+o(1))}. \tag{3.59}$$

For a d -dimensional random vector X and arbitrary $\varepsilon > 0$ one can find a finite collection of unit vectors $\theta_1, \dots, \theta_N, N = N(\varepsilon, d)$ such that

$$\{|X| > |x|\} \subset \bigcup_{i=1}^N \{\theta_i \cdot X > (1 - \varepsilon)|x|\}.$$

Then

$$P\{|X| > |x|\} \leq N(\varepsilon, d) P\{\theta \cdot X > (1 - \varepsilon)|x|\},$$

and together with (3.59) it gives the desired asymptotic estimate (3.58) for $P(|S_k| > |x|)$ in the multi-dimensional case. □

We proceed with obtaining pointwise estimates for $a^{*k}(x)$. Denote by $F_k(s)$ the distribution function of $|S_k|$, then in the case $p = 1$ we have

$$\begin{aligned} a^{*(k+1)}(x) &\leq C_1 \int_0^\infty e^{-b(|x|-s)} dF_k(s) \\ &= C_1 \int_0^{\frac{1}{2}|x|} e^{-b(|x|-s)} dF_k(s) + C_1 \int_{\frac{1}{2}|x|}^\infty e^{-b(|x|-s)} dF_k(s) \leq C_1 e^{-\frac{1}{2}b|x|} + C_1 P\{|S_k| \geq \frac{1}{2}|x|\}. \end{aligned}$$

Together with estimates (3.58) and (3.53), where $\varkappa_1 = \frac{b}{4}$, this yields for all $k \leq \alpha_1|x|$:

$$a^{*k}(x) \leq 2C_1 e^{-\frac{b}{8}|x|}. \tag{3.60}$$

The case $p > 1$ can be treated similarly, and for any $k \leq \alpha_p|x|$ we obtain

$$\begin{aligned} a^{*(k+1)}(x) &\leq C_1 \int_0^\infty e^{-b(|x|-s)^p} dF_k(s) \\ &\leq C_1 e^{-\frac{b}{2p}x^p} + C_1 e^{-\varkappa_p \frac{x^p}{2pk^{p-1}}} \leq C_2 e^{-\tilde{\varkappa}_p \frac{x^p}{k^{p-1}}}. \end{aligned} \tag{3.61}$$

In order to obtain upper bounds for the terms of the sum in (1.11) we use estimates (3.60)–(3.61). We denote

$$\begin{aligned} S(z, t) &= z \ln t - z \ln z + z + \ln a^{*z}(x), \\ S_0(z, t) &= z \ln t - z \ln z + z = z \ln \frac{t}{z} + z. \end{aligned} \tag{3.62}$$

Note that

$$\max_z S_0(z, t) = S_0(t, t) = t,$$

and $S_0(z, t)$ is decreasing in z as $z > t$. Consequently, for any $c > 0$ and for sufficiently large t we have

$$\max_{z \geq ct^{\frac{\delta+1}{2}}} S_0(z, t) = S_0\left(ct^{\frac{\delta+1}{2}}, t\right) < -\tilde{c}t^{\frac{\delta+1}{2}} \ln t. \tag{3.63}$$

In the case $p = 1$, considering the upper bound $a^{*k}(x) \leq C_1$, we get

$$\max_{z \geq ct^{\frac{\delta+1}{2}}} S(z, t) \leq \max_{z \geq ct^{\frac{\delta+1}{2}}} S_0(z, t) + \ln C_1 < -\tilde{c}_1 t^{\frac{\delta+1}{2}} \ln t. \tag{3.64}$$

If $k < \alpha_1|x| = \alpha_1 r t^{\frac{\delta+1}{2}}(1 + o(1))$, then estimate (3.60) implies the following uniform in k upper bound

$$a^{*k}(x) < C_2 e^{-\frac{b}{8}|x|} = C_2 e^{-\frac{b}{8}rt^{\frac{\delta+1}{2}}(1+o(1))}. \tag{3.65}$$

Consequently,

$$\max_{k < \alpha_1 r t^{\frac{\delta+1}{2}}} S(k, t) \leq S_0(t, t) + \max_{k < \alpha_1 r t^{\frac{\delta+1}{2}}} \ln a^{*k}(x) \leq t - \frac{b}{8}rt^{\frac{\delta+1}{2}}. \tag{3.66}$$

Finally, using (3.64) and (3.66), we conclude that in the case $p = 1$ the asymptotic estimate (3.5) holds with $c_1 = \frac{b}{8}r$. Indeed,

$$v(x, t) = e^{-t} \sum_{k=1}^\infty \frac{t^k a^{*k}(x)}{k!} = e^{-t} \sum_{k=1}^{\alpha_1 r t^{\frac{\delta+1}{2}}} \frac{t^k a^{*k}(x)}{k!} + e^{-t} \sum_{k > \alpha_1 r t^{\frac{\delta+1}{2}}} \frac{t^k a^{*k}(x)}{k!}$$

$$\leq \alpha_1 r t^{\frac{\delta+1}{2}} e^{-\frac{b}{8} r t^{\frac{\delta+1}{2}}} + O\left(e^{-t^{\frac{\delta+1}{2}} \ln t}\right) \leq e^{-\frac{b}{8} r t^{\frac{\delta+1}{2}}} (1+o(1)) \quad \text{as } t \rightarrow \infty.$$

In the case $p > 1$, using estimates (3.63) and (3.61), for all $k \leq \alpha_p |x| = \alpha_p r t^{\frac{\delta+1}{2}} (1 + o(1))$ we have

$$\begin{aligned} \max_{k \leq \alpha_p |x|} S(k, t) &= \max_{k \leq \alpha_p |x|} \{S_0(k, t) + \ln a^{*k}(x)\} \\ &\leq \max_{k \leq \alpha_p |x|} \left\{ S_0(k, t) - \tilde{\alpha}_p r^p \frac{t^{\frac{\delta+1}{2} p}}{k^{p-1}} \right\} \leq -c_p^{(1)} t^{\frac{\delta+1}{2}} (\ln t)^{\frac{p-1}{p}}. \end{aligned}$$

In order to justify the last inequality we note that

$$\hat{k}(t) := \operatorname{argmax} \left\{ S_0(k, t) - \tilde{\alpha}_p r^p \frac{t^{\frac{\delta+1}{2} p}}{k^{p-1}} \right\} = \hat{c} \frac{t^{\frac{\delta+1}{2}}}{(\ln t)^{\frac{1}{p}}} = o(t^{\frac{\delta+1}{2}})$$

with a constant $\hat{c} = \hat{c}(p, \delta, r)$. Then $\hat{k}(t) < \alpha_p |x|$ and, therefore,

$$\frac{t^k a^{*k}(x)}{k!} \leq e^{-c_p^{(1)} t^{\frac{\delta+1}{2}} (\ln t)^{\frac{p-1}{p}}}, \quad k < \hat{k}(t).$$

Estimating $S(k, t)$ for $k > \hat{k}(t)$ relies on the inequalities for the function $S_0(k, t)$ similar to those in (3.63) and the upper bound $a^{*k}(x) \leq C_1$. We have for $k > \hat{k}(t)$

$$\frac{t^k a^{*k}(x)}{k!} < C_1 e^{S_0(\hat{k}(t), t)} = C_1 e^{\hat{k}(t)(\ln t - \ln \hat{k}(t) + 1)} \leq e^{-c_p^{(2)} t^{\frac{\delta+1}{2}} (\ln t)^{\frac{p-1}{p}} (1+o(1))}.$$

Finally, taking into account the fact that $\frac{t}{k} < \frac{1}{2}$ for all $k > \hat{k}(t) = \hat{c} t^{\frac{\delta+1}{2}} (\ln t)^{-\frac{1}{p}}$, we conclude that in the case $p > 1$

$$\begin{aligned} v(x, t) &= e^{-t} \sum_{k \leq \hat{k}(t)} \frac{t^k a^{*k}(x)}{k!} + e^{-t} \sum_{k > \hat{k}(t)} \frac{t^k a^{*k}(x)}{k!} \leq C_1 t^{\frac{\delta+1}{2}} e^{-t - c_p^{(1)} t^{\frac{\delta+1}{2}} (\ln t)^{\frac{p-1}{p}}} \\ &\quad + e^{-t - c_p^{(2)} t^{\frac{\delta+1}{2}} (\ln t)^{\frac{p-1}{p}} (1+o(1))} \leq e^{-c_p t^{\frac{\delta+1}{2}} (\ln t)^{\frac{p-1}{p}} (1+o(1))}. \end{aligned}$$

Now let us consider the case of $a(x)$ with a compact support. If (3.2) holds then

$$v(x, t) = e^{-t} \sum_{k > \frac{r}{\mu} t^{\frac{\delta+1}{2}}} \frac{t^k a^{*k}(x)}{k!} \leq C_1 e^{-t} \exp \left\{ \max_{k > \frac{r}{\mu} t^{\frac{\delta+1}{2}}} S_0(k, t) \right\}, \tag{3.67}$$

where S_0 has been defined in (3.62). The function $S_0(k, t)$ is decreasing in variable k as $k > t$, hence

$$\max_{k > \frac{r}{\mu} t^{\frac{\delta+1}{2}}} S_0(k, t) = S_0\left(\frac{r}{\mu} t^{\frac{\delta+1}{2}}, t\right) = -\frac{r}{\mu} \frac{\delta - 1}{2} t^{\frac{\delta+1}{2}} \ln t (1 + o(1)). \tag{3.68}$$

Finally, (3.67) and (3.68) imply estimate (3.6). This completes the proof of Theorem 3.2.

3.5. The region of large deviations: Proof of Theorem 3.4

The main step of the proof is obtaining pointwise estimates for a^{*k} .

LEMMA 3.18. Let $a(x)$ satisfy (1.8)–(1.9) and (3.1), and assume that, in the case $p = 1$, condition \mathbf{A}_1 holds. Then for $x = rt(1 + o(1))$, $r \in R^d \setminus \{0\}$, and for any positive constants $\alpha_1 < \alpha_2$ we have

$$a^{*k}(x) \leq e^{-I(\frac{x}{k})k(1+o(1))}, \quad \text{if } \alpha_1 t \leq k \leq \alpha_2 t. \tag{3.69}$$

Furthermore, there exists a positive constant $\alpha_1 > 0$ such that

$$a^{*k}(x) \leq e^{-I(r)t(1+o(1))}; \quad \text{if } 1 \leq k \leq \alpha_1 t. \tag{3.70}$$

If condition (3.2) is fulfilled, then for any α_2

$$a^{*k}(x) = 0, \text{ if } k \leq \frac{|x|}{\mu}, \text{ and } a^{*k}(x) \leq e^{-I(\frac{x}{k})k(1+o(1))}, \text{ if } \frac{|x|}{\mu} \leq k \leq \alpha_2 t. \tag{3.71}$$

Proof. We start with the case $p = 1$, $\alpha_1 t \leq k \leq \alpha_2 t$. The kernel $a^{*(k+1)}(x)$ can be written as follows:

$$a^{*(k+1)}(x) = \int_{\{z: I(\frac{z}{k}) \leq I(\frac{x}{k})\}} a^{*k}(z)a(x-z)dz + \int_{\{z: I(\frac{z}{k}) > I(\frac{x}{k})\}} a^{*k}(z)a(x-z)dz. \tag{3.72}$$

Denote by $A_1 = \{z : I(\frac{z}{k}) \leq I(\frac{x}{k})\}$, $A_2 = \{z : I(\frac{z}{k}) > I(\frac{x}{k})\}$. Using the large deviations principle for the sum of i.i.d. random vectors, see [9], we obtain the upper estimate for the second integral in (3.72), when $\alpha_1 t \leq k \leq \alpha_2 t$:

$$\int_{A_2} a^{*k}(z)a(x-z)dz \leq C_1 P(S_k \in A_2) \leq C_1 \exp \left\{ - \inf_{\rho \in A_2} I(\rho) k + o(k) \right\} = e^{-I(\frac{x}{k})k + o(k)}. \tag{3.73}$$

To estimate the first integral in (3.72) we define $F_k(s) = \int_{\{z: I(\frac{z}{k})k \geq s\}} a^{*k}(z)dz$ with $s \in (0, I(\frac{x}{k})k)$,

then we have

$$F_k(0) = 1, \quad F_k(\infty) = 0, \quad F_k(s) = e^{-s(1+o(1))} \quad (k \rightarrow \infty). \tag{3.74}$$

If we denote $L_s = \{z : I(\frac{z}{k})k = s\}$, then $\text{dist}(x, L_s)$ is a decreasing continuous function on $[0, I(\frac{x}{k})k]$, it is smooth on $(0, I(\frac{x}{k})k]$. For each $s \in [0, I(\frac{x}{k})k]$ there exists a unique $z_s \in L_s$ such that $\text{dist}(x, L_s) = |x - z_s|$. All these assertions are elementary consequences of convexity of the function I . Clearly, $z_0 = 0$, $z_{I(\frac{x}{k})k} = x$.

It follows from Proposition 3.10 and estimate (3.1) that for any $x, z \in R^d$, such that $I(\frac{z}{k}) < I(\frac{x}{k})$, the following inequality holds true:

$$\left(I\left(\frac{x}{k}\right) - I\left(\frac{z}{k}\right) \right) k \leq \max_{r \in I(\frac{x}{k}, \frac{z}{k})} |\nabla I(r)| |x - z| \leq b |x - z| \leq -\ln a(x - z),$$

where by $l(x, y)$ we denote the segment connecting points x and y . Consequently,

$$e^{-I(\frac{z}{k})k} a(x - z) \leq e^{-I(\frac{x}{k})k}. \tag{3.75}$$

Then using (3.74) and inequality (3.75) we rewrite the first integral in (3.72) as follows:

$$\begin{aligned} \int_{A_1} a^{*k}(z)a(x-z)dz &\leq C_1 \int_{A_1} a^{*k}(z)e^{-b|x-z|}dz \leq C_1 \int_0^{I(\frac{x}{k})k} e^{-b \text{dist}(x, L_s)} d(-F_k(s)) \\ &= C_1 e^{-b \text{dist}(x, L_s)} F_k(s) \Big|_{I(\frac{x}{k})k}^0 - C_1 b \int_0^{I(\frac{x}{k})k} F_k(s) e^{-b \text{dist}(x, L_s)} \frac{d}{ds} \text{dist}(x, L_s) ds \end{aligned}$$

$$\begin{aligned}
 &= C_1 e^{-b|x|} - C_1 e^{-I(\frac{x}{k})k(1+o(1))} - C_1 b \int_0^{I(\frac{x}{k})k} e^{-s(1+o(1))} e^{-b|x-z_s|} \frac{d}{ds} \text{dist}(x, L_s) ds \\
 &\leq e^{-I(\frac{x}{k})k(1+o(1))} - C_1 b \int_0^{I(\frac{x}{k})k} e^{-I(\frac{zs}{k})k(1+o(1))} e^{-b|x-z_s|} \frac{d}{ds} \text{dist}(x, L_s) ds \\
 &\leq e^{-I(\frac{x}{k})k(1+o(1))} - C_1 b e^{-I(\frac{x}{k})k(1+o(1))} \int_0^{I(\frac{x}{k})k} \frac{d}{ds} \text{dist}(x, L_s) ds \\
 &\leq e^{-I(\frac{x}{k})k(1+o(1))} + C_1 b e^{-I(\frac{x}{k})k(1+o(1))} |x| \leq e^{-I(\frac{x}{k})k(1+o(1))}.
 \end{aligned}$$

This inequality together with (3.73) imply (3.69) in the case $p = 1$.

To prove the upper bound (3.69) for $p > 1$ and $\alpha_1 t \leq k \leq \alpha_2 t$ we rewrite $a^{*(k+1)}(x)$ as a sum

$$a^{*(k+1)}(x) = \int_{|z-x| < hk^{1/p}} a^{*k}(z)a(x-z)dz + \int_{|z-x| \geq hk^{1/p}} a^{*k}(z)a(x-z)dz. \tag{3.76}$$

The second integral in (3.76) has an upper bound

$$\max_{|u| \geq hk^{1/p}} a(u) \leq C_1 e^{-bh^p k}.$$

If the constant $h > 0$ is taken in such a way that $bh^p > I(\frac{r}{\alpha_1})$, then $bh^p > I(\frac{r}{\alpha_1}) > I(\frac{x}{k})$ for any $k \in [\alpha_1 t, \alpha_2 t]$. Thus, the second term in (3.76) is bounded by (3.69).

For $k \sim t$ and for arbitrary $\varkappa > 0$ the first term in (3.76) can be estimated from above as

$$\begin{aligned}
 \int_{|z-x| < hk^{1/p}} a^{*k}(z)a(x-z)dz &\leq C_1 \int_{|z-x| < hk^{1/p}} a^{*k}(z)dz \leq C_1 \Pr\{|S_k - x| < \varkappa k\} \\
 &\leq C_1 \exp \left\{ - \inf_{\rho \in A_\varkappa} I(\rho)k + o(k) \right\} \leq e^{-I(\frac{x}{k})k+o(k)},
 \end{aligned}$$

where $A_\varkappa = \{z : |z - \frac{x}{k}| < \varkappa\}$. Here we used the large deviations principle for estimating $\Pr\{|S_k - x| < \varkappa k\}$ and continuity of $I(r)$.

In the case $p = 1$ and $k \leq \alpha_1 t$ we apply the upper bound

$$a(x) \leq A \tilde{a}(|x|), \quad \text{with some constant } A > 1, \tag{3.77}$$

where $\tilde{a}(|x|) = a_1 e^{-b|x|}$ is a spherically symmetric kernel satisfying (3.1) with the same b . Next we need the following statement for one-dimensional random variables.

PROPOSITION 3.19. *Let $a(x), x \in \mathbb{R}$, satisfy (3.1) with $p = 1$ and condition \mathbf{A}_1 holds. Then there exists positive constant \tilde{C}_1 such that*

$$a^{*k}(x) \leq \tilde{C}_1 e^{-I(\frac{x}{k})k(1+o(1))}; \quad \text{for all } k \geq 1. \tag{3.78}$$

Proof. We represent $a^{*(k+1)}(x), x \in \mathbb{R}$ as follows:

$$a^{*(k+1)}(x) = \int_{-\infty}^0 a^{*k}(z)a(x-z)dz + \int_0^x a^{*k}(z)a(x-z)dz + \int_x^\infty a^{*k}(z)a(x-z)dz. \tag{3.79}$$

Since $I(\frac{x}{k})k < bx$, the first integral in (3.79) admits the estimate

$$\int_{-\infty}^0 a^{*k}(z)a(x-z)dz \leq C_1 e^{-bx} \leq e^{-I(\frac{x}{k})k}.$$

For the last integral in (3.79) we apply the Markov inequality:

$$\int_x^\infty a^{*k}(z)dz = P(S_k > x) \leq \inf_\gamma \frac{(\mathbb{E}e^{\gamma X_1})^k}{e^{k\gamma\frac{x}{k}}} = \inf_\gamma e^{kL(\gamma)-k\gamma\frac{x}{k}} = e^{-I(\frac{x}{k})k}. \tag{3.80}$$

Then we get for any k and any $x > 0$

$$\int_x^\infty a^{*k}(z)a(x-z)dz \leq C_1 P(S_k > x) \leq C_1 e^{-I(\frac{x}{k})k}.$$

Denote $\tilde{F}_k(x) = \int_x^\infty a^{*k}(z)dz = P(S_k > x)$. Then the second integral in (3.79) admits the estimate

$$\begin{aligned} \int_0^x a^{*k}(z)a(x-z)dz &\leq C_1 \int_0^x e^{-b(x-z)}d(-\tilde{F}_k(z)) = C_1 e^{-b(x-z)}\tilde{F}_k(z)\Big|_x^0 + C_1 b \int_0^x e^{-b(x-z)}\tilde{F}_k(z)dz \\ &\leq C_1 e^{-bx} + C_1 b \int_0^x e^{-I(\frac{x}{k})k-b(x-z)}dz \leq C_1 e^{-bx} + C_1 bx e^{-I(\frac{x}{k})k} \leq C_2 x e^{-I(\frac{x}{k})k}; \end{aligned}$$

we have used here the inequalities

$$I\left(\frac{x}{k}\right) - I\left(\frac{z}{k}\right) < b \frac{x-z}{k} \quad \text{for all } z \in (0, x), \quad \text{and } I\left(\frac{x}{k}\right)k < bx.$$

Considering $x = rt(1 + o(1))$ we obtain estimate (3.78) for all $k \geq 1$. □

Then using (3.78) we have

$$a^{*k}(x) \leq A^k \tilde{a}^{*k}(|x|) \leq A^k e^{-I_a(\frac{|x|}{k})k(1+o(1))}.$$

Since $k < \alpha_1 t$ with a small α_1 , then $\frac{|x|}{k} > \frac{|r|}{\alpha_1} \gg 1$, and using asymptotic representation (3.26) for $I_a(s)$ as $s \rightarrow \infty$ and inequality (3.33), we conclude that for any $\delta > 0$ there exists $\alpha_1 > 0$ such that

$$a^{*k}(x) \leq A^k \tilde{a}^{*k}(|x|) \leq A^k e^{-b|x|(1-\delta)} = A^k e^{-b|r|t(1-\delta)} \leq e^{-b|r|t(1-\delta)+\alpha_1 t \ln A} \leq e^{-I(r)t}. \tag{3.81}$$

In order to obtain the last inequality we chose $\delta = \frac{(b|r|-I(r))}{4b|r|}$. Thus (3.70) is proved for $p = 1$.

If $p > 1$ and $k \leq \alpha_1 t$, then for sufficiently small α_1 recalling that $x = rt(1 + o(1))$, from the Markov inequality (3.80) we have

$$P\{|S_k| > \frac{1}{2}|x|\} \leq e^{-\tilde{I}(\frac{|x|}{2k})k} \leq e^{-\tilde{I}(\frac{|r|}{2\alpha_1})\alpha_1 t} \leq e^{-2I(r)t}, \tag{3.82}$$

where $\tilde{I}(s)$ is the rate function for the one-dimensional random variable $|X|$. Here we used the fact that the function $J(\alpha) = \alpha\tilde{I}(\frac{s}{\alpha})$ is decreasing in $\alpha \in (0, 1]$, that is a consequence of convexity of $\tilde{I}(s)$. Moreover, by (3.27) we have $J(\alpha) \rightarrow \infty$ as $\alpha \rightarrow 0+$. Then using (3.82) we conclude that for a small enough constant $\alpha_1 > 0$ we get

$$\begin{aligned} a^{*(k+1)}(x) &= \int_{|z| \leq \frac{1}{2}|x|} a^{*k}(z) a(x-z) dz + \int_{|z| > \frac{1}{2}|x|} a^{*k}(z) a(x-z) dz \\ &\leq C_1 e^{-b(\frac{|x|}{2})^p} + C_1 e^{-\tilde{I}(\frac{|r|}{2\alpha_1})\alpha_1 t} \leq C_1 e^{-\tilde{b}(\frac{|r|}{2})^p t^p} + C_1 e^{-\tilde{I}(\frac{|r|}{2\alpha_1})\alpha_1 t} \leq \tilde{C}_2 e^{-I(r)t}. \end{aligned}$$

This completes the proof of estimates (3.69) and (3.70).

The first relation in (3.71) is evident. The proof of the second one is based on the same arguments as those used in the case $p > 1$. □

Combining Stirling's formula with the estimates of Lemma 3.18 we obtain the following statement.

COROLLARY 3.20. *Let the assumptions of Lemma 3.18 be fulfilled. If $x = rt(1 + o(1))$, then for all k such that $\alpha_1 t \leq k \leq \alpha_2 t$ with arbitrary positive numbers α_2 and α_1 , estimate (3.69) implies that*

$$\frac{t^k a^{*k}(x)}{k!} \leq \exp\left\{k \ln t - k \ln k + k - I\left(\frac{x}{k}\right)k + o(t)\right\} = \exp\{S(k, t) + o(t)\}, \tag{3.83}$$

as $t \rightarrow \infty$, where $S(k, t) = k \ln t - k \ln k + k - I\left(\frac{rt}{k}\right)k$.

Recalling the definition of ξ_r in (3.8) and the function Φ in (3.9) we have

$$e^{-t+S(\hat{z}, t)} = \exp\left\{t\left(-1 + \frac{1}{\xi_r}(1 + \ln \xi_r - I(\xi_r r))\right)\right\} = \exp\{-\Phi(r)t\}, \tag{3.84}$$

where $\hat{z} = \operatorname{argmax} S(z, t)$. If $x = rt(1 + o(1))$ as $t \rightarrow \infty$, then the following upper bound

$$e^{-t} \frac{t^k a^{*k}(x)}{k!} \leq e^{-\Phi(r)t(1+o(1))} \tag{3.85}$$

is valid for all k from the interval $k \in (\alpha_1 t, \alpha_2 t)$.

To estimate $v(x, t)$ from above we decompose the sum in (1.11) into three parts:

$$v(x, t) = e^{-t} \sum_{k < \alpha_1 t} \frac{t^k a^{*k}(x)}{k!} + e^{-t} \sum_{k = \alpha_1 t}^{\alpha_2 t} \frac{t^k a^{*k}(x)}{k!} + e^{-t} \sum_{k > \alpha_2 t} \frac{t^k a^{*k}(x)}{k!}. \tag{3.86}$$

For the first sum in (3.86) we apply upper bound (3.70). This together with (3.84) yield

$$a^{*k}(x) \leq e^{-I(r)t+o(t)} = e^{-t+S(t,t)+o(t)} \leq e^{-t+S(\hat{z},t)+o(t)} = e^{-\Phi(r)t(1+o(1))} \tag{3.87}$$

because $-I(r)t = S(t, t) - t$, and $\max_z S(z, t) = S(\hat{z}, t)$ with $\hat{z} > t$. Consequently,

$$e^{-t} \sum_{k=1}^{\alpha_1 t} \frac{t^k a^{*k}(x)}{k!} \leq e^{-\Phi(r)t(1+o(1))}, \quad \text{as } t \rightarrow \infty. \tag{3.88}$$

For the third sum, if $k > \alpha_2 t$ with $\alpha_2 > 2$ then we have

$$\frac{t^k}{k!} < \frac{t^{\alpha_2 t}}{(\alpha_2 t)!} < e^{(\alpha_2 - \alpha_2 \ln \alpha_2)t}.$$

Choosing $\alpha_2 > 2$ such that $1 - \alpha_2 + \alpha_2 \ln \alpha_2 > \Phi(r)$, we obtain

$$e^{-t} \sum_{k > 2t} \frac{t^k a^{*k}(x)}{k!} \leq C_1 e^{(-1 + \alpha_2 - \alpha_2 \ln \alpha_2)t} < e^{-\Phi(r)t}. \tag{3.89}$$

It remains to estimate the second sum on the right-hand side of (3.86). To this end we use (3.85), then

$$e^{-t} \sum_{k = \alpha_1 t}^{\alpha_2 t} \frac{t^k a^{*k}(x)}{k!} \leq \alpha_2 t e^{-\Phi(r)t(1+o(1))} = e^{-\Phi(r)t(1+o(1))}, \quad t \rightarrow \infty. \tag{3.90}$$

Finally, in the region $x = rt(1 + o(1))$, $r \neq 0$, from (3.88)–(3.90) we deduce:

$$v(x, t) \leq e^{-\Phi(r)t(1+o(1))}, \quad t \rightarrow \infty. \tag{3.91}$$

For $a(x)$ with a finite support we take $\alpha_1 = r/\mu$ in (3.86). Then the first sum on the right-hand side of (3.86) does not contribute. Estimating the two other sums relies on (3.85), (3.89) and (3.90) like in the case $p > 1$. This completes the proof of (3.12).

It remains to show that the function $\Phi(r)$ satisfies the asymptotic relations in (3.13)–(3.17). Considering the properties of the function $I(r)$, in particular (3.30), it is easy to see that $\xi_r = 1 - \frac{r^2}{2\sigma} + o(r^2)$, as $r \rightarrow 0$. Recalling now the definition of $\Phi(r)$ in (3.9), we finally obtain asymptotic formula (3.13).

The asymptotics of $\Phi(r)$ for large r depends crucially on the rate of decay of $a(x)$ at infinity. We start with the case, when $a(x)$ satisfies (3.2). Then from Proposition 3.10 it follows that $I(\xi r) = \infty$ for all $\xi|r| > \mu$. Then the solution ξ_r of equation (3.8) satisfies the inequality $\xi_r < \frac{\mu}{|r|}$. By the definition of $\Phi(r)$ we have

$$\Phi(r) = 1 - \frac{1}{\xi_r} + \frac{1}{\xi_r} \ln \frac{1}{\xi_r} + \frac{1}{\xi_r} I(r\xi_r) \geq \frac{1}{\xi_r} \left(\ln \frac{1}{\xi_r} - 1 \right),$$

Therefore, for large enough r ,

$$\Phi(r) \geq \min_{x \in (\frac{|r|}{\mu}, \infty)} x(\ln x - 1) = \frac{|r|}{\mu} \left(\ln \frac{|r|}{\mu} - 1 \right),$$

and we obtain (3.17).

Since the principal term on the right-hand side of (3.32) only depends on $|r|$ as $r \rightarrow \infty$, then in the case $p > 1$ for the solution ξ_r of equation (3.8) we have $\xi_r = \xi_{|r|}(1 + o(1))$, as $r \rightarrow \infty$. Therefore, we can reduce the general case to the spherically symmetric case (or the one-dimensional case). Note that for any $p \geq 1$ condition \mathbf{A}_p implies (3.37). The next statement describes the asymptotic behaviour of ξ_r for large r under the assumption that (3.1) and (3.37) hold true.

PROPOSITION 3.21. *Let (3.1) and (3.37) hold. Then*

$$u(s) := s \xi_s \rightarrow \infty, \quad \text{as } s \rightarrow \infty. \tag{3.92}$$

If $p > 1$ and condition \mathbf{A}_p is fulfilled, then

$$\xi_s = h_p \frac{(\ln s)^{1/p}}{s} (1 + o(1)), \quad \text{as } s \rightarrow \infty, \tag{3.93}$$

where h_p is a constant depending on p and b .

Proof. We first prove (3.92). If we assume that $u(s)$ is bounded: $u = u(s) < a$, then for all $s > 0$ the function $\ln \xi_s$ is bounded from below:

$$\ln \xi_s = I(u) - uI'(u) > I(a) - aI'(a) > -\infty. \tag{3.94}$$

We have used here the facts that $J(u) = I(u) - uI'(u)$ is a decreasing function on $[0, +\infty)$, and due to condition (3.37) the functions $I(u), I'(u)$ are finite for all $u > 0$. On the other hand,

$$\ln \xi_s = \ln \frac{u(s)}{s} < \ln a - \ln s.$$

For large s this inequality contradicts (3.94). This proves (3.92).

The function $J(u) = I(u) - uI'(u) < 0$ is negative for all $u > 0$, because $J(0) = 0$, $J'(u) \leq 0$ for $u \geq 0$, and $J'(u) < 0$ for $0 \leq u < \kappa_0$ with some $\kappa_0 > 0$. In the case $p > 1$ combining this inequality with (3.8), (3.26) and (3.92) we conclude that

$$\ln \frac{1}{\xi_s} = b(p-1)(r\xi_s)^p(1+o(1)), \quad s \rightarrow \infty.$$

Consequently, we get (3.93) in the case $p > 1$: $\xi_s = h_p \frac{(\ln s)^{1/p}}{s}(1+o(1))$ with $h_p = (b(p-1))^{-1/p}$. □

Inserting (3.93) into (3.9), we finally obtain asymptotic formulas (3.16) and (3.22).

In the case $p = 1$ using (3.8) and (3.31) for large r we get

$$\Phi(r) = 1 - \frac{1}{\xi_r} (1 + \ln \xi_r - I(\xi_r, r)) = 1 - \frac{1}{\xi_r} + b|r|(1+o(1)).$$

According to (3.92) we have $u(r) = |r|\xi_r \rightarrow \infty$ as $|r| \rightarrow \infty$, consequently, $\frac{1}{|r|\xi_r} \rightarrow 0$, and $\frac{1}{\xi_r} = o(|r|)$. Thus,

$$\Phi(r) = b|r|(1+o(1)) \quad \text{as } |r| \rightarrow \infty,$$

and asymptotic formula (3.15) is proved. Theorem 3.4 is completely proved.

3.6. The region of large deviations: Proof of Theorem 3.8

In order to justify the asymptotics in (3.24) it suffices to prove that for $x = rt(1+o(1))$ we have

$$e^{-\Phi(r)t(1+\nu_1(t))} \leq v(x, t) \leq e^{-\Phi(r)t(1+\nu_2(t))}, \tag{3.95}$$

where $\nu_j(t) \rightarrow 0$ as $t \rightarrow \infty$, $j = 1, 2$. Since the upper bound has already been proved, see (3.12), we proceed with the lower bound. Denote $\hat{r} = x/t$. Then $\hat{r} = r(1+o(1))$ as $t \rightarrow \infty$.

From the definition of ξ_r in (3.8) by the implicit function theorem we obtain that ξ_r is a smooth function of r . So is $r\xi_r$. Letting $r_0^* = \xi_r r$ and $r^* = \xi_{\hat{r}} \hat{r}$, we then have $r^* = r_0^*(1+o(1))$.

We define $\gamma^*(r^*) \in \mathbb{R}^d$ as a solution to the equation $\nabla L(\gamma) = r^*$. By Proposition 3.13 this equation has a unique solution. Moreover, $\gamma^*(r^*)$ is a smooth function of r^* . In particular, $\gamma_0^* = \gamma^*(r_0^*) = \gamma^*(r^*)(1+o(1))$, as $t \rightarrow \infty$. We recall, see Proposition 3.13 again, that for a random variable X_{γ^*} with the density $a_{\gamma^*}(x) = \frac{a(x)e^{\gamma^* \cdot x}}{\Lambda(\gamma^*)}$ its expectation is equal to r^* .

Consider a family of densities $\tilde{a}_{\gamma^*}(x) = a_{\gamma^*}(x+r^*)$ and the corresponding random variables $\tilde{X}_{\gamma^*} = X_{\gamma^*} - r^*$.

LEMMA 3.22. *There exists a neighbourhood \mathcal{O} of γ_0^* in \mathbb{R}^d such that for all $\gamma^* \in \mathcal{O}$ the density \tilde{a}_{γ^*} possesses the following properties:*

- (a) $\tilde{a}_{\gamma^*}(x) \leq Ce^{-\mu|x|}$ for some $\mu > 0$ and $C > 0$;
- (b) the matrix

$$\sigma_{ij}(\gamma^*) = \int_{\mathbb{R}^d} x_i x_j \tilde{a}_{\gamma^*}(x) dx$$

is positive definite, $\sigma(\gamma^*)\zeta \cdot \zeta \geq \mu_1|\zeta|^2$ for some $\mu_1 > 0$ and for all $\zeta \in \mathbb{R}^d$.

The constants μ , μ_1 and C do not depend on the choice of $\gamma^* \in \mathcal{O}$.

Proof. If $p = 1$ then under condition \mathbf{A}_1^s we have $|\gamma_0^*| < b$. We can choose sufficiently small neighbourhood \mathcal{O} of γ_0^* in such a way that the inequality $b - |\gamma^*| \geq \frac{1}{2}(b - |\gamma_0^*|)$ holds for all

$\gamma^* \in \mathcal{O}$. It is clear that $\Lambda(\gamma^*) \geq C > 0$ for all $\gamma^* \in \mathbb{R}^d$. This implies, in view of (3.1) and the definition of \tilde{a}_{γ^*} , the first statement of lemma with $\mu_1 = \frac{1}{2}(b - |\gamma_0^*|)$. If $p > 1$, then this statement is obvious.

The second statement of Lemma is a straightforward consequence of the first one. Indeed, it follows from (a) that there exists $R_0 > 0$ such that

$$\int_{Q_{R_0}} \tilde{a}_{\gamma^*}(x) dx \geq \frac{1}{2}$$

for all $\gamma^* \in \mathcal{O}$, here Q_{R_0} stands for the ball of radius R_0 centred at the origin. Then for any $\theta \in S^{d-1}$ we have

$$\sigma_{ij}(\gamma^*)\theta_i \cdot \theta_j = \int_{\mathbb{R}^d} (x \cdot \theta)^2 \tilde{a}_{\gamma^*}(x) dx \geq \int_{Q_{R_0} \setminus \Pi_\delta} (x \cdot \theta)^2 \tilde{a}_{\gamma^*}(x) dx,$$

where $\Pi_\delta = \{x \in Q_{R_0} : |x \cdot \theta| < \delta\}$. Due to (a) there exists $\delta_0 > 0$ such that $\int_{\Pi_{\delta_0}} \tilde{a}_{\gamma^*}(x) dx \leq \frac{1}{4}$ for all γ^* and for all $\theta \in S^{d-1}$. Therefore,

$$\int_{Q_{R_0} \setminus \Pi_{\delta_0}} (x \cdot \theta)^2 \tilde{a}_{\gamma^*}(x) dx \geq \frac{1}{4}\delta_0^2.$$

This yields (b). □

It follows from Lemma 3.22 that the local limit theorem applies to a family of i.i.d. random variables with the density \tilde{a}_{γ^*} , see [4, Theorems 19.1 and 19.2]. Therefore,

$$\tilde{a}_{\gamma^*}^{*k}(0) = (2\pi k)^{-\frac{d}{2}} |\sigma(\gamma^*)|^{-1} (1 + o(1)),$$

as $k \rightarrow \infty$, and

$$a_{\gamma^*}^{*k}(kr^*) = (2\pi k)^{-\frac{d}{2}} |\sigma(\gamma^*)|^{-1} (1 + o(1)). \tag{3.96}$$

Moreover, by [4, Theorem 19.2], the convergence is uniform in $\gamma^* \in \mathcal{O}$.

According to (3.41),

$$a^{*k}(kr^*) = a_{\gamma^*}^{*k}(kr^*) e^{-I(r^*)k}.$$

Take $k = [\frac{t}{\xi \hat{r}}]$, where $[\cdot]$ stands for the integer part. Then $kr^* = rt(1 + o(1)) = x(1 + o(1))$, as $t \rightarrow \infty$. Considering (3.9), (3.84) and (3.96) and the fact that the convergence in (3.96) is uniform in $\gamma^* \in \mathcal{O}$, we conclude that, under this choice of k ,

$$\frac{t^k}{k!} a^{*k}(kr^*) = a_{\gamma^*}^{*k}(kr^*) e^{-\Phi(\hat{r})t(1+o(1))} = e^{-\Phi(r)t(1+o(1))}.$$

This yields the desired lower bound in (3.95).

Acknowledgements. The authors are grateful to Prof. M. Lifshits for fruitful discussions on large deviation principle. The third and the fourth authors would like to thank the Department of Mathematics of the University of Bielefeld for hospitality.

References

1. F. ANDREU-VAILLO, J. M. MAZON, J. D. ROSSI and J. J. TOLEDO-MELERO, *Nonlocal diffusion problems* (American Mathematical Society, Providence, RI, 2010).
2. D. G. ARONSON, ‘Bounds for the fundamental solution of a parabolic equation’, *Bull. Amer. Math. Soc. (N.S.)* 73 (1967) 890–896.
3. A. BENDIKOV, ‘Asymptotic formulas for symmetric stable semigroups’, *Expo. Math.* 12 (1994) 381–384.
4. R. N. BHATTACHARIA and R. RANGA RAO, *Normal approximations and asymptotic expansions* (John Wiley & Sons, Hoboken, NJ, 1976).

5. C. BRANDLE, E. CHASSEIGNE and R. FERREIRA, 'Unbounded solutions of the nonlocal heat equation', *Comm. Pure Appl. Anal.* 10 (2011) 1663–1686.
6. E. CHASSEIGNE, M. CHAVES and J. D. ROSSI, 'Asymptotic behavior for nonlocal diffusion equations', *J. Math. Pures Appl.* 86 (2006) 271–291.
7. E. CHASSEIGNE, P. FELMER, J. D. ROSSI and E. TOPP, 'Fractional decay bounds for nonlocal zero order heat equations', *Bull. Lond. Math. Soc.* 46 (2014) 943–952.
8. E. B. DAVIES, 'Large deviations for heat kernels on graphs', *J. Lond. Math. Soc.* (2) 47 (1993) 65–72.
9. A. DEMBO and O. ZEITOUNI, *Large deviation techniques and applications*, 2nd edn (Springer, Berlin, 1998).
10. D. FINKELSTEIN and P. TKACHOV, 'Kesten's bound for sub-exponential densities on the real line and its multi-dimensional analogues', *Adv. in Appl. Probab.* 50 (2018), to appear.
11. A. GETAN, S. MOLCHANOV and B. VAINBERG, 'Intermittency for branching walks with heavy tails', *Stoch. Dyn.* 17 (2017) 1750044-1–1750044-14.
12. A. GRIGOR'YAN, J. HU and K.-S. LAU, 'Estimates of heat kernels for non-local regular Dirichlet forms', *Trans. Amer. Math. Soc.* 366 (2014) 6397–6441.
13. YU. KONDRATIEV, S. MOLCHANOV, A. PIATNITSKI and E. ZHIZHINA, 'Resolvent bounds for jump generators', *Appl. Anal.* 97 (2018) 323–336.
14. YU. KONDRATIEV, S. MOLCHANOV, S. PIROGOV and E. ZHIZHINA, 'On ground state of some non-local Schrödinger operators', *Appl. Anal.* 96 (2017) 1390–1400.
15. YU. KONDRATIEV, S. MOLCHANOV and B. VAINBERG, 'Spectral analysis of non-local Schrödinger operators', *J. Funct. Anal.* 273 (2017) 1020–1048.
16. J.-L. LIONS, *Quelques methodes de resolution des problemes aux limites nonlineaires* (Dunod Gauthier-Villars, Paris, 1969).
17. A. A. MOGULSKII, 'The large deviations principle for a generalized Poisson process', *Siberian Adv. Math.* 27 (2017) 160–186.
18. S. A. MOLCHANOV and E. B. YAROVAYA, 'Large deviations for a symmetric branching random walk on a multidimensional lattice', *Proc. Steklov Inst. Math.* 282 (2013) 186–201.
19. M. M. H. PANG, 'Heat kernels on graphs', *J. Lond. Math. Soc.* (2) 47 (1993) 50–64.
20. V. M. ZOLOTAREV, *One-dimensional stable distributions*, Translations of Mathematical Monographs 65 (American Mathematical Society, Providence, RI, 1986).

Alexander Grigor'yan
 School of Mathematical Sciences and LPMC
 Nankai University
 300071 Tianjin
 P.R. China

and

Department of Mathematics
 University of Bielefeld
 33501 Bielefeld
 Germany
alexander.grigorian@googlegmail.com

Andrey Piatnitski
 The Arctic University of Norway
 Campus Narvik, Postbox 385
 8505 Narvik
 Norway

and

Institute for Information Transmission
 Problems RAS
 Moscow 127051
 Russia
apiatnitski@gmail.com

Yuri Kondratiev
 Department of Mathematics
 University of Bielefeld
 33501 Bielefeld
 Germany

yukondrat@gmail.com

Elena Zhizhina
 Institute for Information Transmission
 Problems RAS
 Moscow 127051
 Russia

ejj@iitp.ru