

## Approximation in law of locally $\alpha$ -stable Lévy-type processes by non-linear regressions\*

Alexei Kulik<sup>†</sup>

### Abstract

We study a real-valued Lévy-type process  $X$ , which is *locally  $\alpha$ -stable* in the sense that its jump kernel is a combination of a ‘principal’ (state dependent)  $\alpha$ -stable part with a ‘residual’ lower order part. We show that under mild conditions on the local characteristics of a process (the jump kernel and the velocity field) the process is uniquely defined, is Markov, and has the strong Feller property. We approximate  $X$  in law by a non-linear regression  $\tilde{X}_t^x = f_t(x) + t^{1/\alpha}U_t^x$  with a deterministic *regressor term*  $f_t(x)$  and  $\alpha$ -stable *innovation term*  $U_t^x$ , and provide error estimates for such an approximation. A case study is performed, revealing different types of assumptions which lead to various choices of regressor/innovation terms and various types of the estimates. The assumptions are quite general, cover the *super-critical case*  $\alpha < 1$ , and allow non-symmetry of the Lévy kernel and unboundedness of the drift coefficient.

**Keywords:** Lévy-type process; locally  $\alpha$ -stable process; conditionally  $\alpha$ -stable approximation; parametrix method.

**AMS MSC 2010:** 60J35; 60J75; 35S05; 35S10; 47G30.

Submitted to EJP on January 23, 2019, final version accepted on July 7, 2019.

## 1 Introduction

Lévy processes are used nowadays in a wide variety of models in physics, biology, finance etc., where the random noise — by different reasons — can not be assumed Gaussian, and thus the entire model does not fit to the diffusion framework. For instance, the famous Ditlevsen model of the millennial climate changes [3] is based on the observation that the available ice-core data necessarily requires non-Gaussian noise to be included into the model. In the basic Ditlevsen model the non-Gaussian noise is  $\alpha$ -stable; nowadays it is understood that it would be physically more realistic to have the parameters of the noise state-dependent; e.g. the skewness parameter should be positive in the cold glacial periods and negative in the warmer interstadials. The similar

\*Supported by Wrocław University of Science and Technology, grant 0401/0155/18.

<sup>†</sup>Wrocław University of Science and Technology, Faculty of Pure and Applied Mathematics, Wybrzeże Wyspiańskiego Str. 27, 50-370 Wrocław, Poland. E-mail: [oleksii.kulyk@pwr.edu.pl](mailto:oleksii.kulyk@pwr.edu.pl)

problem appears in many other models with state-dependent parameters, which gives a natural background for the notion of a *Lévy-type process*. The latter is understood as a (kind of) a Lévy process whose characteristic triplet is allowed to depend on the current value of the process; we refer to [2] for a detailed introduction, see also Section 2 below. The definition of a Lévy-type process has the same spirit with the classical Kolmogorov's definition of a diffusion process as a location-dependent Brownian motion with a drift. However, to the contrast with the classical theory of diffusions, in the general theory of such Lévy-type process some principal questions remain unsolved in general, e.g.

- (I) for a given set of local characteristics, is the corresponding Lévy-type process uniquely defined?
- (II) what kind of local properties of the law of the process can be derived, and under which assumptions on characteristics?

Not being able to discuss in details a considerable list of references devoted to these questions, we refer to [10], [20], [13] for such a discussion, and only note that the available methods contain a list of limitations, which exclude from the consideration many natural and physically relevant Lévy-type models.

In this paper we provide a detailed study of one class of Lévy-type processes, which is highly relevant for applications and, on the other hand, reveals numerous hidden challenges which one encounters while trying to resolve the above questions (I), (II) in general Lévy-type setting. The class to be studied can be shortly described as a mixture of a real-valued  $\alpha$ -stable-type process with state dependent *drift*, *intensity*, and *skewness* parameters on one hand, and a certain (state dependent) lower order 'nuisance' part on the other hand; see a detailed definition in Section 2. The  $\alpha$ -stable noise, because of its scaling property, has an exceptional importance in physical applications, and at the same time there are strong reasons to require the parameters of the noise to be state-dependent, likewise to the Ditlevsen model discussed above. Presence of the 'residual' lower order part is quite reasonable, as well. Namely, this part allows one to introduce a wide spectrum of tempering/damping effects for the tails of the noise, which combines both the  $\alpha$ -stable and Gaussian regimes (see [21]) and thus appear frequently in physical models (see [22] and references therein). On the other hand, a lower order *microstructural* noise terms without a specified inner structure appear quite naturally in finance models; see [1] for a detailed discussion.

For such a *locally  $\alpha$ -stable* Lévy-type model we prove the corresponding process to be uniquely defined and to be a Markov process with strong Feller property, thus resolving the general question (I). To approach the question (II), we specify a family of  $\alpha$ -stable probability densities  $g^{t,x}, t > 0, x \in \mathbb{R}$  and a function  $f_t(x), t \geq 0, x \in \mathbb{R}$  such that the transition density  $p_t(x, y)$  of  $X$  has representation

$$p_t(x, y) = \frac{1}{t^{1/\alpha}} g^{t,x} \left( \frac{y - f_t(x)}{t^{1/\alpha}} \right) + R_t(x, y), \tag{1.1}$$

where the residual kernel  $R_t(x, y)$  is negligible (in a certain sense) as  $t \rightarrow 0$ . This representation essentially means that, conditioned by  $X_0 = x$ , process  $X$  admits approximation in law by the non-linear regression

$$\tilde{X}_t^x = f_t(x) + t^{1/\alpha} U_t^x, \quad t > 0, \tag{1.2}$$

where  $U^{t,x}$  is a random variable with the  $\alpha$ -stable distribution density  $g^{t,x}$ . We call  $f_t(x)$  a (deterministic) *regressor term* for  $X$ , and  $U^{t,x}$  an  *$\alpha$ -stable innovation term*. It is natural to call (1.2) a *conditionally  $\alpha$ -stable approximation* to  $X$ , in the same spirit with the standard *conditionally Gaussian approximation* for a diffusion. However, we will see

that the regressor term  $f_t(x)$  in general should have a more sophisticated form than just  $x + b(x)t$ , typical for the diffusion case.

Our study is based on the *parametrix method*, which in the diffusion case is a classical analytical tool to construct and investigate transition densities. To apply this method in the (non-Gaussian) Lévy-type setting, we modify it substantially; here we outline the most crucial change. The classical parametrix method relies on the fact that a (properly chosen) ‘zero order approximation’ to the unknown  $p_t(x, y)$  and corresponding ‘differential error term’ (see Section 4.1 below for these definitions) follow certain prior bounds, which then propagate to the transition density  $p_t(x, y)$ . For diffusions these kernels are Gaussian; for certain  $\alpha$ -stable-type models similar kernel estimates with  $\alpha$ -stable kernels are available as well; see [12] and [14] for the cases  $\alpha > 1$  and  $\alpha \leq 1, b \equiv 0$  respectively, and [19], where in the technically more involved *super-critical regime* with  $\alpha < 1$  and non-trivial  $b$  the kernel estimates are obtained as a combination of stable kernels with deterministic flows. However, all these models are ‘regular’ in the sense that the Lévy kernel of the noise is assumed to have a density w.r.t. the Lebesgue measure. Presence of singular terms may change situation drastically; see Example 3.7 below, where  $p_t(x, \cdot)$  is unbounded and thus kernel estimates simply fail. The same effect have been discussed in the recent preprint [18] for solutions of multidimensional SDEs with cylindrical  $\alpha$ -stable noise and non-trivial rotation, see [18, Remark 4.23].

To study such highly singular settings, we adopt the following two-stage scheme. First, we establish integral-in- $y$  estimates (actually, operator norm estimates in  $C_\infty$ ) and perform the parametrix method with the convergence of corresponding series understood in this ( $L_1$ ) sense. This resolves question (I) and gives  $L_1$ -estimates for the error term  $R_t(x, y)$  in (1.1). Second, we analyse the series representation for  $p_t(x, y)$  and clarify additional assumptions, which one should require in order to get stronger types of estimates for  $R_t(x, y)$ : uniform-in- $(x, y)$  and kernel estimates. This scheme is motivated by perspective applications, where the choice among several types of estimates will allow one to avoid limitations in the model’s assumptions when a particular application is considered. We plan to use integral-in- $y$  estimates in the proof of Local Asymptotic (Mixed) Normality property for statistical models with discretely observed Lévy-type processes (this is an ongoing project with A. Kohatsu-Higa) and, combined with uniform-in- $(x, y)$  estimates, in the asymptotic study of the Least Absolute Deviation estimator for a drift parameter (this is an ongoing project with H. Masuda). Motivated by these applications, we restrict the current exposition by the real-valued case with constant  $\alpha$ . The multidimensional locally stable-like model with state-dependent  $\alpha = \alpha(x)$  is considered in the widest generality in the companion paper [11].

The structure of the paper is the following. In Section 2 we introduce the notation and specify the model. In Section 3 we specify the conditions and formulate the main results. For these results we also provide a discussion, including examples, possible extensions, and related references. In Section 4 we separately explain the essence of the parametrix method and derive the corresponding integral representation of the (candidate for) the transition probability density of the required process. Sections 5–7 respectively contain the proofs of three main results, Theorem 3.3–Theorem 3.5. The proofs of certain technicalities, which otherwise would make the reading much more difficult, are postponed to Appendix.

## Acknowledgments

The author is grateful to Arturo Kohatsu-Higa and Hiroki Masuda for insightful discussions which clarified for him the perspectives of the parametrix technique in application to statistics, and for numerous helpful comments about the previous version of the draft. The author gladly expresses a particular respect to Victoria Knopova and

René Schilling: the numerous discussion on this paper and on the companion one [11] made a deep impact on the style of the final exposition and saved the author from several pitiful mistakes. The work on the project has been principally finished during the visit of the author to the Technical University of Dresden (Germany); the author is very grateful to the Technical University of Dresden and especially to René Schilling for their support and hospitality. Finally, the author is deeply grateful to referees for their attention to the paper and very helpful comments.

## 2 Notation and preliminaries

In what follows,  $C_\infty$  denotes the class of continuous functions  $\mathbb{R} \rightarrow \mathbb{R}$  vanishing at  $\infty$ , and  $C_0$  denotes the class of continuous functions with compact support. By  $C_\infty^2, C_0^2$  we denote the classes of twice differentiable functions  $f$  such that  $f, f', f''$  belong to  $C_\infty$  or  $C_0$ , respectively. A Lévy-type operator  $L$  with the domain  $C_\infty^2$  is defined by

$$Lf(x) = b(x)f'(x) + \frac{1}{2}a(x)f''(x) + \int_{\mathbb{R}} \left( f(x+u) - f(x) - uf'(x)1_{|u|\leq 1} \right) \mu(x; du), \quad f \in C_\infty^2. \tag{2.1}$$

Here  $b : \mathbb{R} \rightarrow \mathbb{R}, a : \mathbb{R} \rightarrow \mathbb{R}^+$  are given measurable functions, and  $\mu(x; du)$  is a Lévy kernel; that is, a measurable function w.r.t.  $x$  and a Lévy measure w.r.t.  $du$ .

There are two natural and closely related ways to associate a Lévy-type process  $X$  with the Lévy type operator  $L$ . Within the first one,  $X$  is a time-homogeneous Markov process which generates a Feller semigroup (that is, a strongly continuous semigroup in  $C_\infty$ ) such that its generator  $A$  coincides with  $L$  on  $C_\infty^2$  (or, which is slightly more general, on  $C_0^2$ ). The second way is based on the notion of the Martingale Problem (MP). Recall that a process  $X$  is said to be a solution to the martingale problem  $(L, \mathcal{D})$ , if for every  $f \in \mathcal{D}$  the process

$$f(X_t) - \int_0^t Lf(X_s) ds, \quad t \geq 0$$

is a martingale w.r.t. the natural filtration for  $X$ . A martingale problem  $(L, \mathcal{D})$  is said to be well posed in  $D(\mathbb{R}^+)$  (the Skorokhod space of càdlàg functions), if for any probability measure  $\pi$  on  $\mathbb{R}$  there exists a solution  $X$  to this problem with càdlàg trajectories and  $\text{Law}(X_0) = \pi$ , and for any two such solutions their distributions in  $D(\mathbb{R}^+)$  coincide. By the second definition, Lévy-type process associated to  $L$  is a solution to the MP  $(L, \mathcal{D})$  with  $L$  given by (2.1) and  $\mathcal{D} = C_\infty^2$  or  $C_0^2$ .

Arbitrary Lévy process  $X$  satisfies both of the above definitions; the corresponding operator  $L$  is defined by (2.1) with  $b(x) \equiv b, a(x) \equiv a, \mu(x, \cdot) \equiv \mu$ , where  $(b, a, \mu)$  is the characteristic triplet for  $X$ . This explains the name Lévy-type process, which we use systematically. The principal problem (I) outlined in the Introduction can be now formulated precisely: given a triplet  $b(x), a(x), \mu(x, du)$ , is the Lévy-type process, associated with  $L$ , uniquely defined in either/both of two ways explained above? That is, does there exist a unique Feller process with the prescribed restriction of the generator, or/and is the MP  $(L, \mathcal{D})$  well posed? The problem (II) then would be to describe — in the most explicit way it is possible — the transition probability  $P_t(x, dy)$  of the process  $X$ .

We will study these two questions in the particular setting of locally  $\alpha$ -stable Lévy-type operators/ processes, which we now introduce. A real-valued  $\alpha$ -stable process is a Lévy process which lacks the diffusion term ( $a = 0$ ), may contain a non-trivial shift ( $b \neq 0$ ), and has the Lévy measure

$$\mu(du) = \mu^{(\alpha; \lambda, \rho)}(du) := \lambda \frac{1 + \rho \operatorname{sgn} u}{|u|^{\alpha+1}} du.$$

Taking the intensity and skewness parameters state dependent,  $\lambda : \mathbb{R} \rightarrow \mathbb{R}^+, \rho : \mathbb{R} \rightarrow$

$[-1, 1]$ , we obtain an  $\alpha$ -stable Lévy kernel

$$\mu^{(\alpha)}(x; du) := \mu^{(\alpha; \lambda(x), \rho(x))}(du) = \lambda(x) \frac{1 + \rho(x) \operatorname{sgn} u}{|u|^{\alpha+1}} du.$$

Our actual Lévy kernel has the form

$$\mu(x; du) = \mu^{(\alpha)}(x; du) + \nu(x; du); \tag{2.2}$$

that is, it is a perturbation of an  $\alpha$ -stable kernel by a certain ‘residual’ kernel  $\nu(x; du)$ . The residual kernel  $\nu(x; du)$  is allowed to be signed, and we denote by  $\nu_+(x; du)$ ,  $\nu_-(x; du)$  the positive (resp. the negative) parts of its Hahn decomposition  $\nu(x; du) = \nu_+(x; du) - \nu_-(x; du)$ . The negative part  $\nu_-(x; du)$  is assumed to be dominated by  $\mu^{(\alpha)}(x; du)$ , and  $|\nu|(x; du) = \nu_+(x; du) + \nu_-(x; du)$  — the variation of  $\nu(x; du)$  — is assumed to be a Lévy kernel. The main assumption imposed on the residual kernel is that, uniformly in  $x$ , the Blumenthal-Gettoor activity index for  $|\nu|$  is strictly smaller than  $\alpha$ ; that is, for some  $\beta < \alpha$

$$|\nu|(x; \{|u| > r\}) \leq Cr^{-\beta}, \quad r \in (0, 1]. \tag{2.3}$$

Since the Blumenthal-Gettoor index for an  $\alpha$ -stable Lévy measure equals  $\alpha$ , this condition actually means that the small jump behavior of  $\mu(x; du)$  is asymptotically the same as for its  $\alpha$ -stable part  $\mu^{(\alpha)}(x; du)$ , and this is our reason to call the kernel (2.2) *locally  $\alpha$ -stable*.

Summarizing all the above, we specify the *locally  $\alpha$ -stable* Lévy-type operator as an operator of the form (2.1) with  $\mu(x, du)$  given by (2.2),  $a(x) \equiv 0$ , and possibly non-trivial  $b(x)$ ; that is,

$$Lf(x) = b(x)f'(x) + \int_{\mathbb{R}} \left( f(x+u) - f(x) - uf'(x)1_{|u|\leq 1} \right) \left( \mu^{(\alpha)}(x; du) + \nu(x; du) \right). \tag{2.4}$$

### 3 The main results

In this section we specify the conditions imposed on the model, formulate the main results, and make a discussion which includes examples, possible extensions, and related references.

#### 3.1 Conditions

In what follows,  $L$  is the Lévy-type operator defined by (2.4), and (2.3) is assumed. Throughout the paper we denote by  $C$  a generic constant whose particular value may vary from place to place. We define the *compensated drift coefficient* by

$$\tilde{b}(x) = b(x) - 1_{\alpha < 1} \int_{|u|\leq 1} u \mu^{(\alpha)}(x; du) - 1_{\beta < 1} \int_{|u|\leq 1} u \nu(x; du),$$

and assume the following.

**H<sup>drift</sup>.** (On the compensated drift coefficient). There exists index  $\eta \in [0, 1]$  satisfying the *balance condition*

$$\alpha + \eta > 1, \tag{3.1}$$

such that

$$|\tilde{b}(x) - \tilde{b}(y)| \leq C|x - y|^\eta, \quad |x - y| \leq 1. \tag{3.2}$$

**H<sup>( $\alpha$ )</sup>.** (On coefficients  $\lambda, \rho$  of the kernel  $\mu^\alpha$ ).

- (i)  $\lambda, \rho$  are Hölder continuous with some index  $\zeta \in (0, \alpha)$ ;

(ii) for some  $0 < \lambda_{\min} < \lambda_{\max}$ ,

$$\lambda_{\min} \leq \lambda(x) \leq \lambda_{\max}.$$

$\mathbf{H}^\nu$ . (On the residual kernel  $\nu$ ). We deal with two types of upper bounds:

(i) (weak bound) the kernel  $\nu(x, du)$  satisfies (2.3) and the following ‘tail condition’:

$$\sup_{x \in \mathbb{R}} |\nu|(x, \{|u| \geq R\}) \rightarrow 0, \quad R \rightarrow \infty; \tag{3.3}$$

(ii) (strong bound) the kernel has the density

$$q_\nu(x, u) = \frac{\nu(x, du)}{du},$$

which satisfies

$$|q_\nu(x, u)| \leq C|u|^{-\beta-1}1_{|u| \leq 1} + C|u|^{-\gamma-1}1_{|u| > 1} \tag{3.4}$$

with some  $\beta \in (0, \alpha), \gamma > 0$ .

$\mathbf{H}^{cont}$ . (Continuity assumptions). The kernel  $\nu(x, du)$  is assumed to have the following weak continuity property: for any  $f \in C(\mathbb{R})$  with compact support in  $\mathbb{R} \setminus \{0\}$ , the function

$$x \mapsto \int_{\mathbb{R}} f(u)\nu(x, du) \tag{3.5}$$

is continuous. The drift coefficient  $b$  is assumed to be continuous.

Note that, thanks to condition  $\mathbf{H}^{(\alpha)}$ , the continuity of (3.5) yields similar continuity for the entire kernel  $\mu(x, du) = \mu^{(\alpha)}(x, du) + \nu(x, du)$ .

**Remark 3.1.** In the *super-critical regime*  $\alpha < 1$ , the balance condition (3.1) is close to the necessary one for the process to be well defined. This observation dates back to [24], where a natural example of an SDE driven by a symmetric additive  $\alpha$ -stable noise with  $\eta$ -Hölder continuous  $b$  is given, which has two different weak solutions. We emphasise that in the current setting the balance condition involves the compensated drift coefficient  $\tilde{b}$  instead of the original  $b$ .

**Remark 3.2.** A good way to understand the role of the continuity condition  $\mathbf{H}^{cont}$  is to observe that, if (say)  $\nu \equiv 0$  and  $b$  is discontinuous, it is impossible for the operator (2.4) that  $Lf$  is continuous for all  $f \in C_0^2$ , and thus the first definition of the Lévy-type process becomes inappropriate. This complication is of a technical kind, which is not related to our main goal to derive representation (1.1) for the transition probability of the process. Thus we adopt  $\mathbf{H}^{cont}$  and avoid further technical complications.

### 3.2 The main statements

Our first main result uniquely identifies a locally  $\alpha$ -stable Lévy type process with given characteristics.

**Theorem 3.3.** *Let  $L$  be given by (2.4) and conditions  $\mathbf{H}^{drift}$ ,  $\mathbf{H}^{(\alpha)}$ ,  $\mathbf{H}^\nu$  (i), and  $\mathbf{H}^{cont}$  hold true. Then the martingale problem  $(L, C_0^2)$  is well posed in  $D(\mathbb{R}^+)$  and, at the same time, the solution  $X$  of this martingale problem is the unique Feller process, whose generator  $A$  restricted to  $C_0^\infty$  coincides with  $L$ . This process is strong Feller and possesses a transition probability density  $p_t(x, y)$ .*

Next, we provide several versions of the representation (1.1) with different types of bounds on the residual kernel  $R_t(x, y)$ , depending on the actual assumptions imposed in the characteristics of the process. Following the two-stage scheme outlined in the Introduction, we first do this under the basic set of conditions  $\mathbf{H}^{drift}$ ,  $\mathbf{H}^{(\alpha)}$ ,  $\mathbf{H}^\nu$  (i), and then discuss modifications under additional assumptions.

Let us introduce more notation. By  $g^{(\lambda, \rho, v)}(w)$  we denote the density of the  $\alpha$ -stable distribution with the intensity  $\lambda$ , skewness  $\rho$ , and a shift  $v$ :

$$g^{(\lambda, \rho, v)}(w) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iw\xi + \Psi_{\alpha}^{(\lambda, \rho, v)}(\xi)} d\xi, \tag{3.6}$$

$$\Psi_{\alpha}^{(\lambda, \rho, v)}(\xi) = i\xi v + \int_{\mathbb{R}} \left( e^{iu\xi} - 1 - iu\xi 1_{|u| \leq 1} \right) \mu^{(\alpha, \lambda, \rho)}(du). \tag{3.7}$$

Next, we denote

$$\delta_{\eta} = \frac{\eta + \alpha - 1}{\alpha} > 0, \quad \delta_{\zeta} = \frac{\zeta}{\alpha} > 0, \quad \delta_{\beta} = \frac{\alpha - \beta}{\alpha} > 0, \quad \delta_{\eta, \zeta, \beta} = \min(\delta_{\eta}, \delta_{\zeta}, \delta_{\beta});$$

note that the positivity of  $\delta_{\eta}$  is just the balance condition (3.1). We fix (arbitrary) positive  $\delta < \delta_{\eta, \zeta, \beta}$ . We also fix (arbitrary)  $T > 0$  and furthermore consider  $t \leq T$ , only. Denote

$$m_t^{\mu}(x) = \int_{t^{1/\alpha} < |u| \leq 1} u \mu(x, du), \quad b_t(x) = b(x) - m_t^{\mu}(x), \tag{3.8}$$

the *partial compensator* of the kernel (2.2) with the truncation level  $t^{1/\alpha}$ , and *partially compensated* drift coefficient, respectively. Define the corresponding *mollified* coefficient

$$B_t(x) = \int_{\mathbb{R}} b_t(x - z) \frac{1}{2\sqrt{\pi}t^{1/\alpha}} e^{-z^2 t^{-2/\alpha}} dz.$$

This coefficient is chosen in such a way that

$$\sup_x |b_t(x) - B_t(x)| \leq Ct^{-1+1/\alpha+\delta}, \tag{3.9}$$

$$\text{Lip}(B_t) = \sup_{x \neq y} \frac{|B_t(x) - B_t(y)|}{|x - y|} \leq Ct^{-1+\delta}, \tag{3.10}$$

see Appendix A.1 (recall that  $\delta < \delta_{\beta} < 1$ ). We define  $\chi_s(x)$ ,  $s \geq 0$ ,  $x \in \mathbb{R}$  as the solution to the Cauchy problem

$$\frac{d}{ds} \chi_s(x) = B_s(\chi_s(x)), \quad \chi_0(x) = x. \tag{3.11}$$

Note that by (3.10) the family of Lipschitz constants  $\text{Lip}(B_t)$ ,  $t > 0$  is integrable on any finite segment, thus  $\chi_t(x)$  is uniquely defined by the classical Picard successful approximation procedure. We define

$$\lambda_t(x) = \frac{1}{t} \int_0^t \lambda(\chi_s(x)) ds, \quad \rho_t(x) = \frac{1}{t\lambda_t(x)} \int_0^t \lambda(\chi_s(x)) \rho(\chi_s(x)) ds;$$

that is,  $\lambda_t(x)$  and  $\lambda_t(x)\rho_t(x)$  are the averages of the functions  $\lambda(\cdot)$ ,  $\lambda(\cdot)\rho(\cdot)$  along the trajectory  $\chi_{\cdot}(x)$  on the segment  $[0, t]$ . We also denote  $v(x) = 2\lambda(x)\rho(x)$ ,

$$W_{\alpha}(t; s) = t^{-1/\alpha} \int_{s^{1/\alpha}}^{t^{1/\alpha}} \frac{dr}{r^{\alpha}} = \begin{cases} \frac{1}{1-\alpha} t^{-1/\alpha} (t^{1/\alpha-1} - s^{1/\alpha-1}), & \alpha \neq 1, \\ t^{-1} (\log t - \log s), & \alpha = 1, \end{cases} \quad 0 \leq s \leq t, \tag{3.12}$$

and put

$$v_t(x) = \int_0^t v(\chi_s(x)) W_{\alpha}(t; s) ds.$$

Note that

$$\int_0^t W_\alpha(t, s) ds = 1; \tag{3.13}$$

that is,  $v_t(x)$  is also an average of  $v(\cdot)$  along the trajectory  $\chi_\cdot(x)$ , but with respect to a certain (non-uniform) probability distribution on  $[0, t]$ . We finally define

$$g^{t,x}(w) = g^{(\lambda_t(x), \rho_t(x), v_t(x))}(w), \tag{3.14}$$

the  $\alpha$ -stable density with the ‘ $\chi$ -averaged’ parameters  $\lambda_t(x), \rho_t(x), v_t(x)$  defined above.

Now we are ready to state our second main result. Recall that we consider  $t \in [0, T]$ , where  $T$  is arbitrary but fixed; the particular values of the constants  $C$  below may depend on  $T$  and particular choice of  $\delta < \delta_{\eta, \zeta, \beta}$ .

**Theorem 3.4.** *I. Let conditions  $\mathbf{H}^{drift}$ ,  $\mathbf{H}^{(\alpha)}$ ,  $\mathbf{H}^\nu$  (i), and  $\mathbf{H}^{cont}$  hold true. Then*

$$p_t(x, y) = \frac{1}{t^{1/\alpha}} g^{t,x} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) + R_t(x, y), \tag{3.15}$$

where

$$\sup_x \int_{\mathbb{R}} |R_t(x, y)| dy \leq Ct^\delta. \tag{3.16}$$

II. Assume in addition that for some  $\delta_\nu > 0$

$$\sup_{w \in \mathbb{R}} \left| t^{-1/\alpha} \int_{\mathbb{R}} \nu \left( x; \{ |u| > t^{1/\alpha}, |x + u - w| \leq t^{1/\alpha} \} \right) dx \right| \leq Ct^{-1+\delta_\nu}. \tag{3.17}$$

Then

$$\sup_{x,y} |R_t(x, y)| \leq Ct^{-1/\alpha+\delta_\infty}, \quad \delta_\infty = \min(\delta, \delta_\nu). \tag{3.18}$$

Our last main result provides a point-wise kernel estimate for the residual term  $R_t(x, y)$  under the stronger assumption  $\mathbf{H}^\nu$  (ii). Denote

$$G_t^{(\alpha, \beta, \gamma)}(x, y) = \begin{cases} t^{-1/\alpha}, & |y - x| \leq (t^{1/\alpha} \wedge 1), \\ t^{\beta/\alpha} |y - x|^{-\beta-1}, & (t^{1/\alpha} \wedge 1) < |y - x| \leq 1, \\ t^{\beta/\alpha} |y - x|^{-\gamma-1}, & |y - x| > 1. \end{cases} \tag{3.19}$$

**Theorem 3.5.** *Let conditions  $\mathbf{H}^{drift}$ ,  $\mathbf{H}^{(\alpha)}$ ,  $\mathbf{H}^\nu$  (ii), and  $\mathbf{H}^{cont}$  hold true. Then*

$$|R_t(x, y)| \leq Ct^\delta G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y) + Ct^{\delta'} G_t^{(\alpha, \beta', \gamma')}(\chi_t(x), y) \tag{3.20}$$

with

$$\beta' = \max(\beta, \alpha - \zeta), \quad \gamma' = \min(\alpha, \gamma), \quad \delta' = \frac{\alpha - \beta'}{\alpha} > 0. \tag{3.21}$$

As a direct corollary, we get an upper bound for the entire transition probability density  $p_t(x, y)$ . Denote  $G^{(\alpha)}(x) = |x|^{-\alpha-1} \wedge 1$ .

**Corollary 3.6.** *Let conditions  $\mathbf{H}^{drift}$ ,  $\mathbf{H}^{(\alpha)}$ ,  $\mathbf{H}^\nu$  (ii), and  $\mathbf{H}^{cont}$  hold true. Then*

$$p_t(x, y) \leq \frac{C}{t^{1/\alpha}} G^{(\alpha)} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) + Ct |y - \chi_t(x)|^{-\gamma'-1} 1_{|y - \chi_t(x)| > 1}, \quad t \in (0, T]. \tag{3.22}$$

The following two examples show that there is a the substantial difference between three types of the estimates given above: (i) integral-in- $y$  (Theorem 3.4, I); (ii) uniform-in- $(x, y)$  (Theorem 3.4, II); (iii) kernel (Theorem 3.5). The first example shows that, for singular kernels  $\mu(x; du)$ , the estimates (ii), (iii) may simply fail.



**Example 3.7.** Let the ‘nuisance part’ of the noise correspond to the possibility of the process  $X_t$  to jump, at Poisson time instants, to the point 0; that is,  $\nu(x, du) = \delta_{-x}(du)$ . Then

$$p_t(x, y) \geq e^{-1} \int_0^t p_{t-s}^{(\alpha)}(0, y) ds,$$

where  $p_t^{(\alpha)}(x, y)$  denotes the transition probability density for the process with the kernel  $\mu^{(\alpha)}(x, du)$ . If  $b \equiv 0$ ,  $\lambda \equiv 1$ ,  $\rho \equiv 0$ , then

$$p_t^{(\alpha)}(x, y) \asymp t^{-1/\alpha} G^{(\alpha)}\left(\frac{y-x}{t^{1/\alpha}}\right),$$

and thus for  $\alpha \leq 1$  the function  $p_t(x, y)$  is unbounded at the vicinity of the point  $y = 0$ .

The difference between the kernel and uniform-in- $(x, y)$  estimates is more subtle. Of course, the kernel estimates yield both the integral-in- $y$  and uniform-in- $(x, y)$  estimates, but the cost is that the (strong) condition  $\mathbf{H}^\nu$  (ii) is needed, which in particular requires  $\nu(x, du)$  to be smooth. This may be too restrictive when a model with a microstructural residual noise in the spirit of [1] is considered. Our second example shows that the additional assumption (3.17), which guarantees uniform-in- $(x, y)$  bounds, is substantially weaker than  $\mathbf{H}^\nu$  (ii), and can hold true for singular nuisance kernels.

**Example 3.8.** Let  $\nu(x, du) = \nu(du)$ , then simply by the Fubini theorem and (2.3) we have

$$\begin{aligned} & \left| t^{-1/\alpha} \int_{\mathbb{R}} \nu\left(x; \{|u| > t^{1/\alpha}, |x+u-w| \leq t^{1/\alpha}\}\right) dx \right| \\ & \leq \int_{|u| > t^{1/\alpha}} \left( t^{-1/\alpha} \int_{|x+u-w| \leq t^{1/\alpha}} dx \right) |\nu|(du) \leq Ct^{-\beta/\alpha} = Ct^{-1+\delta_\beta}. \end{aligned}$$

More generally, let  $\nu(x; du)$  possess a bound

$$|\nu|(x, du) \leq \nu'(v : c(x, v) \in du),$$

where  $\nu'$  is a Lévy measure satisfying (2.3) and  $c(x, u)$  satisfies  $|c(x, u)| \leq C|u|$ , for each  $u$  the function  $x + c(x, u)$  is  $C^1$  and is invertible (in  $x$ ), and

$$|1 + c'_x(x, u)|^{-1} \leq C. \tag{3.23}$$

Then we can obtain (3.17) first changing the variables  $x' = x + c(x, u)$  and then using the Fubini theorem and (2.3) in the same way as above.

### 3.3 SDEs

For the reader’s convenience, we formulate separately the version of the above results in the case where the process  $X$  is a solution to an SDE. Consider the SDE

$$dX_t = b(X_t) dt + \sigma(X_{t-}) dZ_t + \int_{|u| \leq 1} c(X_{t-}, u) \tilde{N}(dt, du) + \int_{|u| > 1} c(X_{t-}, u) N(dt, du) \tag{3.24}$$

where  $Z$  is an  $\alpha$ -stable process,  $N(dt, du)$  is an independent of  $Z$  Poisson point measure with the compensator  $dt\nu'(du)$ , and  $\tilde{N}(dt, du) = N(dt, du) - dt\nu'(du)$  is the corresponding martingale measure. Assume that  $Z$  has the characteristic triplet  $(0, 0, \mu^{(\alpha; \lambda, \rho)})$  and  $|c(x, u)| \leq C|u|$ . Denote

$$\tilde{b}(x) = b(x) - 1_{\alpha < 1} \frac{2\lambda\rho}{1-\alpha} \sigma(x) - 1_{\beta < 1} \int_{|u| \leq 1} c(x, u) \nu'(du).$$

**Proposition 3.9.** *Let the following assumptions hold:*

- $\tilde{b}$  satisfies  $\mathbf{H}^{drift}$ ;
- $\sigma$  is  $\zeta$ -Hölder continuous and for some  $c_1, c_2 > 0$

$$c_1 \leq \sigma(x) \leq c_2;$$

- for some  $\beta < \alpha$ ,

$$\nu'(x; \{|u| > r\}) \leq Cr^{-\beta}, \quad r \in (0, 1];$$

- the functions  $b(x)$  and  $x \mapsto c(x, \cdot) \in L_1((u^2 \wedge 1)\nu'(du))$  are continuous.

Then the SDE (3.24) has unique weak solution  $X$ , and this solution is a strong Feller Markov process. The transition probability of this process has a density  $p_t(x, y)$  which has representation (1.1), where

- the regressor  $\mathfrak{f}_t(x) = \chi_t(x)$  is defined by (3.11) with  $B_t(x)$  which corresponds to

$$b_t(x) = b(x) - 2\lambda\rho\sigma(x) \int_{t^{1/\alpha}}^1 \frac{du}{u^\alpha} - \int_{t^{1/\alpha} < |u| \leq 1} c(x, u) \nu'(du)$$

- the density of the  $\alpha$ -stable innovation term has the form  $g^{t,x}(w) = g^{(\lambda_t(x), \rho, \nu_t(x))}(w)$  with

$$\lambda_t(x) = \frac{\lambda}{t} \int_0^t \sigma(\chi_s(x))^\alpha ds, \quad \nu_t(x) = 2\lambda\rho \int_0^t \sigma(\chi_s(x))^\alpha W_\alpha(t; s) ds;$$

- the residual term  $R_t(x, y)$  satisfies (3.16).

In addition,

- if the function  $x + c(x, u)$  is  $C^1$ , is invertible in  $x$  and (3.23) holds, then the residual term  $R_t(x, y)$  satisfies (3.18);
- if

$$\frac{\nu'(du)}{du} = C|u|^{-\beta-1} 1_{|u| \leq 1} + C|u|^{-\gamma-1} 1_{|u| > 1},$$

$c(x, \cdot) \in C^1$  and

$$\inf_{x,u} |c'_u(x, u)| > 0,$$

then (3.20) holds.

The Lévy-type operator, which formally corresponds to the SDE (3.24) is given by

$$\begin{aligned} Lf(x) &= b(x)f'(x) + \int_{\mathbb{R}} \left( f(x + \sigma(x)u) - f(x) - \sigma(x)uf'(x)1_{|u| \leq 1} \right) \mu^{(\alpha; \lambda, \rho)}(du) \\ &\quad + \int_{\mathbb{R}} \left( f(x + c(x, u)) - f(x) - c(x, u)uf'(x)1_{|u| \leq 1} \right) \nu'(du). \end{aligned}$$

Then the uniqueness of the weak solution to the SDE is close to the well posedness of the MP  $(L, C_0^\infty)$ ; for a (simple) formal argument which connects these two notions see e.g. [19, Section 4.3]. Thus the required statements follow from Theorems 3.3–3.5 by simple re-arrangements.

### 3.4 Possible extensions

Let us briefly discuss several possible modifications and extensions of the main results. First, let us note that the case of state-dependent  $\alpha = \alpha(x)$  can be treated similarly, but with a more sophisticated and less transparent estimates. We postpone its study to the companion paper [11], where the multidimensional locally  $\alpha$ -stable model is considered in the widest possible generality. It is also visible that the sensitivities (i.e. derivatives) of  $p_t(x, y)$  w.r.t.  $t$  and external parameters can be treated with the same method; in particular we refer to [10], [19] for representations and bounds for  $\partial_t p_t(x, y)$  and to [6] for an application of such bounds in the accuracy bounds for approximation of integral functionals. In order not to overextend the exposition, in the current paper we do not address the sensitivities, leaving their study to a further research.

Next, let us mention that the particular form of the conditionally  $\alpha$ -stable approximation (1.1) obtained in Theorem 3.4 is not the only possible one. Namely, one can change consistently the regressor  $f_t(x) = \chi_t(x)$  and the  $\alpha$ -stable innovation term, providing the following alternative representation, which may be more convenient e.g. for simulation purposes. Define for a given  $t > 0$  the family  $\bar{\chi}_s^t(x)$ ,  $s \in [0, t]$ ,  $x \in \mathbb{R}$  as the solution to the Cauchy problem

$$\frac{d}{ds} \bar{\chi}_s^t(x) = B_t(\bar{\chi}_s^t(x)), \quad \bar{\chi}_0^t(x) = x,$$

and put

$$\bar{\lambda}_t(x) = \frac{1}{t} \int_0^t \lambda(\bar{\chi}_s^t(x)) ds, \quad \bar{\rho}_t(x) = \frac{1}{t\lambda_t(x)} \int_0^t \lambda(\bar{\chi}_s^t(x))\rho(\bar{\chi}_s^t(x)) ds,$$

$$\bar{g}^{t,x}(w) = g^{(\bar{\lambda}_t(x), \bar{\rho}_t(x), 0)}(w).$$

**Proposition 3.10.** *Let conditions  $\mathbf{H}^{drift}$ ,  $\mathbf{H}^{(\alpha)}$ ,  $\mathbf{H}^\nu$  (i), and  $\mathbf{H}^{cont}$  hold true. Then*

$$p_t(x, y) = \frac{1}{t^{1/\alpha}} \bar{g}^{t,x} \left( \frac{y - \bar{\chi}_t^t(x)}{t^{1/\alpha}} \right) + \bar{R}_t(x, y), \tag{3.25}$$

where  $\bar{R}_t(x, y)$  satisfies (3.16). Under the additional condition (3.17)  $\bar{R}_t(x, y)$  satisfies (3.18), and under the condition  $\mathbf{H}^\nu$  (ii) the term  $\bar{R}_t(x, y)$  satisfies (3.20). In the latter case,  $\chi_t(x)$  in the right hand side of (3.20) can be replaced by  $\bar{\chi}_t^t(x)$ .

*Sketch of the proof.* It is clear from the definition of the density  $\bar{g}^{t,x}$  that

$$g^{(\bar{\lambda}_t(x), \bar{\rho}_t(x), \bar{v}_t(x))}(w) = \bar{g}^{t,x}(w - \bar{v}_t(x)),$$

where we denote

$$\bar{v}_t(x) = \int_0^t v(\chi_s^t(x)) W_\alpha(t; s) ds.$$

On the other hand, one can show similarly to (A.36) that

$$|\bar{\chi}_s^t(x) - \chi_s(x)| \leq Ct^{1/\alpha}, \quad s \leq t, \quad \text{and} \quad \left| (\bar{\chi}_t^t(x) - t^{1/\alpha} \bar{v}_t(x)) - \chi_t(x) \right| \leq Ct^{1/\alpha+\delta}. \tag{3.26}$$

It follows from the Hölder continuity of  $\lambda, \rho, v$  and the first inequality in (3.26) that

$$|\bar{\lambda}_t(x) - \lambda_t(x)| + |\bar{\rho}_t(x) - \rho_t(x)| + |\bar{v}_t(x) - v_t(x)| \leq Ct^{\zeta/\alpha} \leq Ct^\delta.$$

Then the required bounds for

$$\begin{aligned} \bar{R}_t(x, y) &= R_t(x, y) + \frac{1}{t^{1/\alpha}} g^{(\lambda_t(x), \rho_t(x), v_t(x))} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) \\ &\quad - \frac{1}{t^{1/\alpha}} g^{(\lambda_t(x), \bar{\rho}_t(x), \bar{v}_t(x))} \left( \frac{y - \bar{\chi}_t^t(x) + t^{1/\alpha} \bar{v}_t(x)}{t^{1/\alpha}} \right) \end{aligned}$$

follow by respective bounds for  $R_t(x, y)$  and the basic properties of stable densities (e.g. (A.30)).  $\square$

In the above representations, we define the regressor as the solution to the ODE driven by the (mollified) partially compensated drift, and then determine the parameters of the  $\alpha$ -stable density of the innovation term by averaging of the correspondent space dependent parameters of the model w.r.t. the solution to the ODE on the time interval  $[0, t]$ . These principal components can be further simplified by the cost of making the bounds less precise and (possibly) under additional assumptions. First, let us mention briefly that the *true* solution  $\chi_t$  to (3.11) can be replaced by its  $k$ -th iteration  $\chi_t^{(k)}$  in the Picard approximation procedure. The situation here is similar to the one studied in [19, Section 2.2], thus we omit a detailed discussion and just mention that for such an approximation to be successful one needs

$$1 + \eta + \dots + \eta^k > \frac{1}{\alpha}.$$

In particular, the naive choice of the regressor  $f_t(x) = x + b(x)t$  mentioned in Introduction corresponds to the case  $k = 1$ . That is, for such a choice to be successful it is required that  $\alpha > (1 + \eta)^{-1}$ , which in particular excludes small values  $\alpha \leq 1/2$ .

Next, in the case of *bounded*  $\tilde{b}$ , the innovation term can be further simplified. Namely, in this case it is easy to verify that

$$|\chi_t(x) - x| \leq C(t + t^{1/\alpha})$$

if  $\alpha \neq 1$  (in the exceptional case  $\alpha = 1$  an additional logarithmic term should appear). Since  $\lambda$  is  $\zeta$ -Hölder continuous, this yields

$$|\lambda_t(x) - \lambda(x)| \leq C(t^\zeta + t^{\delta\zeta}),$$

and the similar bounds hold true for  $\rho_t, v_t, \bar{\lambda}_t, \bar{\rho}_t, \bar{v}_t$ . Then essentially the same argument as in the proof of (4.33) (see Appendix A.4) makes it possible to deduce representations

$$p_t(x, y) = \frac{1}{t^{1/\alpha}} g^x \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) + R_t^{frozen}(x, y) = \frac{1}{t^{1/\alpha}} \bar{g}^x \left( \frac{y - \bar{\chi}_t(x)}{t^{1/\alpha}} \right) + \bar{R}_t^{frozen}(x, y) \tag{3.27}$$

with the  $\alpha$ -stable densities

$$g^x(w) = g^{(\lambda(x), \rho(x), v(x))}(w), \quad \bar{g}^x(w) = g^{(\bar{\lambda}(x), \bar{\rho}(x), 0)}(w),$$

which just correspond to the values of the parameters ‘frozen’ at the initial point  $x$ . The error terms  $R_t^{frozen}(x, y), \bar{R}_t^{frozen}(x, y)$  under the corresponding conditions satisfy analogues of (3.16), (3.18), and (3.20) with  $\delta$  changed to  $\delta \wedge \zeta$ . Note that  $\delta < \zeta/\alpha$ ; that is, for  $\alpha \geq 1$  the bounds actually remain unchanged.

### 3.5 Some related results

We do not give a wide overview of the related results in this extensively developing domain, referring an interested reader to [10], [20], and a survey paper [13] for such reviews. Instead, we focus on a discussion of references directly related to the particular issues treated in the current paper.

1. *Various types of estimates.* We have already mentioned that the most attention in the available literature is devoted to kernel-type estimates, see detailed surveys in [10], [13]. The separate study of integral-in- $y$  and uniform-in- $(x, y)$  estimates is apparently

new; note however the forthcoming book [15], Sections 5.4, 5.5, where a systematic treatment is given, which leads to a pair of dual  $L_1$ - $C_\infty$  estimates. These estimates are of the same spirit with ours; however, one should note that the additive-in-space bounds (see [15, (5.69)]) adopted there as the main assumption, in certain settings, may become too restrictive. Namely, it will become clear from the proof of uniform-in- $(x, y)$  estimate in Section 5 below that the main property required for such estimate to hold is the integral-in- $x$  bound (6.4) which is actually a ‘dual’ analogue of the ‘direct’ integral-in- $y$  estimate. Example (3.7) shows that, for singular Lévy kernels, the ‘direct’ and the ‘dual’ estimates should be treated separately. On the other hand, the additive structure of [15, (5.69)] makes the integral-in- $x$  and the integral-in- $y$  estimates synonymic, which does not allow one to approach singular Lévy-type models. In the recent preprint [18], another (mixed  $L_1$ - $C_\infty$ ) type of estimates is proposed to treat the singular model Lévy-type based on the multidimensional SDEs with cylindrical  $\alpha$ -stable noise and non-trivial rotation.

Let us mention that the  $L_1$ -approach, based on integral-in- $y$  estimates only, has a deep connection, at least on the level of the principle ideas, with the approach to the well-posedness of the martingale problem for integro-differential operators which dates back to [7] and [16], [17].

2. *Non-symmetry of the Lévy noise.* The heat kernel estimates for Lévy and Lévy-type processes were mainly studied for symmetric noises; the non-symmetric setting becomes the subject of a study just in the few last years. The most advanced study in this direction available to the author is given by the recent preprint [23]; we refer there for an overview of few other recent results in the same direction. In the model from [23], the external drift (our  $b$ ) is not included, as well as the nuisance kernel  $\nu$ . On the other hand, the class of the kernels treated therein is substantially wider than our class of  $\alpha$ -stable principal parts.

3. *Non-boundedness of the drift coefficient.* It is traditional for the literature exploiting the analytical parametrix-type methods that the coefficients are assumed to be globally bounded. On the other hand, it was specially pointed to the author by H. Masuda that, for various applications esp. in statistics it is highly desirable for the theory to cover mean reverting models of the Ornstein-Uhlenbeck type. This explains the special attention paid in the paper to the case of unbounded  $b$ . The only reference known to the author, where such non-boundedness is allowed, is an apparently yet not published preprint [8].

## 4 Preliminaries to the proofs: the parametrix method and an integral representation for $p_t(x, y)$

In this section we make preparation for the proofs of the main results. We introduce an integral equation whose unique solution  $p_t(x, y)$  later on will be proved to be the transition probability density of the target process  $X$ . Such a construction is motivated by the *parametrix method*, which is a classical tool for constructing fundamental solutions to parabolic Cauchy problems. We present here only the rigorous step-by-step exposition without additional discussion of the heuristics behind the method; for such a discussion e.g. [10], [19].

### 4.1 The parametrix method: an outline, and the choice of the zero order approximation

In this section, we introduce the main objects and explain the method. We will repeatedly use the following notation for space- and time-space convolutions of functions:

$$(f * g)_t(x, y) := \int_{\mathbb{R}^d} f_t(x, z)g_t(z, y) dz, \quad (f \otimes g)_t(x, y) := \int_0^t \int_{\mathbb{R}^d} f_{t-s}(x, z)g_s(z, y) dz ds.$$

We will fix a function  $p_t^0(x, y)$ , a ‘zero order approximation’ to the unknown  $p_t(x, y)$ , which will belong to  $C^1(0, \infty)$  in  $t$  and to  $C_\infty^2$  in  $x$ . In particular, the following ‘differential error term’ will be well defined point-wisely:

$$\Phi_t(x, y) := -(\partial_t - L_x)p_t^0(x, y), \quad x, y \in \mathbb{R}, \tag{4.1}$$

here and below the lower index of an operator indicates the variable at which the operator is applied. Under the proper choice of  $p_t^0(x, y)$ , the kernel  $\Phi_t(x, y)$  will satisfy

$$\sup_{x \in \mathbb{R}} \int_{\mathbb{R}} |\Phi_t(x, y)| dy \leq Ct^{-1+\delta}. \tag{4.2}$$

The cornerstone of the construction is given by the 2nd type Fredholm integral equation

$$p_t(x, y) = p_t^0(x, y) + (p \otimes \Phi)_t(x, y), \tag{4.3}$$

which we interpret in the following way. With the time horizon  $T > 0$  being fixed, consider the Banach space of the kernels  $\Upsilon_t(x, y)$  on  $[0, T] \times \mathbb{R} \times \mathbb{R}$  with the norm

$$\|\Upsilon\|_{\infty,1,1} = \sup_{x \in \mathbb{R}} \int_0^T \int_{\mathbb{R}} |\Upsilon_t(x, y)| dy dt.$$

Consider also the Banach space  $L_{\infty,\infty,1}^T$  of functions  $f_t(x, y)$  with the norm

$$\|f\|_{\infty,\infty,1} = \sup_{x \in \mathbb{R}, t \in [0, T]} \int_{\mathbb{R}} |f_t(x, y)| dy.$$

Any kernel  $\Upsilon \in L_{\infty,1,1}^T$  generates a bounded linear operator in  $L_{\infty,\infty,1}^T$

$$(A^\Upsilon f)_t(x, y) = (f \otimes \Upsilon)_t(x, y),$$

with the operator norm of  $A^\Upsilon$  bounded by  $\|\Upsilon\|_{\infty,1,1}$ . By (4.2), the kernel  $\Phi_t(x, y)$  belongs to  $L_{\infty,1,1}^T$ . Then we naturally interpret (4.3) as an equation

$$p = p^0 + A^\Phi p \tag{4.4}$$

in the Banach space  $L_{\infty,\infty,1}^T$ . It is an easy calculation that (4.2) yields

$$\sup_x \int_{\mathbb{R}} |\Phi_t^{\otimes k}(x, y)| dy \leq t^{-1+k\delta} \frac{C^k \Gamma(\delta)^k}{\Gamma(k\delta)}, \quad \Phi^{\otimes k} = \underbrace{\Phi \otimes \dots \otimes \Phi}_k, \quad k \geq 1, \tag{4.5}$$

see Section 4.3 below. Then

$$\sum_{k=1}^{\infty} \|(A^\Phi)^k\| = \sum_{k=1}^{\infty} \|A^{\Phi^{\otimes k}}\| \leq \sum_{k=1}^{\infty} \|\Phi^{\otimes k}\|_{\infty,1,1} \leq \sum_{k=1}^{\infty} T^{k\delta} \frac{C^k \Gamma(\delta)^k}{\Gamma((k+1)\delta)} < \infty,$$

and therefore the solution to the equation (4.4) in  $L_{\infty,\infty,1}^T$  is uniquely specified by the classical von Neumann series representation:

$$p_t(x, y) = p_t^0(x, y) + \sum_{k \geq 1} (p^0 \otimes \Phi^{\otimes k})_t(x, y) = p_t^0(x, y) + (p^0 \otimes \Psi)_t(x, y), \tag{4.6}$$

$$\Psi_t(x, y) = \sum_{k \geq 1} \Phi_t^{\otimes k}(x, y), \tag{4.7}$$

with the series convergent in  $L_{\infty,\infty,1}^T$  and  $L_{\infty,1,1}^T$ , respectively.

Now, let us proceed with specification of the zero-order approximation  $p_t^0(x, y)$  for our particular model. We define the function  $\kappa_s(y), s \geq 0, y \in \mathbb{R}$  as the solution to the Cauchy problem

$$\frac{d}{ds} \kappa_s(y) = -B_s(\kappa_s(y)), \quad s \geq 0, \quad \kappa_0(y) = y, \quad y \in \mathbb{R}.$$

Define for  $z \in \mathbb{R}, t > 0$

$$\Psi_\alpha(t, z; \xi) = \int_0^t \int_{\mathbb{R}} \left( e^{iu\xi} - 1 - iu\xi 1_{|u| \leq s^{1/\alpha}} \right) \mu^{(\alpha)}(\kappa_s(z); du) ds,$$

which has representation in the form

$$\Psi_\alpha(t, z; \xi) = \Psi_\alpha(\tilde{\lambda}_t(z), \tilde{\rho}_t(z), \tilde{v}_t(z))(t^{1/\alpha} \xi) \tag{4.8}$$

with

$$\begin{aligned} \tilde{\lambda}_t(z) &= \frac{1}{t} \int_0^t \lambda(\kappa_s(z)) ds, & \tilde{\rho}_t(z) &= \frac{1}{t \tilde{\lambda}_t(z)} \int_0^t \lambda(\kappa_s(z)) \rho(\kappa_s(z)) ds, \\ \tilde{v}_t(z) &= \int_0^t v(\kappa_s(z)) W_\alpha(t; s) ds, \end{aligned}$$

recall that  $W_\alpha(t; s)$  is defined in (3.12). We will prove (4.8) in Appendix A.4; this identity actually means that  $\Psi_\alpha(t, z; \xi)$  is a characteristic exponent of an  $\alpha$ -stable law. We denote by  $h^{t,z}(w)$  the corresponding  $\alpha$ -stable distribution density

$$h^{t,z}(w) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iw\xi + \Psi_\alpha(t,z;\xi)} d\xi,$$

and define

$$p_t^0(x, y) = h^{t,y}(\kappa_t(y) - x). \tag{4.9}$$

Denote

$$\tilde{g}^{t,z}(w) = g^{(\tilde{\lambda}_t(z), \tilde{\rho}_t(z), \tilde{v}_t(z))}(w), \tag{4.10}$$

then by (4.8) the formula can be written as

$$p_t^0(x, y) = \frac{1}{t^{1/\alpha}} \tilde{g}^{t,y} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right). \tag{4.11}$$

#### 4.2 Kernel $\Phi_t(x, y)$ : decomposition and estimates

Define an auxiliary operator

$$\tilde{L}^{(\alpha),z,t} f(x) = \int_{\mathbb{R}} \left( f(x+u) - f(x) - uf'(x) 1_{|u| \leq t^{1/\alpha}} \right) \mu^{(\alpha)}(\kappa_t(z); du), \quad f \in C_\infty^2.$$

The following identity is crucial for the entire construction.

$$(\partial_t - \tilde{L}_x^{(\alpha),z,t}) h^{t,z}(w-x) = 0, \quad t > 0, \quad x, w, z \in \mathbb{R}. \tag{4.12}$$

This identity can be verified using the formula (4.9) and a standard Fourier analysis-based argument; see Appendix A.4. We have

$$\begin{aligned} \partial_t p_t^0(x, y) &= \partial_t h^{t,y}(\kappa_t(y) - x) = \partial_t h^{t,y}(w-x) \Big|_{w=\kappa_t(y)} + \partial_w h^{t,y}(w-x) \Big|_{w=\kappa_t(y)} \partial_t \kappa_t(y) \\ &= \partial_t h^{t,y}(w-x) \Big|_{w=\kappa_t(y)} - \partial_x h^{t,y}(w-x) \Big|_{w=\kappa_t(y)} \partial_t \kappa_t(y). \end{aligned}$$

Thus, combining (4.12) and the fact that  $\partial_t(\kappa_t(y)) = -B_t(\kappa_t(y))$ , we get

$$\begin{aligned} \partial_t p_t^0(x, y) &= \tilde{L}_x^{(\alpha),y,t} h^{t,y}(\kappa_t(y) - x) + \partial_x h^{t,y}(\kappa_t(y) - x) B_t(\kappa_t(y)) \\ &= \tilde{L}_x^{(\alpha),y,t} p_t^0(x, y) + B_t(\kappa_t(y)) \partial_x p_t^0(x, y). \end{aligned} \tag{4.13}$$

On the other hand, for the operator  $L$  defined by (2.4) we have the following decomposition:

$$L = b\partial_x + L^{(\alpha),x,1} + L^{\nu,x,1} = b_t\partial_x + L^{(\alpha),x,t} + L^{\nu,x,t}, \tag{4.14}$$

where

$$\begin{aligned} L^{(\alpha),z,t} f(x) &= \int_{\mathbb{R}} \left( f(x+u) - f(x) - uf'(x) 1_{|u| \leq t^{1/\alpha}} \right) \mu^{(\alpha)}(z; du), \quad f \in C_{\infty}^2, \\ L^{\nu,z,t} f(x) &= \int_{\mathbb{R}} \left( f(x+u) - f(x) - uf'(x) 1_{|u| \leq t^{1/\alpha}} \right) \nu(z; du), \quad f \in C_{\infty}^2. \end{aligned} \tag{4.15}$$

Now we can represent  $\Phi$  in the following form:

$$\begin{aligned} \Phi_t(x, y) &= (L_x - \partial_t) p_t^0(x, y) \\ &= \left( b_t(x) - B_t(\kappa_t(y)) \right) \partial_x p_t^0(x, y) + \left( L_x^{(\alpha),x,t} - \tilde{L}_x^{(\alpha),y,t} \right) p_t^0(x, y) + L_x^{\nu,x,t} p_t^0(x, y) \\ &=: \Phi_t^{drift}(x, y) + \Phi_t^{(\alpha)}(x, y) + \Phi_t^{\nu}(x, y). \end{aligned} \tag{4.16}$$

In what follows, we estimate separately the components of  $\Phi$  in the decomposition (4.16) and deduce an integral estimate for the entire  $\Phi$ , which holds true under  $\mathbf{H}^{\nu}$  (i). We will repeatedly use representation (4.11) and the following observation. The functions  $\tilde{\lambda}_t(z), \tilde{\rho}_t(z), \tilde{\nu}_t(z)$  are bounded since they are obtained by averaging of bounded functions w.r.t. probability measures. In addition,  $\tilde{\lambda}_t(z)$  is uniformly separated from zero. That is, for the function (4.10) with  $z = y$  the bounds (A.23), (A.30)–(A.33) can be used.

*Step 1: Estimate for  $\Phi^{drift}$ .* By (3.9), (3.10) we have

$$|b_t(x) - B_t(\kappa_t(y))| \leq |b_t(x) - B_t(x)| + |B_t(x) - B_t(\kappa_t(y))| \leq C \left( 1 + \left| \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right| \right) t^{-1+1/\alpha+\delta}.$$

We have

$$\partial_x p_t^0(x, y) = -\frac{1}{t^{2/\alpha}} \left( \tilde{g}^{t,y} \right)' \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right).$$

Applying (A.30), and then (A.18), (A.19), we easily get

$$|\Phi_t^{drift}(x, y)| \leq C t^{-1-1/\alpha+\delta} G^{(\alpha)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) = C t^{-1+\delta} G_t^{(\alpha,\alpha,\alpha)}(x, \kappa_t(y)). \tag{4.17}$$

*Step 2: Estimate for  $\Phi^{(\alpha)}$ .* Denote for  $f \in C_{\infty}^2$

$$L^{(\alpha),sym} f(x) = \int_{\mathbb{R}} \left( f(x+u) - f(x) - uf'(x) 1_{|u| \leq 1} \right) \frac{du}{|u|^{\alpha+1}}, \tag{4.18}$$

$$L^{(\alpha),asym} f(x) = \int_{\mathbb{R}} \left( f(x+u) - f(x) - uf'(x) 1_{|u| \leq 1} \right) \operatorname{sgn}(u) \frac{du}{|u|^{\alpha+1}}. \tag{4.19}$$

Then

$$\begin{aligned} L^{(\alpha),z,t} f \left( \frac{x}{t^{1/\alpha}} \right) &= \frac{\lambda(z)}{t} (L^{(\alpha),sym} f) \left( \frac{x}{t^{1/\alpha}} \right) + \frac{\lambda(z)\rho(z)}{t} (L^{(\alpha),asym} f) \left( \frac{x}{t^{1/\alpha}} \right), \\ \tilde{L}^{(\alpha),z,t} f \left( \frac{x}{t^{1/\alpha}} \right) &= \frac{\lambda(\kappa_t(z))}{t} (L^{(\alpha),sym} f) \left( \frac{x}{t^{1/\alpha}} \right) + \frac{\lambda(\kappa_t(z))\rho(\kappa_t(z))}{t} (L^{(\alpha),asym} f) \left( \frac{x}{t^{1/\alpha}} \right), \end{aligned}$$



and thus

$$\begin{aligned} \Phi_t^{(\alpha)}(x, y) &= \frac{\lambda(x) - \lambda(\kappa_t(y))}{t} (L^{(\alpha, sym)} \tilde{g}^{t, y}) \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) \\ &\quad + \frac{\lambda(x)\rho(x) - \lambda(\kappa_t(y))\rho(\kappa_t(y))}{t} (L^{(\alpha, asym)} \tilde{g}^{t, y}) \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right). \end{aligned} \tag{4.20}$$

On the other hand, we have by (A.32), (A.33)

$$|L^{(\alpha, sym)} \tilde{g}^{t, z}(x)| + |L^{(\alpha, asym)} \tilde{g}^{t, z}(x)| \leq CG^{(\alpha)}(x).$$

Since the functions  $\lambda(x)$  and  $\rho(x)$  are bounded and  $\zeta$ -Holder continuous, this gives

$$\begin{aligned} \Phi_t^{(\alpha)}(x, y) &\leq C(|x - \kappa_t(y)|^\zeta \wedge 1) t^{-1} G_t^{(\alpha, \alpha, \alpha)}(x, \kappa_t(y)) = Ct^{-1+\zeta/\alpha} G_t^{(\alpha, \alpha-\zeta, \alpha)}(x, \kappa_t(y)) \\ &= Ct^{-1+\delta_\zeta} G_t^{(\alpha, \alpha-\zeta, \alpha)}(x, \kappa_t(y)). \end{aligned}$$

*Step 3: Estimate for  $\Phi^\nu$ .* We decompose

$$\begin{aligned} \Phi_t^\nu(x, y) &= \int_{\mathbb{R}} \left( p_t^0(x+u, y) - p_t^0(x, y) - u \partial_x p_t^0(x+u, y) 1_{|u| \leq t^{1/\alpha}} \right) \nu(x; du) \\ &= \int_{|u| \leq t^{1/\alpha}} \left( p_t^0(x+u, y) - p_t^0(x, y) - u \partial_x p_t^0(x+u, y) \right) \nu(x; du) \\ &\quad + \int_{|u| > t^{1/\alpha}} p_t^0(x+u, y) \nu(x; du) - \int_{|u| > t^{1/\alpha}} p_t^0(x, y) \nu(x; du) \\ &=: \Phi_t^{\nu, small}(x, y) + \Phi_t^{\nu, large, +}(x, y) + \Phi_t^{\nu, large, -}(x, y). \end{aligned}$$

We have by (A.30)

$$|\partial_{xx}^2 p_t^0(x, y)| \leq Ct^{-3/\alpha} G^{(\alpha+2)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right),$$

which gives

$$\begin{aligned} |\Phi_t^{\nu, small}(x, y)| &\leq Ct^{-3/\alpha} \sup_{|v| \leq t^{1/\alpha}} G^{(\alpha+2)} \left( \frac{\kappa_t(y) - x - v}{t^{1/\alpha}} \right) \int_{|u| \leq t^{1/\alpha}} u^2 |\nu|(x; du) \\ &\leq Ct^{-3/\alpha} G^{(\alpha+2)} \left( \frac{\kappa_t(y) - x - v}{t^{1/\alpha}} \right) (t^{1/\alpha})^{2-\beta}, \end{aligned}$$

in the last inequality we used (A.20), condition (2.3), and (A.3). Next, we have by (A.23)

$$p_t^0(x, y) \leq Ct^{-1/\alpha} G^{(\alpha)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right),$$

thus by (2.3)

$$|\Phi_t^{\nu, large, -}(x, y)| \leq p_t^0(x, y) \int_{|u| > t^{1/\alpha}} |\nu|(x; du) \leq Ct^{-1/\alpha} G^{(\alpha)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) (t^{1/\alpha})^{-\beta}.$$

Then by (A.18)

$$\begin{aligned} |\Phi_t^{\nu, small}(x, y)| + |\Phi_t^{\nu, large, -}(x, y)| &\leq Ct^{-3/\alpha} G^{(\alpha+2)} \left( \frac{\kappa_t(y) - x - v}{t^{1/\alpha}} \right) (t^{1/\alpha})^{2-\beta} \\ &\quad + Ct^{-1/\alpha} G^{(\alpha)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) (t^{1/\alpha})^{-\beta} \\ &\leq Ct^{-1/\alpha-\beta/\alpha} G^{(\alpha)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) = Ct^{-\beta/\alpha} G_t^{(\alpha, \alpha, \alpha)}(x, \kappa_t(y)) \\ &= Ct^{-1+\delta_\beta} G_t^{(\alpha, \alpha, \alpha)}(x, \kappa_t(y)). \end{aligned}$$

That is, the first and the third parts in the above decomposition of  $\Phi^\nu$  satisfy a bound similar to the bound (4.17) for  $\Phi^{drift}$ . For the second part, we simply write

$$|\Phi_t^{\nu,large,+}(x,y)| \leq t^{-1+\delta\beta} Q_t(x,y),$$

where

$$Q_t(x,y) = t^{\beta/\alpha} |\Phi_t^{\nu,large,+}(x,y)| = t^{1-\delta\beta} |\Phi_t^{\nu,large,+}(x,y)| \tag{4.21}$$

is just a notation. This gives

$$|\Phi_t^\nu(x,y)| \leq Ct^{-1+\delta\beta} G_t^{(\alpha,\alpha,\alpha)}(x,\kappa_t(y)) + Ct^{-1+\delta\beta} Q_t(x,y).$$

*Summary: Proof of (4.2).* The above calculation gives

$$|\Phi_t(x,y)| \leq Ct^{-1+\delta} G_t^{(\alpha,\alpha,\alpha)}(x,\kappa_t(y)) + C^{-1+\delta\zeta} G_t^{(\alpha,\alpha-\zeta,\alpha)}(x,\kappa_t(y)) + Ct^{-1+\delta\beta} Q_t(x,y), \tag{4.22}$$

and thus

$$|\Phi_t(x,y)| \leq Ct^{-1+\delta} H_t(x,y), \quad H_t(x,y) = G_t^{(\alpha,\alpha-\zeta,\alpha)}(x,\kappa_t(y)) + Q_t(x,y). \tag{4.23}$$

We have for any  $\alpha, \beta, \gamma > 0$

$$\sup_x \int_{\mathbb{R}} G_t^{(\alpha,\beta,\gamma)}(x,\kappa_t(y)) dy \leq C, \tag{4.24}$$

see Appendix A.5. Since

$$\tilde{g}^{t,y}(w) \leq CG^{(\alpha)}(w), \quad w \in \mathbb{R}$$

by (A.23), we have then

$$p_t^0(x,y) \leq CG_t^{(\alpha,\alpha,\alpha)}(x+u,\kappa_t(y)). \tag{4.25}$$

Applying (4.24) with  $\alpha = \beta = \gamma$ , we get by (2.3)

$$\begin{aligned} \sup_x \int_{\mathbb{R}} Q_t(x,y) dy &\leq Ct^{-\beta/\alpha} \int_{|u|>t^{1/\alpha}} \left( \int_{\mathbb{R}} G_t^{(\alpha,\alpha,\alpha)}(x+u,\kappa_t(y)) dy \right) |\nu|(x;du) \\ &\leq Ct^{\beta/\alpha} \int_{|u|>t^{1/\alpha}} |\nu|(x;du) \leq C. \end{aligned} \tag{4.26}$$

Applying once again (4.24) with  $\gamma = \alpha, \beta = \alpha - \zeta$ , we get

$$\sup_x \int_{\mathbb{R}} H_t(x,y) dy \leq C, \tag{4.27}$$

which combined with (4.23) completes the proof of (4.2).

**Remark 4.1.** Using (3.3), we can also get

$$\sup_{x \in \mathbb{R}, t \in (0,T]} \int_{\{y:|y-x|>R\}} H_t(x,y) dy \rightarrow 0, \quad R \rightarrow \infty. \tag{4.28}$$

The proof is completely analogous and is omitted.

**4.3 Solution to (4.3): specification and further re-arrangement**

For any  $k > 1$  we have

$$\Phi_t^{\otimes k}(x, y) = \int_{0 < s_1 < \dots < s_{k-1} < t} \Phi_{s_1-s_0, \dots, s_k-s_{k-1}}(x, y) ds_1 \dots ds_{k-1},$$

where we denote  $s_0 = 0, s_k = t,$

$$\Phi_{\tau_1, \dots, \tau_k}(x, y) = \left( \Phi_{\tau_1} * \dots * \Phi_{\tau_k} \right)(x, y) = \int_{\mathbb{R}^{k-1}} \Phi_{\tau_1}(x, w_1) \dots \Phi_{\tau_k}(w_{k-1}, y) dw_1 \dots dw_{k-1}.$$

By (4.23),

$$\begin{aligned} \int_{\mathbb{R}} |\Phi_{\tau_1, \dots, \tau_k}(x, y)| dy &\leq \int_{\mathbb{R}^{k-1}} \int_{\mathbb{R}} |\Phi_{\tau_1}(x, w_1) \dots \Phi_{\tau_k}(w_{k-1}, y)| dw_1 \dots dw_{k-1} dy \\ &\leq C \tau_k^{-1+\delta} \int_{\mathbb{R}^{k-1}} |\Phi_{\tau_1}(x, w_1) \dots \Phi_{\tau_{k-1}}(w_{k-2}, w_{k-1})| dw_1 \dots dw_{k-1} \leq \dots \leq C^k \prod_{j=1}^k \tau_j^{-1+\delta}. \end{aligned}$$

Thus

$$\begin{aligned} \sup_x \int_{\mathbb{R}} |\Phi_t^{\otimes k}(x, y)| dy &\leq C^k \int_{0 < s_1 < \dots < s_{k-1} < t} \prod_{j=1}^k (s_j - s_{j-1})^{-1+\delta} ds_1 \dots ds_{k-1} \\ &= C^k t^{-1+k\delta} \int_{0 < v_1 < \dots < v_{k-1} < 1} \prod_{j=1}^k (v_j - v_{j-1})^{-1+\delta} dv_1 \dots dv_{k-1} = t^{-1+k\delta} \frac{C^k \Gamma(\delta)^k}{\Gamma(k\delta)}, \end{aligned}$$

which is just (4.5). That is, the solution  $p_t(x, y)$  to the integral equation (4.3) is uniquely defined by (4.6).

Note that the resolvent kernel  $\Psi_t(x, y)$  for the integral equation (4.3) inherits from  $\Phi_t(x, y)$  the integral bounds and the tail behavior. Namely, we have

$$\sup_x \int_{\mathbb{R}} |\Psi_t(x, y)| dy \leq \sum_{k=1}^{\infty} t^{-1+k\delta} \frac{C^k \Gamma(\delta)^k}{\Gamma(k\delta)} \leq C t^{-1+\delta}. \tag{4.29}$$

Next, by (4.28) we have

$$\sup_{x \in \mathbb{R}, t \in (0, T]} t^{1-\delta} \int_{\{y: |y-x| > R\}} |\Phi_t(x, y)| dy \rightarrow 0, \quad R \rightarrow \infty.$$

Then it is easy to show by induction that, for any  $k,$

$$\sup_{x \in \mathbb{R}, t \in (0, T]} t^{1-\delta} \int_{\{y: |y-x| > R\}} |\Phi_t^{\otimes k}(x, y)| dy \rightarrow 0, \quad R \rightarrow \infty.$$

These bounds combined with (4.5) yield the similar tail behavior of the kernel  $\Psi_t(x, y):$

$$\sup_{x \in \mathbb{R}, t \in (0, T]} t^{1-\delta} \int_{\{y: |y-x| > R\}} |\Psi_t(x, y)| dy \rightarrow 0, \quad R \rightarrow \infty. \tag{4.30}$$

The solution to (4.3) can be written as

$$p_t(x, y) = p_t^0(x, y) + r_t(x, y), \quad r_t(x, y) = (p^0 \otimes \Psi)_t(x, y), \tag{4.31}$$

and by (4.29), (4.25), and (4.24) we have

$$\begin{aligned} \sup_x \int_{\mathbb{R}} |r_t(x, y)| dy &\leq \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} p_{t-s}^0(x, z) |\Psi_s(z, y)| dz dy ds \\ &\leq C \int_0^t \left( \int_{\mathbb{R}} p_{t-s}^0(x, z) dz \right) s^{-1+\delta} ds \leq C t^\delta. \end{aligned} \tag{4.32}$$

Note that representation (4.31) differs from the one claimed in Theorem 3.4; in particular, zero order term  $p_t^0(x, y)$  in (4.31) is not equal to the principal term

$$p_t^{main}(x, y) = \frac{1}{t^{1/\alpha}} g^{t,x} \left( \frac{y - f_t(x)}{t^{1/\alpha}} \right)$$

in (1.1). The difference between these two terms admits the following bound; the proof is postponed to Appendix A.4:

$$|p_t^{main}(x, y) - p_t^0(x, y)| \leq Ct^{\delta\zeta} G_t^{(\alpha, \alpha-\zeta, \alpha)}(\chi_t(x), y) + Ct^\delta G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y). \quad (4.33)$$

We have

$$\sup_x \int_{\mathbb{R}} G_t^{(\alpha, \beta, \gamma)}(x, y) dy \leq C, \quad (4.34)$$

see Appendix A.5. That is, by (4.33)

$$\int_{\mathbb{R}} |p_t^{main}(x, y) - p_t^0(x, y)| dy \leq Ct^\delta. \quad (4.35)$$

Now it is easy to prove the following.

**Lemma 4.2.** For any  $f \in C_\infty$ ,

$$\sup_x \left| \int_{\mathbb{R}} p_t^{main}(x, y) f(y) dy - f(x) \right| \rightarrow 0, \quad \sup_x \left| \int_{\mathbb{R}} p_t^0(x, y) f(y) dy - f(x) \right| \rightarrow 0, \quad t \rightarrow 0,$$

*Proof.* We first note that there exists  $C > 1$  such that, for  $|x|$  large enough,

$$C^{-1}|x| \leq |\chi_t(x)| \leq C|x|,$$

see Proposition A.3. Since  $f(x) \rightarrow 0, |x| \rightarrow \infty$ , this gives

$$\sup_x |f(\chi_t(x)) - f(x)| \rightarrow 0, \quad t \rightarrow 0. \quad (4.36)$$

Next,  $g^{t,x}$  are stable densities with uniformly bounded intensities and shifts, and thus for every  $\varepsilon > 0$

$$\sup_x \int_{|w| > \varepsilon} \frac{1}{t^{1/\alpha}} g^{t,x} \left( \frac{w}{t^{1/\alpha}} \right) dw \rightarrow 0, \quad t \rightarrow 0.$$

Since  $f \in C_\infty$  is uniformly continuous, this yields

$$\sup_x \left| \int_{\mathbb{R}} \frac{1}{t^{1/\alpha}} g^{t,x} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) f(y) dy - f(\chi_t(x)) \right| \rightarrow 0, \quad t \rightarrow 0,$$

which proves the first assertion. The second assertion follows from the first one by (4.35).  $\square$

## 5 Proof of Theorem 3.3

We have defined the function  $p_t(x, y)$  as a solution to the integral equation (4.3). In this section we make a further analysis of its representation (4.6) and prove that function  $p_t(x, y)$ , in a certain approximate sense, provides a fundamental solution to the Cauchy problem for the operator  $\partial_t - L$ . This fact will be a cornerstone for the proof of Theorem 3.3.

**5.1 Continuity properties and approximate fundamental solution**

Denote

$$P_t f(x) = \int_{\mathbb{R}} f(y) p_t(x, y) dy, \quad t > 0, \quad P_0 f(x) = f(x).$$

**Lemma 5.1.** *For a given bounded measurable  $f$ , the function  $P_t f(x)$  is continuous w.r.t.  $(t, x) \in (0, \infty) \times \mathbb{R}$ .*

*For  $f \in C_\infty$ , one has  $P_t f \in C_\infty, t \geq 0$ , and  $P_t, t \geq 0$  is a continuous family of bounded linear operators in  $C_\infty$ .*

*Proof.* The proof is fairly standard, thus we just sketch it. We have

$$P_t f(x) = \int_{\mathbb{R}} f(y) p_t^0(x, y) dy + \int_0^t \int_{\mathbb{R}} p_{t-s}^0(x, y) \Psi_s^f(y) dy ds, \tag{5.1}$$

$$\Psi_t^f(y) = \int_{\mathbb{R}} \Psi_t(y, z) f(z) dz. \tag{5.2}$$

The function  $p_t^0(x, y)$ , given by an explicit formula (4.9), is continuous w.r.t.  $x, t$  for any  $y$ . Then one can deduce continuity of  $P_t f(x)$  using the bounds (4.25), (4.29) and a standard domination convergence argument; e.g. [10, Section 3.3]. Using (4.30), one can show in addition that

$$P_t f(x) \rightarrow 0, \quad |x| \rightarrow \infty \tag{5.3}$$

uniformly in  $t \in [0, T]$ . Combined with continuity of  $P_t f(x)$  in  $(t, x) \in (0, T] \times \mathbb{R}$  and Lemma 4.2, this yields continuity in  $t \in [0, T]$  of the family  $\{P_t f\} \in C_\infty$ . Clearly, each  $P_t$  is a linear operator; these operators are bounded thanks to (4.25), (4.29).  $\square$

**Lemma 5.2.** *For a given  $f \in C_\infty$ , the function  $\Psi_t^f(x)$  is continuous w.r.t.  $(t, x) \in (0, \infty) \times \mathbb{R}$ . In addition, for any  $0 < \tau < T$*

$$\Psi_t^f(x) \rightarrow 0, \quad |x| \rightarrow \infty \tag{5.4}$$

*uniformly on  $t \in [\tau, T]$ .*

*Proof.* The argument here is close to the one from the previous proof, with  $p_t^0$  changed to  $\Phi$ ; recall that  $\Psi_t(x, y)$  satisfies

$$\Psi_t(x, y) = \Phi_t(x, y) + \int_0^t \int_{\mathbb{R}} \Phi_{t-s}(x, y') \Psi_s(y', y) dy. \tag{5.5}$$

Therefore we omit the details, and only discuss two points which make the difference with the previous proof. First, the bound (4.23), when compared to (4.25), contains an extra term  $t^{-1+\delta}$ . This is the reason why (5.4) is stated for  $t \in [\tau, T]$  with *positive*  $\tau$ . Next, we yet have to verify that  $\Phi_t(x, y)$  is continuous in  $x, t$ . Recall the decomposition (4.16), and observe that the term  $\Phi^{(\alpha)}$  has the required continuity. However, two other terms in the decomposition (4.16) may fail to be continuous. Namely, since the function  $1_{|u|>t^{1/\alpha}}$  is discontinuous, weak continuity of the kernel  $\nu(x, du)$  does not imply, in general, continuity of the corresponding integral  $m_t^\nu(x)$ . This trouble is artificial, and can be fixed by a proper re-arrangement of the compensating terms in these two summands. Namely, we take function  $\theta \in C(\mathbb{R})$  with

$$\theta(u) = 0, \quad |u| \leq \frac{1}{2}, \quad \theta(u) = 1, \quad |u| \geq 1,$$

and put

$$\widehat{m}_t^\nu(x) = \int_{|u| \leq 1} u \theta(ut^{-1/\alpha}) \nu(x, du), \quad \widehat{b}_t = b - m_t^{(\alpha)} - \widehat{m}_t^\nu,$$

$$\widehat{\Phi}_t^{drift}(x, y) = \left( \widehat{b}_t(x) - B_t(\kappa_t(y)) \right) \partial_x p_t^0(x, y).$$

$$\widehat{\Phi}_t^\nu(x, y) = \int_{\mathbb{R}} \left( p_t^0(x + u, y) - p_t^0(x, y) - u(1 - \theta(ut^{-1/\alpha})) \partial_x p_t^0(x, y) 1_{|u| \leq 1} \right) \nu(x; du).$$

Then

$$\widehat{\Phi}_t^{drift}(x, y) + \widehat{\Phi}_t^\nu(x, y) = \Phi_t^{drift}(x, y) + \Phi_t^\nu(x, y),$$

and the terms  $\widehat{\Phi}_t^{drift}$  and  $\widehat{\Phi}_t^\nu$  have the required continuity. The latter can be verified via a routine calculation involving the continuity condition  $\mathbf{H}^{cont}$ , we omit a detailed discussion.  $\square$

The parametrix construction described in Section 4.1 originates in the general interpretation of  $p_t(x, y)$  as a (sort of) fundamental solution to the Cauchy problem for the operator  $\partial_t - L$ ; that is, in other words,  $p_t(x, y)$  should satisfy the backward Kolmogorov equation for the (yet unknown) process  $X$ . In some cases one can show that  $p_t(x, y)$  indeed satisfies

$$(\partial_t - L_x)p_t(x, y) = 0 \tag{5.6}$$

in a classical way; for instance, this is the mainstream approach in the classical diffusive/parabolic setting, see [5]. A necessary pre-requisite for such an approach is to prove that  $p_t(x, y)$  belongs to  $C^1$  w.r.t.  $t$  and to  $C_\infty^2$  (which is just the domain of  $L$ ) w.r.t.  $x$ . In the current setting, zero order approximation  $p_t^0(x, y)$  has the required smoothness properties, however one can hardly extend these properties to  $p_t(x, y)$  using (5.1) in the way used in the proof of Lemma 5.1. The main obstacle is that  $\partial_x p_t^0(x, y), \partial_{xx}^2 p_t^0(x, y)$  exhibit strongly singular behavior as  $t \rightarrow 0$  (see (A.30), (A.31)), which does not allow one to differentiate (5.1). This observation leads to the following auxiliary construction. Define for  $\varepsilon > 0$

$$p_{t,\varepsilon}(x, y) = p_{t+\varepsilon}^0(x, y) + \int_0^t \int_{\mathbb{R}} p_{t-s+\varepsilon}^0(x, y') \Psi_s(y', y) dy' ds, \tag{5.7}$$

$$P_{t,\varepsilon} f(x) = \int_{\mathbb{R}^d} p_{t,\varepsilon}(x, y) f(y) dy, \quad f \in C_\infty. \tag{5.8}$$

The following lemma shows that  $p_{t,\varepsilon}(x, y)$  approximates  $p_t(x, y)$  and satisfies an approximative analogue of (5.6). This is our reason to call the family  $\{p_{t,\varepsilon}(x, y), \varepsilon > 0\}$  an *approximate fundamental solution*.

**Lemma 5.3.** *For every  $f \in C_\infty$  we have the following.*

1. For every  $T > 0$ ,

$$\|P_{t,\varepsilon} f - P_t f\|_\infty \rightarrow 0, \quad \varepsilon \rightarrow 0, \tag{5.9}$$

uniformly in  $t \in (0, T]$ , and

$$P_{t,\varepsilon} f(x) \rightarrow 0, \quad |x| \rightarrow \infty \tag{5.10}$$

uniformly in  $t \in (0, T], \varepsilon \in (0, 1]$ .

- 2.

$$\lim_{t,\varepsilon \rightarrow 0^+} \|P_{t,\varepsilon} f - f\|_\infty = 0. \tag{5.11}$$

3. For every  $\varepsilon > 0$ ,  $P_{t,\varepsilon} f(x)$  belongs to  $C^1$  as a function of  $t$ , to  $C_\infty^2$  as a function of  $x$ , and  $\partial_t P_{t,\varepsilon} f(x), L_x P_{t,\varepsilon} f(x)$  are continuous w.r.t.  $(t, x)$ .

4. For every  $0 < \tau < T, R > 0$

$$Q_{t,\varepsilon}f(x) = (\partial_t - L_x)P_{t,\varepsilon}f(x) \rightarrow 0, \quad \varepsilon \rightarrow 0, \tag{5.12}$$

uniformly in  $(t, x) \in [\tau, T] \times [-R, R]$ . In addition,

$$\int_0^T \sup_{x \in [-R, R]} |Q_{t,\varepsilon}f(x)| dt \rightarrow 0, \quad \varepsilon \rightarrow 0. \tag{5.13}$$

*Proof.* Statements 1–3 follow easily by the same continuity/domination argument which was used in Lemma 5.1 and thus we omit the proof; see [10, Section 4.1] for a detailed exposition of similar group of statements.

To prove statement 4, we apply the argument from the proof of [10, Lemma 5.2]. Since the additional time shift by  $\varepsilon > 0$  removes the singularity at the point  $t = 0$  in (5.7), the continuity/domination argument similar to the one used in Lemma 5.1 allows one to interchange the operator  $(\partial_t - L_x)$  with the integrals in the definition of  $P_{t,\varepsilon}f$ . Then, recalling the definition (4.1) of  $\Phi_t(x, y)$  and (5.2), we get

$$Q_{t,\varepsilon}f(x) = - \int_{\mathbb{R}} \Phi_{t+\varepsilon}(x, y)f(y) dy - \int_0^t \int_{\mathbb{R}} \Phi_{t-s+\varepsilon}(x, y)\Psi_s^f(y) dy ds + \int_{\mathbb{R}} p_\varepsilon^0(x, y)\Psi_t^f(y) dy,$$

see [10, (4.13)] By the continuity of  $\Phi_t(x, y)$  in  $t$ , we have

$$\begin{aligned} \int_{\mathbb{R}} \Phi_{t+\varepsilon}(x, y)f(y) dy + \int_0^t \int_{\mathbb{R}} \Phi_{t-s}(x, y)\Psi_s^f(y) dy ds &\rightarrow \int_{\mathbb{R}} \Phi_t(x, y)f(y) dy \\ &+ \int_0^t \int_{\mathbb{R}} \Phi_{t-s+\varepsilon}(x, y)\Psi_s^f(y) dy ds, \quad \varepsilon \rightarrow 0 \end{aligned}$$

uniformly in  $x \in [-R, R], t \in [\tau, T]$ . On the other hand, since  $\Psi_t^f(x)$  is continuous, we have by Lemma 4.2

$$\int_{\mathbb{R}} p_\varepsilon^0(x, y)\Psi_t^f(y) dy \rightarrow \Psi_t^f(x), \quad \varepsilon \rightarrow 0$$

uniformly in  $x \in [-R, R], t \in [\tau, T]$ , which combined with (5.5) completes the proof of (5.12). On the other hand it follows from (4.25) and (4.29) that

$$\int_0^\tau \sup_x |Q_{t,\varepsilon}f(x)| dt \leq C\tau^\delta \sup_y |f(y)|.$$

Combined with (5.12), this yields (5.13). □

**Definition 5.4.** We say a continuous function  $h(t, x)$  to be approximate harmonic for an operator  $\partial_t - L$ , if there exists a family  $\{h_\varepsilon(t, x), \varepsilon \in (0, 1]\} \in C([0, \infty) \times \mathbb{R})$  such that

(i) for any  $T > 0$

$$\sup_{x \in \mathbb{R}, t \in [0, T]} |h_\varepsilon(t, x) - h(t, x)| \rightarrow 0, \quad \varepsilon \rightarrow 0, \quad \sup_{t \in [0, T], \varepsilon \in (0, 1]} |h_\varepsilon(t, x)| \rightarrow 0, \quad |x| \rightarrow \infty;$$

(ii) each function  $h_\varepsilon(t, x)$  is  $C^1$  w.r.t.  $t, C_\infty^2$  w.r.t.  $x$ , and for every  $0 < \tau <, R > 0$

$$\sup_{x \in [-R, R], t \in [\tau, T]} |(\partial_t - L_x)h_\varepsilon(t, x)| \rightarrow 0, \quad \varepsilon \rightarrow 0.$$

Note that, by Lemma 5.3, for any  $f \in C_\infty$  the function  $h^f(t, x) = P_t f(x)$  is approximate harmonic for  $\partial_t - L$ . The corresponding approximating family is given by

$$h_\varepsilon^f(t, x) = P_{t,\varepsilon}f(x), \quad \varepsilon > 0. \tag{5.14}$$

### 5.2 The Positive Maximum Principle and the semigroup properties

In this section we establish the semigroup properties for the family of the operators  $\{P_t, t \geq 0\}$ . A classical method for this is based on the *Positive Maximum Principle* (PMP) for the operator  $L$ . It is usually applied when  $p_t(x, y)$  is a (true) fundamental solution for  $\partial_t - L$ ; e.g. [12]. In our setting  $p_t(x, y)$  satisfies (5.6) in a weaker approximate sense; however, the classical PMP-based argument admits an extension which is well applicable in such an approximate setting. This extended argument is essentially due to [10, Section 4]. For the reader's and further reference convenience, here we give a systematic version of this argument, based on the notion of approximate harmonic functions.

Recall that an operator  $L$  with a domain  $\mathcal{D}$  is said to satisfy PMP if for any  $f \in \mathcal{D}$  and  $x_0$  such that

$$0 \leq f(x_0) = \max_x f(x)$$

one has

$$Lf(x_0) \leq 0.$$

Clearly, the operator (2.4) with the domain  $\mathcal{D} = C_\infty^2$  satisfies PMP; note that  $Lf$  is continuous for any  $f \in C_\infty^2$ , but does not necessarily belong to  $C_\infty$ .

**Proposition 5.5.** *Let  $h(t, x)$  be an approximate harmonic function for  $\partial_t - L$  such that  $h(0, \cdot) \geq 0$ .*

*Then  $h(t, \cdot) \geq 0, t > 0$ .*

*Proof.* Assuming  $h(t, x)$  being negative at some point, we have that for some  $T > 0$

$$\inf_{t \leq T, x \in \mathbb{R}} h(t, x) < 0. \tag{5.15}$$

Let  $\{h_\varepsilon(t, x), \varepsilon \in (0, 1]\}$  be the approximating family from Definition 5.4, then by assertion (i) there exist  $\nu > 0, \theta > 0, \varepsilon_1 > 0$  such that

$$\inf_{t \leq T, x \in \mathbb{R}} (h_\varepsilon(t, x) + \theta t) < -\nu, \quad \varepsilon < \varepsilon_1.$$

Denote

$$u_\varepsilon(t, x) = h_\varepsilon(t, x) + \theta t, \quad \varepsilon \in (0, 1]$$

these functions are continuous in  $(t, x)$  (because each  $h_\varepsilon$  is continuous) and satisfy

$$u_\varepsilon(t, x) \rightarrow \theta t > 0, \quad |x| \rightarrow \infty$$

uniformly in  $t \in [0, T]$  (because of the assertion (i)). Then for some  $R > 0$  and  $\varepsilon < \varepsilon_1$

$$\inf_{t \leq T, x \in \mathbb{R}} u_\varepsilon(t, x) = \inf_{t \leq T, x \in \mathbb{R}} (h_\varepsilon(t, x) + \theta t) < 0$$

is actually attained at some point in  $[0, T] \times [-R, R]$ ; we fix one such a point for each  $\varepsilon$ , and denote it by  $(t_\varepsilon, x_\varepsilon)$ . We observe that  $t_\varepsilon$  is separated from 0 when  $\varepsilon$  is small enough. Indeed, by the assertion (i) and non-negativity assumption  $h(0, x) \geq 0$ , there exist  $\varepsilon_0 > 0, \tau > 0$  such that

$$u_\varepsilon(t, x) \geq h_\varepsilon(t, x) \geq h_\varepsilon(0, x) - \frac{\nu}{2} \geq -\frac{\nu}{2}, \quad t \leq \tau, \quad \varepsilon < \varepsilon_0, \quad x \in \mathbb{R}.$$

Since

$$u_\varepsilon(t_\varepsilon, x_\varepsilon) = \min_{t \in [0, T], x \in \mathbb{R}} u_\varepsilon(t, x) < -\nu,$$

this yields  $t_\varepsilon > \tau$  for  $\varepsilon < \varepsilon_0$ .



Now we can conclude the proof in a quite standard way. Let  $\varepsilon < \varepsilon_0 \wedge \varepsilon_1$ . Since  $x_\varepsilon$  is the maximal point for  $-u_\varepsilon(t_\varepsilon, \cdot)$  and  $-u_\varepsilon(t_\varepsilon, x_\varepsilon) > 0$ , we have by the PMP

$$L_x u_\varepsilon(t_\varepsilon, x_\varepsilon) \geq 0.$$

Since  $t_\varepsilon$  is the maximal point for  $u_\varepsilon(\cdot, x_\varepsilon)$  and  $t_\varepsilon > \tau$ , we have

$$\partial_t u_\varepsilon(t_\varepsilon, x_\varepsilon) \leq 0,$$

where the sign ' $<$ ' may appear only if  $t_\varepsilon = T$ . Then

$$(\partial_t - L_x)u_\varepsilon(t, x)|_{(t,x)=(t_\varepsilon,x_\varepsilon)} \leq 0. \tag{5.16}$$

On the other hand, we by the assertion (ii) from Definition 5.4

$$(\partial_t - L_x)u_\varepsilon(t, x)|_{(t,x)=(t_\varepsilon,x_\varepsilon)} \geq \theta + \inf_{x \in [-R,R], t \in [\tau,T]} (\partial_t - L_x)h_\varepsilon(t, x) \rightarrow \theta > 0, \quad \varepsilon \rightarrow 0.$$

This gives contradiction and shows that (5.15) fails.  $\square$

Now the semigroup properties for the family  $\{P_t, t \geq 0\}$  can be derived in a standard way.

**Corollary 5.6.** 1. Each operator  $P_t, t \geq 0$  is positivity preserving: for any  $f \geq 0$  one has  $P_t f \geq 0$ .

2. The family  $\{P_t\}$  is a semigroup:

$$P_{t+s}f = P_t P_s f, \quad f \in C_\infty, \quad s, t \geq 0. \tag{5.17}$$

3. For any  $f \in C_0^2(\mathbb{R})$ ,

$$P_t f(x) - f(x) = \int_0^t P_s Lf(x) ds, \quad t \geq 0. \tag{5.18}$$

*Proof.* Statement 1 follows from Proposition 5.5 applied to  $h(t, x) = h^f(t, x)$ , which is already known to be approximate harmonic. To prove statement 2, we fix  $s \geq 0, f \in C_\infty$  and apply Proposition 5.5 to functions

$$h^\pm(t, x) = \pm P_{t+s}f(x) \mp P_t P_s f(x) = \pm h^f(t, x) \mp h^{P_s f}(t, x),$$

which are approximate harmonic and satisfy  $h^\pm(0, \cdot) = 0$ . Finally, to prove statement 3 we apply Proposition 5.5 to the function

$$h(t, x) = P_t f(x) - f(x) - \int_0^t P_s Lf(x) ds,$$

with the approximating family defined by

$$h_\varepsilon(t, x) = P_{t,\varepsilon} f(x) - f(x) - \int_0^t P_{s,\varepsilon} Lf(x) ds.$$

Note that  $h_\varepsilon(t, x)$  satisfies assertion (i) from Definition 5.4 by Lemma 5.3, and

$$\begin{aligned} (\partial_t - L_x)h_\varepsilon(t, x) &= Q_{t,\varepsilon} f(x) - (0 - Lf(x)) - \left( P_{t,\varepsilon} Lf(x) - \int_0^t L(P_{s,\varepsilon} Lf)(x) ds \right) \\ &= Q_{t,\varepsilon} f(x) - \int_0^t \partial_s (P_{s,\varepsilon} Lf)(x) + \int_0^t L(P_{s,\varepsilon} Lf)(x) ds \\ &= Q_{t,\varepsilon} f(x) - \int_0^t Q_{s,\varepsilon} Lf(x) ds. \end{aligned}$$

Applying (5.12) and (5.13), we get assertion (ii) from Definition 5.4.  $\square$

It is easy to deduce from (5.18) that

$$\int_{\mathbb{R}} p_t(x, y) dy = 1, \quad t > 0, \quad x \in \mathbb{R}.$$

Indeed, take  $f \in C_{\infty}^0$  such that  $f(x) = 1, |x| \leq 1$ , and put  $f_k(x) = f(k^{-1}x)$ . Then

$$f_k(x) \rightarrow 1, \quad Lf_k(x) \rightarrow 0, \quad k \rightarrow \infty$$

for every  $x$ , and  $\|Lf_k\| \leq C$ . Using (4.31), (4.25), and (4.32) we can apply the dominated convergence theorem and prove

$$\int_0^t P_s Lf_k(x) ds \rightarrow 0, \quad k \rightarrow \infty,$$

which combined with (5.18) gives the required identity.

Summarizing all the above, we conclude that  $P_t, t \geq 0$  is a *strongly continuous semigroup* in  $C_{\infty}$ , which is *positivity preserving* and *conservative*; that is, this semigroup is *Feller*. It follows from (5.18) that  $C_0^2$  belongs to the domain of its generator, and the restriction of this generator to  $C_0^2$  equals  $L$ . For any probability measure  $\pi$  on  $\mathbb{R}$  there exists a Markov process  $\{X_t\}$  with the transition semigroup  $\{P_t\}$ , càdlàg trajectories, and the initial distribution  $\text{Law}(X_0) = \pi$ ; see [4, Theorem 4.2.7]. Finally, by Lemma 5.1 the process  $X$  is strong Feller.

### 5.3 The martingale problem: uniqueness

Note that any Feller process  $Y$ , whose generator  $A$  restricted to  $C_0^2$  coincides with  $L$ , is a  $D(\mathbb{R}^+)$ -solution to the martingale problem  $(L, C_0^2)$ ; this is essentially the *Dynkin formula* combined with [4, Theorem 4.2.7]. In particular, this is the case for the Markov process  $X$ , constructed in the previous section. In this section, we prove that the  $D(\mathbb{R}^+)$ -solution to the martingale problem  $(L, C_0^2)$  with a given initial distribution  $\pi$  is unique; this will complete the proof of Theorem 3.3. The argument here is principally the same as in [19], with the one important addition which appears because the drift term now is not necessarily bounded.

By [4, Corollary 4.4.3], the required uniqueness holds true if for any two  $D(\mathbb{R}^+)$ -solutions to  $(L, C_0^2)$  with the same initial distribution  $\pi$  corresponding one-dimensional distributions coincide. In what follows, we fix some solution  $Y$  and prove that

$$\mathbf{E}f(Y_T) = \int_{\mathbb{R}} P_T f(x) \pi(dx), \quad f \in C_{\infty}, \quad T > 0. \tag{5.19}$$

It is easy to prove that  $Y_t, t \geq 0$  is stochastically continuous; see [10]. Then for any function  $h(t, x)$  which is differentiable w.r.t.  $t$ , belongs to  $C_0^2$  w.r.t.  $x$ , and has continuous and bounded  $\partial_t h(t, x), L_x h(t, x)$ , the process

$$h(t, Y_t) - \int_0^t \left( \partial_s h(s, Y_s) + L_x h(s, Y_s) \right) ds$$

is a martingale, see [4, Lemma 4.3.4 (a)]. We use this fact for a certain family of functions which approximate

$$h^{T,f}(t, x) = P_{T-t} f(x), \quad t \in [0, T], \quad x \in \mathbb{R};$$

here and below  $f \in C_{\infty}, T > 0$  are fixed. Consider a family of functions  $\{\varphi_R, R > 0\} \subset C^2$  such that  $\|\varphi_R\|_{C^2} \leq C$  and

$$\varphi_R(x) = \begin{cases} 1, & |x| \leq R + 1; \\ 0, & |x| \geq R + 2. \end{cases}$$

Define

$$h_{R,\varepsilon}^{T,f}(t, x) = \varphi_R(x)P_{T-t,\varepsilon}f(x), \quad R > 0, \quad \varepsilon > 0.$$

Recall that  $P_{T-t,\varepsilon}f(x) \in C^2$ , and is bounded together with its derivatives uniformly for  $t \in [0, T_1], |x| \leq R$  for any  $T_1 < T, R > 0$ . Multiplying this function by  $\varphi_R$ , we get a function from the class  $C_0^2$ . That is, we have that

$$M_{R,\varepsilon}^{T,f}(t) = h_{R,\varepsilon}^{T,f}(t, Y_t) - \int_0^t \left( \partial_s h_{R,\varepsilon}^{T,f}(s, Y_s) + L_x h_{R,\varepsilon}^{T,f}(s, Y_s) \right) ds, \quad t \in [0, T_1]$$

is a martingale. Denote  $h_\varepsilon^{T,f}(t, x) = P_{T-t,\varepsilon}f(x)$ . It is clear that

$$\partial_t h_{R,\varepsilon}^{T,f}(t, x) = \varphi_R(x)h_\varepsilon^{T,f}(t, x).$$

In addition, we have

$$L_x h_{R,\varepsilon}^{T,f}(t, x) = \varphi_R(x)L_x h_\varepsilon^{T,f}(t, x) + \int_{|u| \geq 1} \left( \varphi_R(x+u) - \varphi_R(x) \right) h_\varepsilon^{T,f}(t, x+u) \mu(x, du), \quad |x| \leq R.$$

Thus for  $|x| \leq R$  we can write

$$\partial_t h_{R,\varepsilon}^{T,f}(t, x) + L_x h_{R,\varepsilon}^{T,f}(t, x) = -\varphi_R(x)Q_{T-s,\varepsilon}f(x) + \Theta_{R,\varepsilon}^{T,f}(t, x),$$

where  $Q_{t,\varepsilon}f$  is defined in Lemma 5.3, and

$$\Theta_{R,\varepsilon}^{T,f}(t, x) = \int_{|u| \geq 1} \left( \varphi_R(x+u) - \varphi_R(x) \right) h_\varepsilon^{T,f}(t, x+u) \mu(x, du)$$

Observe that, for  $|x| \leq R$ ,

$$\left( \varphi_R(x+u) - \varphi_R(x) \right) \neq 0 \Rightarrow |x+u| \geq R+1,$$

which yields

$$|\Theta_{R,\varepsilon}^{T,f}(t, x)| \leq C \sup_{t \in [0, T], |y| \geq R+1} |h_\varepsilon^{T,f}(t, y)| =: F_{R,\varepsilon}^{T,f}.$$

Now we can finalize the proof. Without loss of generality, we assume that the initial distribution  $\pi$  has a compact support, and take  $R$  large enough, so that  $\text{supp } \pi \subset (-R, R)$ . Denote

$$\tau^R = \inf\{t : |Y_t| \geq R\} > 0,$$

then for any  $T_1 < T$  we have

$$\begin{aligned} & |\mathbf{E}h_{R,\varepsilon}^{T,f}(T_1 \wedge \tau_R, Y_{T_1 \wedge \tau_R}) - \mathbf{E}h_{R,\varepsilon}^{T,f}(0, Y_0)| \\ & \leq \mathbf{E} \int_0^{T_1 \wedge \tau_R} |Q_{T-s,\varepsilon}f(Y_s)| ds + TF_{R,\varepsilon}^{T,f}. \end{aligned}$$

Using Lemma 5.3, we pass to the limit as  $\varepsilon \rightarrow 0$  and get

$$|\mathbf{E}h^{T,f}(T_1 \wedge \tau_R, Y_{T_1 \wedge \tau_R}) - \mathbf{E}h^{T,f}(0, Y_0)| \leq CT \sup_{t \in [0, T], |y| \geq R+1} |h^{T,f}(t, y)|.$$

Taking  $R \rightarrow \infty$  and using Lemma 5.1, we get by the domination convergence theorem

$$\mathbf{E}h^{T,f}(T_1, Y_{T_1}) = \mathbf{E}h^{T,f}(0, Y_0), \quad T_1 < T.$$

Taking  $T_1 \rightarrow T$  and using the domination convergence theorem again, we get

$$\mathbf{E}f(Y_T) = \mathbf{E}P_T f(Y_0) = \int_{\mathbb{R}} P_T f(x) \pi(dx),$$

which proves (5.19).

### 6 Proof of Theorem 3.4

Statement I follows straightforwardly from (4.32) and (4.33). To prove statement II, we further re-arrange decomposition (4.16). Namely, we write

$$\Phi_t(x, y) = \Phi_t^{kernel}(x, y) + \Phi_t^{integral}(x, y), \tag{6.1}$$

where

$$\Phi_t^{integral}(x, y) = \Phi_t^{\nu, large, +} = \int_{|u| > t^{1/\alpha}} p_t^0(x + u, y) \nu(x; du),$$

and the term  $\Phi_t^{kernel}(x, y) = \Phi_t(x, y) - \Phi_t^{integral}(x, y)$  admits a point-wise ('kernel') bound

$$|\Phi_t^{kernel}(x, y)| \leq Ct^{-1+\delta} G_t^{(\alpha, \alpha-\zeta, \alpha)}(x, \kappa_t(y)).$$

Since the kernel  $G_t^{(\alpha, \alpha-\zeta, \alpha)}(u, v)$  is bounded by  $Ct^{-1/\alpha}$ , satisfies (4.34), and is symmetric, one has

$$|\Phi_t^{kernel}(x, y)| \leq Ct^{-1/\alpha-1+\delta}, \quad \sup_y \int_{\mathbb{R}} |\Phi_t^{kernel}(x, y)| dx \leq Ct^{-1+\delta}. \tag{6.2}$$

Next, it is straightforward to see that  $\Phi_t^{integral}(x, y)$  satisfies the similar sup-bound: since  $p_t^0(x, y)$  is bounded by  $Ct^{-1/\alpha}$ , we have by (2.3), (3.3)

$$|\Phi_t^{integral}(x, y)| \leq Ct^{-1/\alpha} \int_{|u| > t^{1/\alpha}} |\nu|(x; du) \leq Ct^{-1/\alpha-1+\delta_\beta}. \tag{6.3}$$

To obtain an integral bound for  $\Phi_t^{integral}(x, y)$ , we recall that

$$p_t^0(x + u, y) \leq \frac{C}{t^{1/\alpha}} G^{(\alpha)}\left(\frac{\kappa_t(y) - x - u}{t^{1/\alpha}}\right),$$

and observe that

$$G^{(\alpha)}(x) \leq C \left( \mathbb{1}_{[-1,1]} + G^{(\alpha)} * \mathbb{1}_{[-1,1]} \right) (x).$$

Then by (3.17)

$$\begin{aligned} \int_{\mathbb{R}} |\Phi_t^{integral}(x, y)| dx &\leq Ct^{-1/\alpha} \left| \int_{\mathbb{R}} \int_{|u| > t^{1/\alpha}} G^{(\alpha)}\left(\frac{w - x - u}{t^{1/\alpha}}\right) \nu(x; du) dx \right| \\ &\leq Ct^{-1/\alpha} \left| \int_{\mathbb{R}} \int_{|u| > t^{1/\alpha}} \mathbb{1}_{[-1,1]}\left(\frac{w - x - u}{t^{1/\alpha}}\right) \nu(x; du) dx \right| \\ &+ Ct^{-1/\alpha} \left| \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \int_{|u| > t^{1/\alpha}} \mathbb{1}_{[-1,1]}\left(\frac{w - x - u}{t^{1/\alpha}} - z\right) G^{(\alpha)}(z) \nu(x; du) dx \right) dz \right| \leq Ct^{-1+\delta_\nu}. \end{aligned}$$

Combined with (6.2), (6.3), this yields

$$|\Phi_t(x, y)| \leq Ct^{-1/\alpha-1+\delta}, \quad \sup_y \int_{\mathbb{R}} |\Phi_t(x, y)| dx \leq Ct^{-1+\delta_\infty}. \tag{6.4}$$

These bounds can be extended to the kernel  $\Psi = \sum_{k \geq 1} \Phi^{\otimes k}$ :

$$|\Psi_t(x, y)| \leq Ct^{-1/\alpha-1+\delta}, \quad \sup_y \int_{\mathbb{R}} |\Psi_t(x, y)| dx \leq Ct^{-1+\delta_\infty}. \tag{6.5}$$

The second bound follows from the second bound in (6.4) literally in the same way with (4.29). To get the first bound, we slightly modify the argument from Section 4.3. In what follows we use the notation of this section. Let  $k \geq 1, \tau_1, \dots, \tau_k \in [0, T]$  be given, and

let  $j \in \{1, \dots, k\}$  be such that  $\tau_j = \max_{i=1, \dots, k} \tau_i$ . Using the first inequality in (6.4) with  $t = \tau_j$ , we get

$$\begin{aligned} |\Phi_{\tau_1, \dots, \tau_k}(x, y)| &\leq \int_{\mathbb{R}^{k-1}} \int_{\mathbb{R}} |\Phi_{\tau_1}(x, w_1) \dots \Phi_{\tau_k}(w_{k-1}, y)| dw_1 \dots dw_{k-1} \\ &\leq C\tau_j^{-1-1/\alpha+\delta} \int_{\mathbb{R}^{j-1}} |\Phi_{\tau_1}(x, w_1) \dots \Phi_{\tau_{j-1}}(w_{j-2}, w_{j-1})| dw_1 \dots dw_{j-1} \\ &\quad \times \int_{\mathbb{R}^{k-j-1}} |\Phi_{\tau_{j+1}}(w_j, w_{j+1}) \dots \Phi_{\tau_{k-1}}(w_{k-2}, w_{k-1})| dw_{j+1} \dots dw_{k-1}. \end{aligned}$$

Then, using repeatedly (4.2) and the second inequality in (6.4) we get

$$\begin{aligned} |\Phi_{\tau_1, \dots, \tau_k}(x, y)| &\leq \\ &\leq C\tau_j^{-1-1/\alpha+\delta} (C\tau_{j-1}^{-1+\delta}) \int_{\mathbb{R}^{j-2}} |\Phi_{\tau_1}(x, w_1) \dots \Phi_{\tau_{j-2}}(w_{j-3}, w_{j-2})| dw_1 \dots dw_{j-2} \\ &\quad \times (C\tau_{j+1}^{-1+\delta_\infty}) \int_{\mathbb{R}^{k-j-2}} |\Phi_{\tau_{j+2}}(w_{j+1}, w_{j+2}) \dots \Phi_{\tau_{k-1}}(w_{k-2}, w_{k-1})| dw_{j+2} \dots dw_{k-1} \\ &\leq \dots \leq C^k \left( \prod_{i=1}^{j-1} \tau_i^{-1+\delta} \right) \tau_j^{-1-1/\alpha+\delta} \left( \prod_{i=j+1}^k \tau_i^{-1+\delta_\infty} \right) \leq \tau_j^{-1-1/\alpha-\delta} \left( \prod_{i \neq j, i \leq k} \tau_i^{-1+\delta_\infty} \right). \end{aligned}$$

Now we take  $0 \leq s_1 \leq \dots \leq s_{k-1} \leq t$  and put  $s_0 = 0, s_k = t, \tau_i = s_i - s_{i-1}, i = 1, \dots, k$ . Then the maximal value  $\tau_j$  is  $\geq t/k$ , and we get

$$\begin{aligned} |\Phi_t^{\otimes k}(x, y)| &\leq k^{1/\alpha} t^{-1/\alpha} C^k \sum_{j=1}^k \int_{0 \leq s_1 \leq \dots \leq s_{k-1} \leq t} \left( \prod_{i \neq j, i \leq k} (s_i - s_{i-1})^{-1+\delta_\infty} \right) (s_j - s_{j-1}) ds_1, \dots, ds_k \\ &\leq t^{-1/\alpha+\delta+(k-1)\delta_\infty} C^k k^{1/\alpha+1} \frac{\Gamma(\delta_\infty)^{k-1} \Gamma(\delta)}{\Gamma((k-1)\delta_\infty + \delta)}. \end{aligned}$$

Taking the sum in  $k \geq 1$ , we obtain the first bound in (6.5).

We also have

$$p_t^0(x, y) \leq Ct^{-1/\alpha}, \quad \sup_y \int_{\mathbb{R}} p_t^0(x, y) dx \leq C.$$

Repeating the calculation used in the proof of (6.5), we get

$$\begin{aligned} |r_t(x, y)| &= \left( \int_0^{t/2} + \int_{t/2}^t \right) |(p_{t-s}^0 * \Psi_s)(x, y)| ds \\ &\leq C2^{1/\alpha} t^{-1/\alpha} t^{\delta_\infty} + C2^{1+1/\alpha-\delta} t^{-1/\alpha+\delta} \leq Ct^{-1/\alpha+\delta_\infty}. \end{aligned}$$

Combined with (4.33), this completes the proof.

### 7 Proof of Theorem 3.5

We further analyze the bound (4.22) under the stronger assumption  $\mathbf{H}^\nu$  (ii). To simplify the notation, we assume  $\gamma \leq \alpha$  and write  $\gamma$  instead of  $\gamma'$ . This does not restrict generality because decreasing  $\gamma$  in the assumption (3.4) leaves this assumption true. We

have

$$\begin{aligned} Q_t(x, y) &= t^{\beta/\alpha} |\Phi_t^{\nu, large, +}(x, y)| \\ &\leq C t^{\beta/\alpha} \int_{|u| > t^{1/\alpha}} t^{-1/\alpha} G^{(\alpha)}\left(\frac{\kappa_t(y) - x - u}{t^{1/\alpha}}\right) \left(|u|^{-\beta-1} 1_{|u| \leq 1} + |u|^{-\gamma-1} 1_{|u| > 1}\right) du \\ &= C \int_{|u| > t^{1/\alpha}} G_t^{(\alpha, \alpha, \alpha)}(x + u, \kappa_t(y)) G_t^{(\alpha, \beta, \gamma)}(0, u) du \\ &\leq C \left(G_t^{(\alpha, \alpha, \alpha)} * G_t^{(\alpha, \beta, \gamma)}\right)(x, \kappa_t(y)), \end{aligned}$$

in the last inequality we have used that  $G_t^{(\alpha, \alpha, \alpha)}(x, y)$  and  $G_t^{(\alpha, \beta, \gamma)}(x, y)$  actually depend on  $|x - y|$ , only. Recall that  $\beta < \alpha, \gamma \leq \alpha$ . Then it is a direct calculation to check that

$$G_t^{(\alpha, \alpha, \alpha)}(x, y) \leq G_t^{(\alpha, \beta, \gamma)}(x, y), \quad G_{2t}^{(\alpha, \beta, \gamma)}(x, y) \leq C G_t^{(\alpha, \beta, \gamma)}(x, y).$$

Then it follows from the sub-convolution property for  $G_t^{(\alpha, \beta, \gamma)}(x, y)$  (see Appendix A.5) that

$$Q_t(x, y) \leq C G_t^{(\alpha, \beta, \gamma)}(x, \kappa_t(y)).$$

That is, by (4.22) we have

$$\begin{aligned} |\Phi_t(x, y)| &\leq C t^{-1+\delta} G_t^{(\alpha, \alpha, \alpha)}(x, \kappa_t(y)) \\ &\quad + C t^{-1+\delta_\zeta} G_t^{(\alpha, \alpha-\zeta, \alpha)}(x, \kappa_t(y)) + C t^{-1+\delta_\beta} G_t^{(\alpha, \beta, \gamma)}(x, \kappa_t(y)). \end{aligned}$$

Since

$$\begin{aligned} t^{-1+\delta_\zeta} G_t^{(\alpha, \alpha-\zeta, \alpha)}(x, y) &= \begin{cases} t^{-1/\alpha-(\alpha-\zeta)/\alpha}, & |y-x| \leq t^{1/\alpha} \\ |y-x|^{-(\alpha-\zeta)-1}, & t^{1/\alpha} < |y-x| \leq 1, \\ |y-x|^{-\alpha-1}, & |y-x| > 1, \end{cases} \\ t^{-1+\delta_\beta} G_t^{(\alpha, \beta, \gamma)}(x, y) &= \begin{cases} t^{-1/\alpha-\beta/\alpha}, & |y-x| \leq t^{1/\alpha} \\ |y-x|^{-\beta-1}, & t^{1/\alpha} < |y-x| \leq 1, \\ |y-x|^{-\gamma-1}, & |y-x| > 1, \end{cases} \end{aligned}$$

the sum of these kernels satisfies

$$\begin{aligned} t^{-1+\delta_\zeta} G_t^{(\alpha, \alpha-\zeta, \alpha)}(x, y) + t^{-1+\delta_\beta} G_t^{(\alpha, \beta, \gamma)}(x, y) &\leq 2 \begin{cases} t^{-1/\alpha-\beta'/\alpha}, & |y-x| \leq t^{1/\alpha} \\ |y-x|^{-\beta'-1}, & t^{1/\alpha} < |y-x| \leq 1, \\ |y-x|^{-\gamma-1}, & |y-x| > 1, \end{cases} \\ &= 2 t^{-1+\delta'} G_t^{(\alpha, \beta', \gamma)}(x, y), \end{aligned}$$

where

$$\beta' = \max(\beta, \alpha - \zeta), \quad \delta' = \frac{\alpha - \beta'}{\alpha} > 0$$

(recall that we have assumed  $\gamma \leq \alpha$ ). This finally gives the bound

$$|\Phi_t(x, y)| \leq C t^{-1+\delta} H_t^1(x, y) + C t^{-1+\delta'} H_t^2(x, y) \tag{7.1}$$

with

$$H_t^1(x, y) = G_t^{(\alpha, \alpha, \alpha)}(x, \kappa_t(y)) = \frac{1}{t^{1/\alpha}} G^{(\alpha)}\left(\frac{\kappa_t(y) - x}{t^{1/\alpha}}\right), \quad H_t^2(x, y) = G_t^{(\alpha, \beta', \gamma)}(x, \kappa_t(y)). \tag{7.2}$$

Denote  $\delta_1 = \delta, \delta_2 = \delta'$ . For any  $k > 1$  we have

$$\begin{aligned} |\Phi_t^{\otimes k}(x, y)| &\leq C^k \sum_{i_1, \dots, i_k \in \{1, 2\}} \int_{0 < s_1 < \dots < s_{k-1} < t} \left( \prod_{j=1}^k (s_j - s_{j-1})^{-1+\delta_{i_j}} \right) \times \\ &\quad \times \left( H_{s_1}^{i_1} * \dots * H_{t-s_k}^{i_k} \right)(x, y) ds_1, \dots, ds_{k-1}. \end{aligned}$$

The kernels  $H^1, H^2$  satisfy  $H_t^1(x, y) \leq H_t^2(x, y)$  and have the sub-convolution property, see Proposition A.10 below. Then for  $t \in (0, T]$

$$\begin{aligned} & \int_{0 < s_1 < \dots < s_{k-1} < t} \left( \prod_{j=1}^k (s_j - s_{j-1})^{-1+\delta_1} \right) \left( H_{s_1}^1 * \dots * H_{t-s_k}^1 \right) (x, y) ds_1, \dots, ds_{k-1} \\ & \leq C^k H_t^1(x, y) \int_{0 < s_1 < \dots < s_{k-1} < t} \left( \prod_{j=1}^k (s_j - s_{j-1})^{-1+\delta_1} \right) ds_1, \dots, ds_{k-1} \\ & = t^{-1+k\delta_1} \frac{C^k \Gamma(\delta_1)^k}{\Gamma(k\delta_1)} H_t^1(x, y) \leq \tilde{C} t^{-1+\delta_1} \frac{C^k \Gamma(\delta_1)^k}{\Gamma(k\delta_1)} H_t^1(x, y), \end{aligned}$$

and (recall that  $\delta_1 < \delta_2$ )

$$\begin{aligned} & \int_{0 < s_1 < \dots < s_{k-1} < t} \left( \prod_{j=1}^k (s_j - s_{j-1})^{-1+\delta_{i_j}} \right) \left( H_{s_1}^{i_1} * \dots * H_{t-s_k}^{i_k} \right) (x, y) ds_1, \dots, ds_{k-1} \\ & \leq C^k H_t^2(x, y) \int_{0 < s_1 < \dots < s_{k-1} < t} \left( \prod_{j=1}^k (s_j - s_{j-1})^{-1+\delta_{i_j}} \right) ds_1, \dots, ds_{k-1} \\ & = t^{-1+\sum_j \delta_{i_j}} \frac{C^k \prod_j \Gamma(\delta_{i_j})}{\Gamma(\sum_j \delta_{i_j})} H_t^2(x, y) \leq \tilde{C} t^{-1+\delta_2} \frac{C^k \Gamma(\delta_2)^k}{\Gamma(k\delta_1)} H_t^2(x, y), \end{aligned}$$

if at least one of the indices  $i_1, \dots, i_k$  equals 2. Thus

$$|\Psi_t(x, y)| \leq \sum_{k \geq 1} |\Phi_t^{\otimes k}(x, y)| \leq C t^{-1+\delta_1} H_t^1(x, y) + C t^{-1+\delta_2} H_t^2(x, y). \tag{7.3}$$

Recall that  $\delta_1 = \delta, \delta_2 = \delta'$  and

$$r_t(x, y) = (p \otimes \Psi)_t(x, y), \quad p_t^0(x, y) \leq C H_t^1(x, y).$$

Then, using the sub-convolution properties of  $H^1, H^2$  and the inequality  $H^1 \leq H^2$  in the same way we did before, we get

$$|r_t(x, y)| \leq C \Delta(t) H_t^1(x, y) + C t^{\delta'} H_t^2(x, y). \tag{7.4}$$

In the notation from the proof of Proposition A.10, we have

$$H_t^1(x, y) = F_t^{(\alpha, \alpha, \alpha)} \left( \frac{x - \kappa_t(y)}{t^{1/\alpha}} \right), \quad H_t^2(x, y) = F_t^{(\alpha, \beta', \gamma)} \left( \frac{x - \kappa_t(y)}{t^{1/\alpha}} \right).$$

Using (A.17) and (A.40), (A.41), we get

$$|r_t(x, y)| \leq C \Delta(t) G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y) + C t^{\delta'} G_t^{(\alpha, \beta', \gamma')}(\chi_t(x), y). \tag{7.5}$$

Combined with (4.33), this completes the proof.

## A Appendix

### A.1 Proof of (3.9), (3.10)

Denote

$$N_\beta(\varepsilon) := \begin{cases} \frac{1}{|1-\beta|} \varepsilon^{1-\beta}, & \beta \in (0, 2), \beta \neq 1; \\ 1 + \log \varepsilon^{-1}, & \beta = 1. \end{cases}$$

The proof of the following statement is easy and omitted.

**Proposition A.1.** Let  $\nu(du)$  be a measure satisfying

$$\nu(|u| > r) \leq C_\nu r^{-\beta}, \quad r \in (0, 1]$$

for some  $\beta \in (0, 2)$ . Then

$$\int_{\varepsilon < |u| \leq 1} |u| \sigma(du) \leq CN_\beta(\varepsilon), \quad \varepsilon \leq 1 \tag{A.1}$$

for  $\beta \in [1, 2)$ , and

$$\int_{|u| \leq \varepsilon} |u| \nu(du) \leq CN_\beta(\varepsilon), \quad \varepsilon \leq 1, \tag{A.2}$$

for  $\beta \in (0, 1)$ . In addition, for any  $\beta \in (0, 2)$

$$\int_{|u| \leq \varepsilon} |u|^2 \nu(du) \leq C\varepsilon^{2-\beta}, \quad \varepsilon \leq 1. \tag{A.3}$$

The constants  $C$  in (A.1), (A.2), and (A.3) depend on  $\beta$  and  $C_\nu$ , only.

**Proposition A.2.** Let  $f$  be such that for some  $\sigma \in [0, 1]$

$$\|f\|_{H_{\sigma,loc}} := \sup_{x \neq y, |x-y| \leq 1} \frac{|f(x) - f(y)|}{|x - y|^\sigma} < \infty. \tag{A.4}$$

Then for each  $t \in (0, T]$

$$F_t(x) := \frac{1}{2\sqrt{\pi}t^{1/\alpha}} \int_{\mathbb{R}} e^{-z^2 t^{-2/\alpha}} f(x - z) dz$$

satisfies

$$\sup_x |f(x) - F_t(x)| \leq C_{\sigma,\alpha,T} t^{\sigma/\alpha} \|f\|_{H_{\sigma,loc}}, \tag{A.5}$$

and  $F_t$  is Lipschitz continuous with  $\text{Lip}(f_t) \leq C_{\sigma,\alpha,T} t^{\sigma/\alpha - 1/\alpha} \|f\|_{H_{\sigma,loc}}$ .

*Proof.* It follows from (A.4) that for  $|x - y| \geq 1$

$$|f(x) - f(y)| \leq 2|x - y| \|f\|_{H_{\sigma,loc}}.$$

This inequality for large  $|x - y|$ , combined with the inequality (A.4) for small  $|x - y|$  yields the following bound valid for all  $x, y \in \mathbb{R}$ :

$$|f(x) - f(y)| \leq 2(|x - y|^\sigma \vee |x - y|) \|f\|_{H_{\sigma,loc}} \leq 2(|x - y|^\sigma + |x - y|) \|f\|_{H_{\sigma,loc}}. \tag{A.6}$$

Then

$$\begin{aligned} |F_t(x) - f(x)| &\leq \frac{1}{2\sqrt{\pi}t^{1/\alpha}} \int_{\mathbb{R}} e^{-z^2 t^{-2/\alpha}} |f(x - z) - f(x)| dz \\ &\leq \frac{1}{2\sqrt{\pi}t^{1/\alpha}} \|f\|_{H_{\sigma,loc}} \int_{\mathbb{R}} e^{-z^2 t^{-2/\alpha}} (C|z|^\sigma + C|z|) dz \\ &= (C_1 t^{\sigma/\alpha} + C_2 t^{1/\alpha}) \|f\|_{H_{\sigma,loc}}, \end{aligned}$$

which proves (A.5). Since

$$\begin{aligned} \partial_x F_t(x) &= \frac{1}{2\sqrt{\pi}t^{1/\alpha}} \int_{\mathbb{R}} e^{-z^2 t^{-2/\alpha}} f'(x - z) dz \\ &= -\frac{1}{\sqrt{\pi}t^{1/\alpha}} \int_{\mathbb{R}} z t^{-2/\alpha} e^{-z^2 t^{-2/\alpha}} f(x - z) dz \\ &= \frac{1}{\sqrt{\pi}t^{1/\alpha}} \int_{\mathbb{R}} z t^{-2/\alpha} e^{-z^2 t^{-2/\alpha}} (f(x) - f(x - z)) dz, \end{aligned}$$



similar calculation gives

$$\text{Lip}(F_t) \leq \sup_x |\partial_t f(x)| \leq \left( C_1 t^{\sigma/\alpha-1} + C_2 t^{1/\alpha-1} \right) \|f\|_{H_{\sigma,loc}}. \quad \square$$

Now we are ready to prove (3.9), (3.10). We decompose

$$b_t(x) = \tilde{b}(x) + \tilde{m}_t^{(\alpha)}(x) + \tilde{m}_t^\nu(x) =: f^1(x) + f^{2,t}(x) + f^{3,t}(x), \quad (\text{A.7})$$

where we denote

$$\tilde{m}_t^{(\alpha)}(x) = \begin{cases} - \int_{t^{1/\alpha} < |u| \leq 1} u \mu^{(\alpha)}(x; du), & \alpha \in [1, 2), \\ \int_{|u| \leq t^{1/\alpha}} u \mu^{(\alpha)}(x; du), & \alpha \in (0, 1), \end{cases}$$

$$\tilde{m}_t^\nu(x) = \begin{cases} - \int_{t^{1/\alpha} < |u| \leq 1} u \nu(x; du), & \beta \in [1, 2), \\ \int_{|u| \leq t^{1/\alpha}} u \nu(x; du), & \beta \in (0, 1). \end{cases}$$

By condition (3.2), we have  $\|f^1\|_{H_{\eta,loc}} \leq C$ . By condition  $\mathbf{H}^{(\alpha)}$ (i) and Proposition A.1, we have  $\|f^{2,t}\|_{H_{\zeta,loc}} \leq CN_\alpha(t^{1/\alpha})$ . Finally, by (2.3), (3.3), and Proposition A.1, we have  $\|f^{3,t}\|_{H_{0,loc}} \leq CN_\beta(t^{1/\alpha})$ . Applying Proposition A.2, we get (3.9), (3.10):

$$|b_t(x) - B_t(x)| \leq Ct^{\eta/\alpha} + Ct^{\zeta/\alpha} N_\alpha(t^{1/\alpha}) + CN_\beta(t^{1/\alpha}) \leq Ct^\delta,$$

$$\text{Lip}(B_t) \leq Ct^{\eta/\alpha-1} + Ct^{\zeta/\alpha-1} N_\alpha(t^{1/\alpha}) + Ct^{-1} N_\beta(t^{1/\alpha}) \leq Ct^{-1+\delta}.$$

Note that the above calculation also gives

$$|b_t(x) - \tilde{b}(x)| \leq CN_\alpha(t^{1/\alpha}) \leq Ct^{-1+1/\alpha} \left( 1 + 1_{\alpha=1} \log_+ t^{-1} \right). \quad (\text{A.8})$$

## A.2 Auxiliary family $\chi_s^t(x)$ and properties of $\chi_t(x)$ , $\kappa_t(y)$

**Proposition A.3.** For any  $T > 0$  there exists  $C > 1$  such that

$$C^{-1}|x| \leq |\chi_t(x)| \leq C|x|, \quad 0 \leq t \leq T, \quad |x| \geq C. \quad (\text{A.9})$$

*Proof.* By  $\mathbf{H}^{drift}$ , the function  $\tilde{b}$  satisfies (A.4) with  $\sigma = \eta$ . Therefore for this function (A.6) with  $\sigma = \eta$  holds. Then by (3.9) and (A.8) the coefficient  $B_t$  in the ODE, which defines  $\chi_t$ , satisfies the following linear growth bound:

$$|B_t(x)| \leq C_1 + C_2 t^{-1+1/\alpha} \left( 1 + 1_{\alpha=1} \log_+ t^{-1} \right) + C_3|x|.$$

This in a standard way provides

$$e^{-C_3}|x| - C_4 \leq |\chi_t(x)| \leq e^{C_3}|x| + C_4. \quad \square$$

In order to relate the families  $\chi_s(x)$ ,  $\kappa_s(y)$ , we introduce an auxiliary family  $\chi_s^t(x)$ , the solution to the Cauchy problem

$$\frac{d}{ds} \chi_s^t(x) = B_{t-s}(\chi_s^t(x)), \quad s \in [0, t], \quad \chi_0^t(x) = x. \quad (\text{A.10})$$

**Proposition A.4.** For any  $T > 0$  there exists  $C$  such that for any  $s \leq t \leq T$

$$e^{-Ct^\delta} |\kappa_t(y) - x| \leq |\kappa_{t-s}(y) - \chi_s^t(x)| \leq e^{Ct^\delta} |\kappa_t(y) - x|. \quad (\text{A.11})$$

*Proof.* Denote  $x_s = \chi_s^t(x), y_s = \kappa_{t-s}(y)$ , then

$$(x_s - y_s)' = (x_s - y_s)q_{t,s}, \quad q_{t,s} = \frac{B_{t-s}(x_s) - B_{t-s}(y_s)}{x_s - y_s}$$

with the convention  $\frac{0}{0} = 1$ . Then

$$\chi_s^t(x) - \kappa_{t-s}(y) = x_s - y_s = (x_0 - y_0) \exp\left(\int_0^s q_{t,r} dr\right) = (x - \kappa_t(y)) \exp\left(\int_0^s q_{t,r} dr\right),$$

which provides the required statement by (3.10) since  $|q_{t,r}| \leq \text{Lip}(B_{t-r})$ .  $\square$

**Proposition A.5.** For any  $T > 0$  and  $\delta < \min(\delta_\eta, \delta_\zeta, \delta_\beta)$  there exist  $C$  such that for

$$\chi_s^t(x) = \chi_s(x) + t^{1/\alpha} \int_0^s v(\chi_r(x)) (W_\alpha(t; r) - W_\alpha(t; t-r)) dr + Q_{s,t}(x), \quad s \leq t \quad (\text{A.12})$$

with

$$|Q_{s,t}(x)| \leq Ct^{1/\alpha+\delta}, \quad s \leq t \leq T. \quad (\text{A.13})$$

*Proof.* Denote  $x_s = \chi_s^t(x), \tilde{x}_s = \chi_s(x)$ , then

$$(x_s - \tilde{x}_s)' = (x_s - \tilde{x}_s)\tilde{q}_{t,s} + \tilde{Q}_{t,s}$$

with

$$\tilde{q}_{t,s} = \frac{B_{t-s}(x_s) - B_{t-s}(\tilde{x}_s)}{x_s - \tilde{x}_s}, \quad \tilde{Q}_{t,s} = B_{t-s}(\tilde{x}_s) - B_s(\tilde{x}_s),$$

and thus

$$x_s - \tilde{x}_s = \int_0^s \tilde{Q}_{t,r} \exp\left(\int_r^s \tilde{q}_{t,w} dw\right) dr. \quad (\text{A.14})$$

By (3.10),

$$\left| \exp\left(\int_r^s \tilde{q}_{t,w} dw\right) - 1 \right| \leq Ct^\delta. \quad (\text{A.15})$$

On the other hand, by (3.9)

$$|\tilde{Q}_{t,r} - (b_{t-r}(\tilde{x}_r) - b_r(\tilde{x}_r))| \leq C((t-r)^{-1+1/\alpha+\delta} + r^{-1+1/\alpha+\delta}),$$

and by (3.8)

$$b_{t-r}(x) - b_r(x) = m_r^\mu(x) - m_{t-r}^\mu(x) = (m_r^{(\alpha)}(x) - m_{t-r}^{(\alpha)}(x)) + (m_r^\nu(x) - m_{t-r}^\nu(x)),$$

where we denote

$$m_r^{(\alpha)}(x) = \int_{r^{1/\alpha} < |u| \leq 1} u\mu^{(\alpha)}(x, du), \quad m_r^\nu(x) = \int_{r^{1/\alpha} < |u| \leq 1} u\nu(x, du).$$

Assume for a while that  $r \leq t-r$ . By Proposition A.1,

$$|m_r^\nu(x) - m_{t-r}^\nu(x)| \leq \int_{r^{1/\alpha} < |u| \leq (t-r)^{1/\alpha}} |u| |\nu|(x, du) \leq \begin{cases} CN_\beta((t-r)^{1/\alpha}), & \beta \in (0, 1); \\ CN_\beta(r^{1/\alpha}), & \beta \in [1, 2). \end{cases}$$

Similarly, for  $t-r \leq r \leq t$

$$|m_r^\nu(x) - m_{t-r}^\nu(x)| \leq \int_{(t-r)^{1/\alpha} < |u| \leq r^{1/\alpha}} |u| |\nu|(x, du) \leq \begin{cases} CN_\beta(r^{1/\alpha}), & \beta \in (0, 1); \\ CN_\beta((t-r)^{1/\alpha}), & \beta \in [1, 2). \end{cases}$$

That is, in any case we have

$$|m_r^\nu(x) - m_{t-r}^\nu(x)| \leq CN_\beta(r^{1/\alpha}) + CN_\beta((t-r)^{1/\alpha}) \leq C\left((t-r)^{-1+1/\alpha+\delta} + r^{-1+1/\alpha+\delta}\right).$$

On the other hand, for  $r \leq t$  we have

$$\begin{aligned} m_r^{(\alpha)}(x) &= m_t^{(\alpha)}(x) + \int_{r^{1/\alpha} < |u| \leq t^{1/\alpha}} u\mu^{(\alpha)}(x, du) \\ &= m_t^{(\alpha)}(x) + v(x) \int_{r^{1/\alpha}}^{t^{1/\alpha}} \frac{dw}{w^\alpha} = m_t^{(\alpha)}(x) + v(x)t^{1/\alpha}W(t; r). \end{aligned}$$

Summarizing these calculations we get

$$\left| \tilde{Q}_{t,r} - v(\tilde{x}_r)t^{1/\alpha}\left(W_\alpha(t; r) - W_\alpha(t; t-r)\right) \right| \leq C\left((t-r)^{-1+1/\alpha+\delta} + r^{-1+1/\alpha+\delta}\right).$$

This bound combined with (A.14) and (A.15), provides (A.12) and (A.13). □

Recall that  $v(\cdot)$  is bounded and  $W_\alpha(t; \cdot)$  is a probability density. That is, directly from (A.12), (A.13) we get the bound

$$|\chi_s^t(x) - \chi_s(x)| \leq Ct^{1/\alpha}, \quad s \leq t \leq T. \tag{A.16}$$

Combined with Proposition A.4, this gives the following.

**Corollary A.6.** *For each  $T > 0$ , there exists  $C$  such that for any  $s \leq t \leq T$*

$$e^{-Ct^\delta} |\kappa_t(y) - x| - Ct^{1/\alpha} \leq |\kappa_{t-s}(y) - \chi_s(x)| \leq e^{Ct^\delta} |\kappa_t(y) - x| + Ct^{1/\alpha}, \quad s \leq t \leq T. \tag{A.17}$$

### A.3 Stable densities

The kernel  $G^{(\alpha)}(x)$  (see the definition before Corollary 3.6) possess the following properties which can be verified straightforwardly:

$$G^{(\alpha)}(x) \leq G^{(\beta)}(x), \quad 0 < \beta < \alpha; \tag{A.18}$$

$$(1 + |x|)^\beta G^{(\alpha)}(x) \leq CG^{(\alpha-\beta)}(x), \quad 0 < \beta < \alpha; \tag{A.19}$$

$$\sup_{|v| \leq 1} G^{(\alpha)}(x + v) \leq CG^{(\alpha)}(x), \tag{A.20}$$

and for any  $c > 0$  there exists  $C$  such that

$$G^{(\alpha)}(cx) \leq CG^{(\alpha)}(x). \tag{A.21}$$

We also have

$$\left(G^{(\alpha)} * G^{(\alpha)}\right)(x) \leq CG^{(\alpha)}(x). \tag{A.22}$$

The following two propositions collect the properties of the  $\alpha$ -stable densities  $g^{(\lambda, \rho, v)}(x)$ , see (3.6) for the definition.

**Proposition A.7.** *The density  $g^{(\lambda, \rho, v)}(x)$  is well defined and belongs to the class  $C^1$  w.r.t.  $(\lambda, \rho)$ , and the class  $C_\infty^2$  w.r.t.  $x$ . The following bounds hold true for each  $\alpha \in (0, 2)$ ,  $0 < \lambda_{\min} \leq \lambda_{\max}$ ,  $R > 0$  uniformly in  $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ ,  $\rho \in [-1, 1]$ ,  $|v| \leq R$ ,  $x \in \mathbb{R}$ :*

$$g^{(\lambda, \rho, v)}(x) \leq CG^{(\alpha)}(x), \tag{A.23}$$

$$|\partial_\lambda g^{(\lambda, \rho, v)}(x)| \leq CG^{(\alpha)}(x), \tag{A.24}$$

$$|\partial_\rho g^{(\lambda, \rho, v)}(x)| \leq CG^{(\alpha)}(x), \tag{A.25}$$

*Proof.* We use the standard trick of a decomposition of an infinitely divisible law into a convolution of ‘small’ and ‘large’ jump parts. Namely, we put

$$\Psi_{\alpha,small}^{(\lambda,\rho,v)}(\xi) = i\xi v + \int_{|u|\leq 1} (e^{iu\xi} - 1 - iu\xi) \mu^{(\alpha,\lambda,\rho)}(du),$$

$$\Psi_{\alpha,large}^{(\lambda,\rho)}(\xi) = \int_{|u|>1} (e^{iu\xi} - 1) \mu^{(\alpha,\lambda,\rho)}(du),$$

and observe that  $\exp \Psi_{\alpha,large}^{(\lambda,\rho)}(\xi)$  is the Fourier transform of a compound Poisson process with the intensity of the ‘Poisson clock’ equal

$$\int_{|u|>1} \mu^{(\alpha,\lambda,\rho)}(du) = \frac{2\lambda}{\alpha},$$

and with the law of a single jump having the density

$$m^{(\alpha,\rho)}(u) = \frac{\alpha(1 + \rho \operatorname{sgn} u)}{2|u|^{\alpha+1}} 1_{|u|>1}.$$

On the other hand,

$$\Psi_{\alpha,small}^{(\lambda,\rho,v)}(\xi) = i\xi v - \lambda \int_{|u|\leq 1} \frac{1 - \cos(u\xi)}{|u|^{\alpha+1}} du - i\lambda\rho \int_{|u|\leq 1} \frac{u\xi \sin(u\xi)}{|u|^{\alpha+1}} du,$$

and in particular

$$\operatorname{Re} \Psi_{\alpha,small}^{(\lambda,\rho,v)}(\xi) = -\lambda \int_{|u|\leq 1} \frac{1 - \cos(u\xi)}{|u|^{\alpha+1}} du \leq -c_1|\xi|^\alpha + c_2 \tag{A.26}$$

with some positive  $c_1, c_2$ . This yields that the inverse Fourier transform

$$f^{(\lambda,\rho,v)}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ix\xi + \Psi_{\alpha,small}^{(\lambda,\rho,v)}(\xi)} d\xi$$

is well defined. Then the density  $g^{(\lambda,\rho,v)}(x)$  is also well defined and possesses the representation

$$g^{(\lambda,\rho,v)}(x) = e^{-2\lambda/\alpha} f^{(\lambda,\rho,v)}(x) + e^{-2\lambda/\alpha} \sum_{k=1}^{\infty} \frac{1}{k!} \left( f^{(\lambda,\rho,v)} * [\tilde{m}^{(\lambda,\rho,v)}]^{*k} \right)(x) \tag{A.27}$$

with

$$\tilde{m}^{(\alpha,\lambda,\rho)}(\xi) = (2\lambda/\alpha) m^{(\alpha,\rho)}(\xi).$$

We claim that, uniformly in  $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ ,  $\rho \in [-1, 1]$ ,  $|v| \leq R$ ,  $x \in \mathbb{R}$ ,

$$|f^{(\lambda,\rho,v)}(x)| + |\partial_\lambda f^{(\lambda,\rho,v)}(x)| + |\partial_\rho f^{(\lambda,\rho,v)}(x)| + |\partial_x f^{(\lambda,\rho,v)}(x)| + |\partial_{xx}^2 f^{(\lambda,\rho,v)}(x)| \leq C e^{-|x|}. \tag{A.28}$$

The argument here is quite standard (e.g. [9]), but for the sake of completeness we outline the proof. The function  $\Psi_{\alpha,small}^{(\lambda,\rho,v)}$  is defined as an integral over the bounded interval  $[-1, 1]$ , and thus has an analytic extension to  $\mathbb{C}$ :

$$\begin{aligned} \Psi_{small}^{(\lambda,\rho,v)}(\xi + i\varphi) &= i\xi v - \varphi v + \int_{|u|\leq 1} (e^{-u\varphi + iu\xi} - 1 - iu\xi + u\varphi) \mu^{(\alpha,\lambda,\rho)}(du) \\ &= \int_{|u|\leq 1} e^{-u\varphi} (e^{iu\xi} - 1 - iu\xi) \mu^{(\alpha,\lambda,\rho)}(du) \\ &\quad + \int_{|u|\leq 1} (e^{-u\varphi} - 1 + u\varphi) \mu^{(\alpha,\lambda,\rho)}(du) \\ &\quad - \varphi v + i\xi v + i \int_{|u|\leq 1} (e^{-u\varphi} - 1) u\xi \mu^{(\alpha,\lambda,\rho)}(du). \end{aligned}$$

Then for  $\varphi \in [-1, 1]$  we have

$$\begin{aligned} \operatorname{Re} \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi) &= - \int_{|u| \leq 1} e^{-u\varphi} (1 - \cos(u\xi)) \mu^{(\alpha, \lambda, \rho)}(du) \\ &\quad + \int_{|u| \leq 1} (e^{-u\varphi} - 1 + u\varphi) \mu^{(\alpha, \lambda, \rho)}(du) - \varphi v \\ &\leq -e^{-1} \lambda \int_{|u| \leq 1} \frac{1 - \cos(u\xi)}{|u|^{\alpha+1}} du \leq -e^{-1} (c_1 |\xi|^\alpha + c_2), \quad \xi \in \mathbb{R}, \end{aligned} \tag{A.29}$$

see (A.26). This makes it possible to change the integration contour in the inverse Fourier transform formula from  $\mathbb{R} = \mathbb{R} + i0$  to  $\mathbb{R} + i\varphi$ , which gives

$$f^{(\lambda, \rho, v)}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ix\xi + x\varphi + \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi)} d\xi.$$

Take  $\varphi = \varphi_x = -\operatorname{sgn} x$ , then

$$f^{(\lambda, \rho, v)}(x) = \frac{e^{-|x|}}{2\pi} \int_{\mathbb{R}} e^{-ix\xi + \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi_x)} d\xi.$$

This representation and (A.29) give

$$\begin{aligned} &|f^{(\lambda, \rho, v)}(x)| + |\partial_\lambda f^{(\lambda, \rho, v)}(x)| + |\partial_\rho f^{(\lambda, \rho, v)}(x)| + |\partial_x f^{(\lambda, \rho, v)}(x)| + |\partial_{xx}^2 f^{(\lambda, \rho, v)}(x)| \\ &\leq C e^{-|x|} \int_{\mathbb{R}} e^{\operatorname{Re} \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi_x)} \\ &\quad \times \left( 1 + |\partial_\lambda \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi_x)| + |\partial_\rho \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi_x)| + |\xi| + \xi^2 \right) d\xi. \end{aligned}$$

It is easy to check that

$$|\partial_\lambda \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi_x)| + |\partial_\rho \Psi_{\alpha, \text{small}}^{(\lambda, \rho, v)}(\xi + i\varphi_x)| \leq C(1 + \xi^2),$$

hence (A.28) follows by (A.29).

Next, we give explicitly the function  $\tilde{m}^{(\alpha, \lambda, \rho)}(u)$  and its derivatives:

$$\begin{aligned} \tilde{m}^{(\alpha, \lambda, \rho)}(u) &= \lambda \frac{1 + \rho \operatorname{sgn} u}{|u|^{\alpha+1}} 1_{|u| > 1}, \\ \partial_\lambda \tilde{m}^{(\alpha, \lambda, \rho)}(u) &= \frac{1 + \rho \operatorname{sgn} u}{|u|^{\alpha+1}} 1_{|u| > 1}, \quad \partial_\rho \tilde{m}^{(\alpha, \lambda, \rho)}(u) = \frac{\lambda \operatorname{sgn} u}{|u|^{\alpha+1}} 1_{|u| > 1}, \end{aligned}$$

and observe that the absolute values of these functions are dominated by  $CG^{(\alpha)}(u)$ .

Now we can finalize the proof. It follows from (A.28) that

$$f^{(\lambda, \rho, v)}(x) \leq CG^{(\alpha)}(x),$$

then taking  $C$  large enough we obtain inductively

$$\left( f^{(\lambda, \rho, v)} * [\tilde{m}^{(\alpha, \lambda, \rho)}]^{*k} \right)(x) \leq C^{2k+1} G^{(\alpha)}(x),$$

and applying (A.27) we complete the proof of (A.23). The proofs of (A.24), (A.25) are essentially the same. The minor difference is that respective derivatives of  $f^{(\lambda, \rho, v)} * [\tilde{m}^{(\alpha, \lambda, \rho)}]^{*k}$  now actually contain  $(k + 1)$  summands, each of them being a  $(k + 1)$ -fold convolution where each term is dominated by  $CG^{(\alpha)}(x)$ ; however the extra multiplier  $(k + 1)$  is not essential thanks to the term  $1/k!$  in (A.27).  $\square$

**Proposition A.8.** *The density  $g^{(\lambda,\rho,v)}(x)$  belongs to the class  $C_\infty^2$  w.r.t.  $x$ . The following bounds hold true for each  $\alpha \in (0, 2)$ ,  $0 < \lambda_{\min} \leq \lambda_{\max}$ ,  $R > 0$  uniformly in  $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ ,  $\rho \in [-1, 1]$ ,  $|v| \leq R$ ,  $x \in \mathbb{R}$ :*

$$|\partial_x g^{(\lambda,\rho,v)}(x)| \leq CG^{(\alpha+1)}(x), \tag{A.30}$$

$$|\partial_{xx}^2 g^{(\lambda,\rho,v)}(x)| \leq CG^{(\alpha+2)}(x). \tag{A.31}$$

$$|L_x^{(\alpha),sym} g^{(\lambda,\rho,v)}(x)| \leq CG^{(\alpha)}(x), \tag{A.32}$$

$$|L_x^{(\alpha),asym} g^{(\lambda,\rho,v)}(x)| \leq CG^{(\alpha)}(x), \tag{A.33}$$

see (4.18), (4.19) for the definition of  $L^{(\alpha),sym}$ ,  $L_x^{(\alpha),sym}$ .

*Proof.* The proofs of (A.32), (A.33) are completely analogous to the previous proof. Namely, using (A.28) it is easy to verify that

$$|L_x^{(\alpha),sym} f^{(\lambda,\rho,v)}(x)| + |L_x^{(\alpha),asym} f^{(\lambda,\rho,v)}(x)| \leq CG^{(\alpha)}(x).$$

Then the required bounds follow by (A.27) and (A.22). The new difficulty in (A.30), (A.31) is that the kernels in the right hand sides has the higher order of decay in  $x$ , and thus cannot be derived simply by (A.22). We will prove the first of these inequalities only: the second one is quite analogous, though the calculation is more cumbersome. Like we did that in the previous proof, we use representation (A.27) and analyze the derivatives of the terms in the right hand side sum. Note that by (A.26) the derivative  $\partial_x f^{(\lambda,\rho,v)}(x)$  is well defined and

$$\partial_x f^{(\lambda,\rho,v)}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} (-i\xi) e^{-ix\xi + \Psi_{\alpha,small}^{(\lambda,\rho,v)}(\xi)} d\xi.$$

Similarly to the previous proof, we deduce that

$$|\partial_x f^{(\lambda,\rho,v)}(x)| \leq C e^{-|x|} \leq CG^{(\alpha+1)}(x).$$

On the other hand, for  $u \neq \pm 1$  there exists

$$\partial_u \tilde{m}^{(\alpha,\lambda,\rho)}(u) = -(\alpha + 1) \lambda \frac{\text{sgn } u + \rho}{|u|^{\alpha+2}} 1_{|u|>1},$$

and the absolute value of the latter function is dominated by  $CG^{(\alpha+1)}(u)$ .

Let us prove the following: there exists  $C$  such that for any  $f \in C^1$  with  $|f(x)| \leq C_f G^{(\alpha)}(x)$ ,  $|\partial_x f(x)| \leq C_f G^{(\alpha+1)}(x)$  the following inequalities hold:

$$|(f * \tilde{m}^{(\alpha,\lambda,\rho)})(x)| \leq C C_f G^{(\alpha)}(x), \quad |\partial_x (f * \tilde{m}^{(\alpha,\lambda,\rho)})(x)| \leq C C_f G^{(\alpha+1)}(x). \tag{A.34}$$

This will easily yield

$$|\partial_x (f^{(\lambda,\rho,v)} * [\tilde{m}^{(\alpha,\lambda,\rho)}]^{*k})(x)| \leq C^{k+1} G^{(\alpha+1)}(x)$$

and complete the proof.

The first inequality in (A.34) follows just from (A.22). To prove the second inequality, we first note that (A.22) also yields that, for some  $C$ ,

$$|\partial_x (f * \tilde{m}^{(\alpha,\lambda,\rho)})(x)| = |(f' * \tilde{m}^{(\alpha,\lambda,\rho)})(x)|$$

is dominated by  $C C_f G^{(\alpha)}(x)$ . Hence it is sufficient to consider the case  $|x| > 2$ , only. Let  $x > 2$ , the case  $x < -2$  is quite analogous. We have

$$\begin{aligned} \partial_x (f * \tilde{m}^{(\alpha,\lambda,\rho)})(x) &= \int_{-\infty}^{x/2} f'(x-u) \tilde{m}^{(\alpha,\lambda,\rho)}(u) du + \int_{x/2}^{\infty} f'(x-u) \tilde{m}^{(\alpha,\lambda,\rho)}(u) du \\ &= \int_{-\infty}^{x/2} f'(x-u) \tilde{m}^{(\alpha,\lambda,\rho)}(u) du + \int_{x/2}^{\infty} f(x-u) \left( \tilde{m}^{(\alpha,\lambda,\rho)} \right)'(u) du \\ &\quad - f(x/2) \tilde{m}^{(\alpha,\lambda,\rho)}(x/2); \end{aligned}$$

note that  $\tilde{m}^{(\alpha,\lambda,\rho)}$  is smooth on  $[x/2, \infty) \subset (1, \infty)$ , hence we can apply the integration by parts formula here. The term  $f'(x - u)$  in the first integral and the term  $\left(\tilde{m}^{(\alpha,\lambda,\rho)}\right)'(u)$  in the second integral are dominated by  $CG^{(\alpha+1)}(x/2)$ . Then both these integrals are dominated by  $CC_f G^{(\alpha+1)}(x/2)$ , and we complete the proof of (A.34) using (A.21).  $\square$

**A.4 Properties of  $p_t^0(x, y)$ : proofs of (4.8), (4.12), and (4.33)**

*Proof of (4.8).* We decompose

$$\begin{aligned} \Psi_\alpha(t, z; \xi) &= \int_0^t \int_{\mathbb{R}} \left( e^{iu\xi} - 1 - iu\xi 1_{|u| \leq t^{1/\alpha}} \right) \mu^{(\alpha)}(\kappa_s(z); du) ds \\ &\quad + i\xi \int_0^t \int_{s^{1/\alpha} < |u| \leq t^{1/\alpha}} u \mu^{(\alpha)}(\kappa_s(z); du) ds. \end{aligned}$$

Since the density of  $\mu^{(\alpha)}(\kappa_s(z); du)$  is a homogeneous function of  $u$  of the order  $-(\alpha + 1)$ , changing the variables  $v = ut^{-1/\alpha}$  we get

$$\begin{aligned} &\int_0^t \int_{\mathbb{R}} \left( e^{iu\xi} - 1 - iu\xi 1_{|u| \leq t^{1/\alpha}} \right) \mu^{(\alpha)}(\kappa_s(z); du) ds \\ &= \frac{1}{t} \int_0^t \int_{\mathbb{R}} \left( e^{ivt^{1/\alpha}\xi} - 1 - ivt^{1/\alpha}\xi 1_{|v| \leq 1} \right) \mu^{(\alpha)}(\kappa_s(z); dv) ds = \Psi_\alpha(\tilde{\lambda}_t(z), \tilde{\rho}_t(z), 0)(t^{1/\alpha}\xi), \end{aligned}$$

see (3.7) for the definition of  $\Psi_\alpha^{\lambda,\rho,v}$ . On the other hand, we have straightforwardly

$$\int_0^t \int_{s^{1/\alpha} < |u| \leq t^{1/\alpha}} u \mu^{(\alpha)}(\kappa_s(z); du) ds = \int_0^t v(\kappa_s(z)) \int_{s^{1/\alpha}}^{t^{1/\alpha}} r \frac{dr}{r^{\alpha+1}} ds = t^{1/\alpha} \tilde{v}_t(z),$$

see Section 4.1 for the definition of  $\tilde{v}_t$ . Thus

$$\Psi_\alpha(t, z; \xi) = i\xi t^{1/\alpha} \tilde{v}_t(z) + \Psi_\alpha(\tilde{\lambda}_t(z), \tilde{\rho}_t(z), 0)(t^{1/\alpha}\xi) = \Psi_\alpha(\tilde{\lambda}_t(z), \tilde{\rho}_t(z), \tilde{v}_t(z))(t^{1/\alpha}\xi).$$

Therefore

$$\begin{aligned} h^{t,z}(w) &= \frac{1}{2\pi} \int_{\mathbb{R}} \exp[-iw\xi + \Psi_\alpha(t, z; \xi)] d\xi \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \exp\left[-iw\xi + \Psi_\alpha(\tilde{\lambda}_t(z), \tilde{\rho}_t(z), \tilde{v}_t(z))(t^{1/\alpha}\xi)\right] d\xi = t^{-1/\alpha} g(\tilde{\lambda}_t(z), \tilde{\rho}_t(z), \tilde{v}_t(z)) \left(\frac{w}{t^{1/\alpha}}\right), \end{aligned}$$

which yields (4.11).  $\square$

*Proof of (4.12).* Denote

$$\psi_\alpha(t, z; \xi) = \partial_t \Psi_\alpha(t, z; \xi) = \int_{\mathbb{R}} \left( e^{iu\xi} - 1 - iu\xi 1_{|u| \leq t^{1/\alpha}} \right) \mu^{(\alpha)}(\kappa_t(z); du).$$

It is easy to show that  $|\psi_\alpha(t, z; \xi)| \leq C(1 + \xi^2)$ . On the other hand, similarly to (A.26), we have that for any  $0 < \tau < T$  there exist constants  $c_1, c_2$  such that

$$\operatorname{Re} \Psi_\alpha(t, z; \xi) \leq -c_1 |\xi|^\alpha + c_2, \quad t \in [\tau, T].$$

Then the dominated convergence gives

$$\begin{aligned} \partial_t h^{t,z}(w - x) &= \lim_{R \rightarrow \infty} \frac{1}{2\pi} \int_{-R}^R \psi_\alpha(t, z; \xi) e^{-iw\xi + ix\xi + \Psi_\alpha(t, z; \xi)} d\xi \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \psi_\alpha(t, z; \xi) e^{-iw\xi + ix\xi + \Psi_\alpha(t, z; \xi)} d\xi, \end{aligned}$$

for  $t \in [\tau, T]$ . Repeating the same argument, we get

$$\begin{aligned} \tilde{L}_x^{(\alpha),z,t} h^{t,z}(w-x) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iw\xi + \Psi_\alpha(t,z;\xi)} \left( \tilde{L}_x^{(\alpha),z,t} e^{ix\xi} \right) d\xi \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iw\xi + \Psi_\alpha(t,z;\xi)} \psi_\alpha(t,z;\xi) e^{ix\xi} d\xi, \end{aligned}$$

$t \in [\tau, T]$ , which proves (4.12) for these values of  $t$ . Since  $0 < \tau < T$  are arbitrary, this completes the proof.  $\square$

*Proof of (4.33).* Denote

$$\overleftarrow{v}_t(x) = \int_0^t v(\chi_s(x)) W_\alpha(t; t-s) ds.$$

We have

$$\begin{aligned} p_t^0(x,y) - p_t^{\text{main}}(x,y) &= \left( \frac{1}{t^{1/\alpha}} g(\tilde{\lambda}_t(y), \tilde{\rho}_t(y), \tilde{v}_t(y)) \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) - \frac{1}{t^{1/\alpha}} g(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x)) \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) \right) \\ &+ \left( \frac{1}{t^{1/\alpha}} g(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x)) \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) - \frac{1}{t^{1/\alpha}} g(\lambda_t(x), \rho_t(x), v_t(x)) \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) \right) \\ &=: R_t^1(x,y) + R_t^2(x,y). \end{aligned}$$

Note that

$$g^{(\lambda,\rho,v)}(w) = g^{(\lambda,\rho,0)}(w-v).$$

Then by (A.24), (A.25), and (A.30) for any  $0 < \lambda_{\min} \leq \lambda_{\max}$ ,  $R > 0$  there exists  $C$  such that for any  $\lambda_1, \lambda_2 \in [\lambda_{\min}, \lambda_{\max}]$ ,  $\rho_1, \rho_2 \in [-1, 1]$ ,  $v_1, v_2 \in [-R, R]$ , and  $x \in \mathbb{R}$

$$|g^{(\lambda_1,\rho_1,v_1)}(x) - g^{(\lambda_2,\rho_2,v_2)}(x)| \leq C \left( |\lambda_1 - \lambda_2| + |\rho_1 - \rho_2| + |v_1 - v_2| \right) G^{(\alpha)}(x). \quad (\text{A.35})$$

We have

$$\begin{aligned} |\tilde{\lambda}_t(y) - \lambda_t(x)| &= \left| \frac{1}{t} \int_0^t \lambda(\kappa_\tau(y)) d\tau - \frac{1}{t} \int_0^t \lambda(\chi_s(x)) ds \right| \\ &= \left| \frac{1}{t} \int_0^t \left( \lambda(\kappa_{t-s}(y)) - \lambda(\chi_s(x)) \right) ds \right|, \end{aligned}$$

in the last identity we changed the variable  $\tau = t - s$ . By (A.17),

$$|\kappa_{t-s}(y) - \chi_s(x)| \leq C|y - \chi_t(x)| + Ct^{1/\alpha}, \quad s \in [0, t].$$

Since function  $\lambda(\cdot)$  is  $\zeta$ -Hölder continuous and bounded, this gives

$$|\tilde{\lambda}_t(y) - \lambda_t(x)| \leq C \left( t^{\delta_\zeta} + |y - \chi_t(x)|^\zeta \wedge 1 \right).$$

Similarly,

$$|\tilde{\rho}_t(y) - \rho_t(x)| \leq C \left( t^{\delta_\zeta} + |y - \chi_t(x)|^\zeta \wedge 1 \right).$$

Finally,

$$\begin{aligned} |\tilde{v}_t(y) - \overleftarrow{v}_t(x)| &= \left| \int_0^t v(\kappa_\tau(y)) W_\alpha(t; \tau) d\tau - \int_0^t v(\chi_s(x)) W_\alpha(t; t-s) ds \right| \\ &\leq \int_0^t |v(\kappa_{t-s}(y)) - v(\chi_s(x))| W_\alpha(t; t-s) ds \leq C \left( t^{\delta_\zeta} + |y - \chi_t(x)|^\zeta \wedge 1 \right). \end{aligned}$$



Thus by (A.35)

$$|R_t^1(x, y)| \leq C \left( t^{\delta\zeta} + |y - \chi_t(x)|^\zeta \wedge 1 \right) \left( \frac{1}{t^{1/\alpha}} G^{(\alpha)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) \right).$$

By (A.17) and (A.20), (A.21),

$$\frac{1}{t^{1/\alpha}} G^{(\alpha)} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) \leq C \frac{1}{t^{1/\alpha}} G^{(\alpha)} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) = G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y).$$

This gives finally

$$|R_t^1(x, y)| \leq C \left( t^{\delta\zeta} + |y - \chi_t(x)|^\zeta \wedge 1 \right) G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y) \leq C t^{\delta\zeta} G_t^{(\alpha, \alpha - \zeta, \alpha)}(\chi_t(x), y).$$

Next, we decompose

$$\begin{aligned} R_t^2(x, y) &= \left( \frac{1}{t^{1/\alpha}} g^{(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x))} \left( \frac{\kappa_t(y) - x}{t^{1/\alpha}} \right) - \frac{1}{t^{1/\alpha}} g^{(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x))} \left( \frac{y - \chi_t^t(x)}{t^{1/\alpha}} \right) \right) \\ &+ \left( \frac{1}{t^{1/\alpha}} g^{(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x))} \left( \frac{y - \chi_t^t(x)}{t^{1/\alpha}} \right) - \frac{1}{t^{1/\alpha}} g^{(\lambda_t(x), \rho_t(x), v_t(x))} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) \right) \\ &=: R_t^{2,1}(x, y) + R_t^{2,2}(x, y). \end{aligned}$$

We have

$$\begin{aligned} |R_t^{2,1}(x, y)| &= \left| \int_0^t \partial_s \left( \frac{1}{t^{1/\alpha}} g^{(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x))} \left( \frac{\kappa_{t-s}(y) - \chi_s^t(x)}{t^{1/\alpha}} \right) \right) ds \right| \\ &= \left| \frac{1}{t^{1/\alpha}} \int_0^t (g^{(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x))})' \left( \frac{\kappa_{t-s}(y) - \chi_s^t(x)}{t^{1/\alpha}} \right) \frac{B_{t-s}(\kappa_{t-s}(y)) - B_{t-s}(\chi_s^t(x))}{t^{1/\alpha}} ds \right|. \end{aligned}$$

Using (A.30), (3.10), (A.11), and (A.16), we get

$$\begin{aligned} &\left| \frac{1}{t^{1/\alpha}} (g^{(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x))})' \left( \frac{\kappa_{t-s}(y) - \chi_s^t(x)}{t^{1/\alpha}} \right) \frac{B_{t-s}(\kappa_{t-s}(y)) - B_{t-s}(\chi_s^t(x))}{t^{1/\alpha}} \right| \\ &\leq \frac{C}{t^{1/\alpha}} G^{(\alpha+1)} \left( \frac{\kappa_{t-s}(y) - \chi_s^t(x)}{t^{1/\alpha}} \right) \text{Lip}(B_{t-s}) \frac{|\kappa_{t-s}(y) - \chi_s^t(x)|}{t^{1/\alpha}} \\ &\leq \frac{C(t-s)^{-1+\delta}}{t^{1/\alpha}} G^{(\alpha)} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right), \end{aligned}$$

which gives

$$|R_t^{2,1}(x, y)| \leq C t^\delta G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y).$$

Finally, we have

$$\begin{aligned} g^{(\lambda_t(x), \rho_t(x), \overleftarrow{v}_t(x))} \left( \frac{y - \chi_t^t(x)}{t^{1/\alpha}} \right) &= g^{(\lambda_t(x), \rho_t(x), 0)} \left( \frac{y - \chi_t^t(x) - t^{1/\alpha} \overleftarrow{v}_t(x)}{t^{1/\alpha}} \right), \\ g^{(\lambda_t(x), \rho_t(x), v_t(x))} \left( \frac{y - \chi_t(x)}{t^{1/\alpha}} \right) &= g^{(\lambda_t(x), \rho_t(x), 0)} \left( \frac{y - \chi_t(x) - t^{1/\alpha} v_t(x)}{t^{1/\alpha}} \right), \end{aligned}$$

and by (A.12)

$$\begin{aligned} \chi_t^t(x) - \chi_t(x) &= t^{1/\alpha} \int_0^t v(\chi_r(x)) \left( W_\alpha(t; r) - W_\alpha(t; t-r) \right) dr + Q_{t,t}(x) \\ &= t^{1/\alpha} v_t(x) - t^{1/\alpha} \overleftarrow{v}_t(x) + Q_{t,t}(x). \end{aligned}$$

Then by (A.13)

$$\left| (\chi_t^t(x) + t^{1/\alpha} \overleftarrow{v}_t(x)) - (\chi_t(x) + t^{1/\alpha} v_t(x)) \right| \leq Ct^{1/\alpha+\delta}, \quad (\text{A.36})$$

and similarly to the above estimates, using (A.30), (A.11), and (A.16), we get

$$|R_t^{2,2}(x, y)| \leq Ct^\delta G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y).$$

That is,

$$\begin{aligned} |p_t^0(x, y) - \tilde{p}_t(x, y)| &\leq |R_t^1(x, y)| + |R_t^{2,1}(x, y)| + |R_t^{2,2}(x, y)| \\ &\leq Ct^\delta G_t^{(\alpha, \alpha-\zeta, \alpha)}(\chi_t(x), y) + Ct^\delta G_t^{(\alpha, \alpha, \alpha)}(\chi_t(x), y), \end{aligned}$$

which is just (4.33). □

### A.5 Properties of the kernels $G^{(\alpha, \beta, \gamma)}$

*Proofs of (4.34) and (4.24).* We have

$$G_t^{(\alpha, \beta, \gamma)}(x, y) = G_t^{(\alpha, \beta, \beta)}(x, y) + t^{\beta/\alpha} \left[ |y-x|^{-\gamma-1} - |y-x|^{-\beta-1} \right] \mathbf{1}_{|y-x|>1},$$

and

$$G_t^{(\alpha, \beta, \beta)}(x, y) \leq t^{-1/\alpha} G^{(\beta)}\left(\frac{y-x}{t^{1/\alpha}}\right).$$

Since  $G^{(\beta)} \in L_1(\mathbb{R})$ , we get (4.34):

$$\int_{\mathbb{R}} G_t^{(\alpha, \beta, \gamma)}(x, y) dx \leq C, \quad \int_{\mathbb{R}} G_t^{(\alpha, \beta, \gamma)}(x, y) dy \leq C, \quad t \in (0, T]. \quad (\text{A.37})$$

Next, the kernel  $G_t^{(\alpha, \beta, \gamma)}$  depends only on  $(y-x)/t^{1/\alpha}$ :

$$G_t^{(\alpha, \beta, \gamma)}(x, y) = F_t^{(\alpha, \beta, \gamma)}\left(\frac{y-x}{t^{1/\alpha}}\right). \quad (\text{A.38})$$

It is straightforward to verify that the corresponding function

$$F_t^{(\alpha, \beta, \gamma)}(x) = \begin{cases} t^{-1/\alpha}, & |x| \leq (1 \wedge t^{-1/\alpha}), \\ t^{-1/\alpha} |x|^{-\beta-1}, & (1 \wedge t^{-1/\alpha}) < |x| \leq t^{-1/\alpha}, \\ t^{(\beta-\gamma-1)/\alpha} |x|^{-\gamma-1}, & |x| > t^{-1/\alpha} \end{cases} \quad (\text{A.39})$$

satisfies the analogues of (A.20), (A.21):

$$F_t^{(\alpha, \beta, \gamma)}(x+v) \leq CF_t^{(\alpha, \beta, \gamma)}(x), \quad |v| \leq 1, \quad (\text{A.40})$$

and for any  $c > 0$  there exists  $C$  such that

$$F_t^{(\alpha, \beta, \gamma)}(cx) \leq CF_t^{(\alpha, \beta, \gamma)}(x) \quad (\text{A.41})$$

(the constants  $C$  can be chosen the same for all  $t \in (0, T]$ ). Using (A.17) with  $s = t$  and (A.40), (A.41), we get

$$G_t^{(\alpha, \beta, \gamma)}(x, \kappa_t(y)) \leq CG_t^{(\alpha, \beta, \gamma)}(\chi_t(x), y). \quad (\text{A.42})$$

Combined with (4.34) this gives (4.24). □

We say that a non-negative kernel  $H_t(x, y)$  has a *sub-convolution property*, if for every  $T > 0$  there exists a constant  $C$  such that

$$(H_{t-s} * H_s)(x, y) \leq CH_t(x, y), \quad t \in (0, T], \quad s \in (0, t), \quad x, y \in \mathbb{R}. \quad (\text{A.43})$$

**Proposition A.9.** For arbitrary  $\alpha, \beta, \gamma > 0$ , the kernel  $G_t^{(\alpha, \beta, \gamma)}(x, y)$  has a sub-convolution property.

*Proof.* We have for  $H_t(x, y) = G_t^{(\alpha, \beta, \gamma)}(x, y)$

$$\sup_{x, y} H_{t-s}(x, y) \leq \left(\frac{2}{t}\right)^{1/\alpha}$$

for  $s < t/2$ , and

$$\sup_{x, y} H_s(x, y) \leq \left(\frac{2}{t}\right)^{1/\alpha}$$

otherwise. In both these cases we have by (A.37)

$$\sup_{x, y} (H_{t-s} * H_s)(x, y) = \int_{\mathbb{R}} H_{t-s}(x, z) H_s(z, y) dz \leq Ct^{-1/\alpha}.$$

This proves (A.43) for  $x, y$  such that  $|x - y| \leq 2t^{1/\alpha}$ . Next,  $H_t(x, y)$  is positive and thus

$$\begin{aligned} (H_{t-s} * H_s)(x, y) &= \int_{\mathbb{R}} H_{t-s}(x, z) H_s(z, y) dz \\ &\leq \int_{|x-z| > |x-y|/2} H_{t-s}(x, z) H_s(z, y) dz + \int_{|y-z| > |x-y|/2} H_{t-s}(x, z) H_s(z, y) dz. \end{aligned}$$

The function  $F_t^{(\alpha, \beta, \gamma)}(x)$  in the presentation (A.38) of  $H_t(x, y) = G_t^{(\alpha, \beta, \gamma)}(x, y)$ , for a fixed  $t$ , depends only on  $|x|$ , and is a non-increasing function of  $|x|$ . Hence

$$H_{t-s}(x, z) \leq H_{t-s}(x/2, y/2), \quad \text{for } |x - z| > |x - y|/2 \Leftrightarrow |x - z| > \left|\frac{x}{2} - \frac{y}{2}\right|,$$

and

$$H_s(z, y) \leq H_s(x/2, y/2), \quad \text{for } |y - z| > |x - y|/2 \Leftrightarrow |y - z| > \left|\frac{x}{2} - \frac{y}{2}\right|.$$

Therefore by (A.37)

$$\begin{aligned} (H_{t-s} * H_s)(x, y) &\leq \int_{|x-z| > |x-y|/2} H_{t-s}(x/2, y/2) H_s(z, y) dz \\ &\quad + \int_{|y-z| > |x-y|/2} H_{t-s}(x, z) H_s(x/2, y/2) dz \\ &\leq C \left( H_{t-s}(x/2, y/2) + H_s(x/2, y/2) \right). \end{aligned}$$

Then for  $|x - y| \geq 2t^{1/\alpha}$  we deduce

$$(H_{t-s} * H_s)(x, y) \leq C((t-s)^{\beta/\alpha} + s^{\beta/\alpha}) F^{(\beta, \gamma)}\left(\frac{y-x}{2}\right), \quad F^{(\beta, \gamma)}(x) = \begin{cases} |x|^{-\beta-1}, & |x| \leq 1; \\ |x|^{-\gamma-1}, & |x| > 1. \end{cases}$$

Clearly,  $(t-s)^{\beta/\alpha} + s^{\beta/\alpha} \leq 2t^{\beta/\alpha}$  and

$$F^{(\beta, \gamma)}(x/2) \leq CF^{(\beta, \gamma)}(x),$$

which completes the proof of (A.43) for  $|x - y| \geq 2t^{1/\alpha}$ . □

**Proposition A.10.** The kernel

$$H_t(x, y) = G_t^{(\alpha, \beta, \gamma)}(x, \kappa_t(y))$$

has the sub-convolution property.

*Proof.* Using (A.42) and Proposition A.9, we get

$$\begin{aligned} (H_s * H_{t-s})(x, y) &\leq C \int_{\mathbb{R}} G_s^{(\alpha, \beta, \gamma)}(\chi_s(x), y') G_{t-s}^{(\alpha, \beta, \gamma)}(y', \kappa_{t-s}(y)) dy' \\ &\leq C G_t^{(\alpha, \beta, \gamma)}(\chi_s(x), \kappa_{t-s}(y)). \end{aligned}$$

Using (A.17), (A.40), and (A.41), we get similarly to (A.42)

$$G_t^{(\alpha, \beta, \gamma)}(\chi_s^s(x), \kappa_{t-s}(y)) \leq C G_t^{(\alpha, \beta, \gamma)}(x, \kappa_t(y)) = C H_t(x, y),$$

which completes the proof.  $\square$

## References

- [1] Y. Aït-Sahalia and J. Jacod (2007). Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.* **35**, no. 1, 355–392. MR-2332279
- [2] B. Böttcher, R. Schilling, and J. Wang (2013). Lévy matters. III, volume 2099 of Lecture Notes in Mathematics. Springer, Cham. *Lévy-type processes: construction, approximation and sample path properties*, With a short biography of Paul Lévy by Jean Jacod, Lévy Matters. MR-3156646
- [3] P. D. Ditlevsen (1999). Observation of  $\alpha$ -stable noise induced millennial climate changes from an ice record. *Geophysical Research Letters* **26**, no. 10, 1441–1444.
- [4] S. N. Ethier, T. G. Kurtz (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. MR-0838085
- [5] A. Friedman (1964). *Partial Differential Equations of Parabolic Type*. Prentice Hall, New York. MR-0181836
- [6] Iu. Ganychenko, V. Knopova, and A. Kulik (2015). Accuracy of discrete approximation for integral functionals of Markov processes, *Modern Stochastics: Theory and Applications* **2**, no. 4, 401–420. MR-3456146
- [7] B. Grigelionis (1968). On a Markov property of Markov processes. (Russian) *Liet. Matem. Rink.* **8**, no. 3, 489–502. MR-0251810
- [8] L. Huang (2015). Density estimates for SDEs driven by tempered stable processes. arXiv:1504.04183 MR-3262494
- [9] V. Knopova and A. Kulik (2013). Intrinsic small time estimates for distribution densities of Lévy processes. *Random Op. Stoch. Eq.* **21**, no. 4, 321–344. MR-3139314
- [10] V. Knopova and A. Kulik (2018). Parametrix construction of the transition probability density of the solution to an SDE driven by  $\alpha$ -stable noise. *Ann. Inst. Henri Poincaré* **54**, no. 1, 100–140. MR-3765882
- [11] V. Knopova, A. Kulik, and R. Schilling. On the construction of a general stable-like Markov process (manuscript)
- [12] A. N. Kochubei (1989). Parabolic pseudo-differential equations, hypersingular integrals and Markov processes. *Math. USSR Izvestija* **33**, 233–259. MR-0972089
- [13] A. N. Kochubei, V. P. Knopova, and A. M. Kulik. *Parametrix methods for equations with fractional Laplacians*, to appear in Handbook Fractional Calculus with Applications. Springer. MR-3965398
- [14] V. N. Kolokoltsov (2000). Symmetric stable laws and stable-like jump-diffusions. *Proc. London Math. Soc.* **80**, 725–768. MR-1744782
- [15] V. N. Kolokoltsov (2018). *Differential Equations on Measures and Functional Spaces*. Birkhäuser.
- [16] T. Komatsu (1984). On the martingale problem for generators of stable processes with perturbations. Markov processes associated with certain integro-differential operators. *Osaka J. Math.* **21**, 113–132 MR-0736974
- [17] T. Komatsu (1984). Pseudo-differential operators and Markov processes. *J. Math. Soc. Japan* **36**, no. 3, 387–418. MR-0746702

- [18] T. Kulczycki, M. Ryznar, and P. Sztonyk (2018). Strong Feller property for SDEs driven by multiplicative cylindrical stable noise. arXiv:1811.05960.
- [19] A. Kulik (2019). On weak uniqueness and distributional properties of a solution to an SDE with  $\alpha$ -stable noise. *Stoch. Proc. Appl.*, **129**, no. 2, 473–506. MR-3907007
- [20] F. Kühn (2017). *Lévy-Type Processes: Moments, Construction and Heat Kernel Estimates*. Lévy Matters, **5**, Springer. MR-3701414
- [21] J. Rosinski (2007). Tempering stable processes, *Stoch. Proc. Appl.* **117**, 677–707. MR-2327834
- [22] I.M. Sokolov, A.V. Chechkin, and J. Klafter (2004). Fractional diffusion equation for a power-law-truncated Lévy process, *Physica A* **336**, 245–251.
- [23] K. Szczypkowski (2018). Fundamental solution for super-critical non-symmetric Lévy-type operators, arXiv:1807.04257.
- [24] H. Tanaka, M. Tsuchiya, and S. Watanabe (1974). Perturbation of drift-type for Lévy processes, *J. Math. Kyoto Univ.* **14**, 73–92. MR-0368146

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS<sup>1</sup>)
- Easy interface (EJMS<sup>2</sup>)

### Economical model of EJP-ECP

- Non profit, sponsored by IMS<sup>3</sup>, BS<sup>4</sup>, ProjectEuclid<sup>5</sup>
- Purely electronic

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>2</sup>EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

<sup>3</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>4</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>5</sup>Project Euclid: <https://projecteuclid.org/>

<sup>6</sup>IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>