

## The cutoff phenomenon in total variation for nonlinear Langevin systems with small layered stable noise\*

Gerardo Barrera<sup>†</sup>    Michael A. Högele<sup>‡</sup>    Juan Carlos Pardo<sup>§</sup>

### Abstract

This paper provides an extended case study of the cutoff phenomenon for a prototypical class of nonlinear Langevin systems with a single stable state perturbed by an additive pure jump Lévy noise of small amplitude  $\varepsilon > 0$ , where the driving noise process is of layered stable type. Under a drift coercivity condition the associated family of processes  $X^\varepsilon$  turns out to be exponentially ergodic with equilibrium distribution  $\mu^\varepsilon$  in total variation distance which extends a result from [60] to arbitrary polynomial moments.

The main results establish the cutoff phenomenon with respect to the total variation, under a sufficient smoothing condition of Blumenthal-Gettoor index  $\alpha > \frac{3}{2}$ . That is to say, in this setting we identify a deterministic time scale  $t_\varepsilon^{\text{cut}}$  satisfying  $t_\varepsilon^{\text{cut}} \rightarrow \infty$ , as  $\varepsilon \rightarrow 0$ , and a respective time window,  $t_\varepsilon^{\text{cut}} \pm o(t_\varepsilon^{\text{cut}})$ , during which the total variation distance between the current state and its equilibrium  $\mu^\varepsilon$  essentially collapses as  $\varepsilon$  tends to zero. In addition, we extend the dynamical characterization under which the latter phenomenon can be described by the convergence of such distance to a unique profile function first established in [9] to the Lévy case for nonlinear drift. This leads to sufficient conditions, which can be verified in examples, such as gradient systems subject to small symmetric  $\alpha$ -stable noise for  $\alpha > \frac{3}{2}$ . The proof techniques differ completely from the Gaussian case due to the absence of a respective Girsanov transform which couples the nonlinear equation and the linear approximation asymptotically even for short times.

**Keywords:** cutoff phenomenon; abrupt thermalization; exponential ergodicity; stable Lévy processes; local limit theorem; nonlinear coupling; short coupling; total variation distance; counterexample to Slutsky's lemma in total variation; Hölder continuity of the characteristic exponent.

**MSC2020 subject classifications:** 37A25; 37A30; 60F05; 60G51; 60G52; 65C30.

---

\*Supported by the Academy of Finland, Facultad de Ciencias at Universidad de los Andes and CONACyT-MEXICO.

<sup>†</sup>Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, 00014, Finland.  
E-mail: [gerardo.barreravargas@helsinki.fi](mailto:gerardo.barreravargas@helsinki.fi)

<sup>‡</sup>Departamento de Matemáticas, Facultad de Ciencias, Universidad de los Andes. Bogotá, Colombia.  
E-mail: [ma.hoegele@uniandes.edu.co](mailto:ma.hoegele@uniandes.edu.co)

<sup>§</sup>CIMAT, Jalisco S/N, Valenciana, CP 36240. Guanajuato, Guanajuato, México.  
E-mail: [jcpardo@cimat.mx](mailto:jcpardo@cimat.mx)

## 1 Exposition

### 1.1 Introduction

Roughly speaking the term *cutoff phenomenon* with respect to a distance  $d_1$  refers to the following asymptotic dynamics: consider the setting of a parametrized family of stochastic processes  $(X^\varepsilon)_{\varepsilon>0}$ ,  $X^\varepsilon = (X_t^\varepsilon)_{t \geq 0}$ , such that for each  $\varepsilon > 0$  the process  $X^\varepsilon$  has a unique limiting distribution  $\mu^\varepsilon$ . Then – as  $\varepsilon$  decreases to 0 – the function  $t \mapsto d_\varepsilon(X_t^\varepsilon, \mu^\varepsilon)$  given by a suitably renormalized distance  $d_\varepsilon$  (of  $d_1$ ) between the law of  $X_t^\varepsilon$  and the corresponding limiting distribution  $\mu^\varepsilon$  essentially resembles the step function  $t \mapsto \text{diam} \cdot \mathbf{1}_{[0, t_\varepsilon^{\text{cut}}]}(t)$ . This function descends from the value  $\text{diam} \in (0, \infty)$  to the value 0, at a deterministic cutoff time scale  $t_\varepsilon^{\text{cut}}$ , which tends to  $\infty$  as  $\varepsilon \rightarrow 0$ , where  $\text{diam} = \limsup_{\varepsilon \rightarrow 0} \text{diameter}(d_\varepsilon)$  in the respective domain of probability distributions over the state space. In other words, there exist positive deterministic functions  $\varepsilon \mapsto t_\varepsilon^{\text{cut}}$  and  $\varepsilon \mapsto \mathfrak{w}_\varepsilon^{\text{cut}}$  satisfying  $t_\varepsilon^{\text{cut}} \rightarrow \infty$  and  $\mathfrak{w}_\varepsilon^{\text{cut}} \ll t_\varepsilon^{\text{cut}}$  such that on  $(t_\varepsilon^{\text{cut}} - \mathfrak{w}_\varepsilon^{\text{cut}}, t_\varepsilon^{\text{cut}} + \mathfrak{w}_\varepsilon^{\text{cut}})$  the transition from  $\text{diam}$  to 0 is bound to happen. In general, this transition may depend on subsequences  $\varepsilon_j \rightarrow 0$  as  $j \rightarrow \infty$ . In certain situations, a proper limit can be taken, and the limiting function gives rise to a so-called cutoff profile function connecting the asymptotic values  $\text{diam}$  and 0 smoothly.

This abrupt convergence phenomenon was first described by Aldous and Diaconis [2] in the early eighties to conceptualize the collapse of the total variation distance between Markov chain marginals related to card shuffling to its uniform limiting distribution. Since then, this behavior has been studied by numerous authors and in different – mainly discrete – settings. For instance we refer to Diaconis [30], Martínez and Ycart [54] and Levin et al. [53] for the Markov chain setting, Chen and Saloff-Coste [25] considered some ergodic Markov processes, Lachaud [50] and Barrera [5] for the case of the Ornstein-Uhlenbeck processes driven by a Brownian motion, to name but a few. Further standard texts on the cutoff phenomenon include [1, 3, 13, 15, 11, 14, 17, 18, 25, 29, 31, 32, 51, 52, 53, 55, 74, 79] and the references therein. The newest developments in this active field of research are found in the recent publications [16, 19, 20, 21, 24, 49, 40, 41].

This article provides a case study on the cutoff phenomenon in the (unnormalized) total variation distance for the strong solution process  $X^\varepsilon$  of a class of stochastic differential equations with nonlinear coercive vector field  $-b$  with a non-degenerate stable state 0 subject to an additive pure jump Lévy process  $L$  at  $\varepsilon$ -small amplitude

$$\begin{cases} dX_t^\varepsilon = -b(X_t^\varepsilon)dt + \varepsilon dL_t & \text{for } t \geq 0, \\ X_0^\varepsilon = x \in \mathbb{R}^d. \end{cases} \quad (1.1)$$

Similar – and in some sense simpler – settings have been studied before: the case of *nonlinear, coercive* vector fields  $(-b)$  subject to *Brownian perturbation*  $L = W$  with respect to the *total variation* [8, 9] and two cases of *linear, asymptotically exponentially stable* drifts  $-b = -Q$  – that is, eigenvalues have negative real part, but the matrix is not necessarily coercive, see [75] – subject to *pure-jump Lévy noises*  $L$  [6, 10] in the *total variation* and the *Wasserstein distance*, respectively. This paper yields the first results on the cutoff phenomenon for *nonlinear coercive, pure-jump Lévy* SDEs in the total variation distance, which is fraught with technical difficulties:

- (a) It inherits the regularity issues from the linear case [10] due to the total variation distance.

- (b) It earns additional challenges due to the nonlinearity. In particular, there is a gap in the literature concerning available (short-time) coupling results between the solution of Lévy SDEs with the nonlinear vector fields and its (linear) Ornstein-Uhlenbeck approximation.

The regularity issue (a) is overcome by the careful choice of the setting of a class of locally layered stable noise processes, by which we generalize the notion of layered stable processes – introduced by Houdré and Kawai [42] – and the equator condition inspired by [71]. Regularity results for densities of SDEs which turn out to be crucial for results in the total variation distance have been extensively studied for instance in [27, 36, 43, 46, 67].

The nonlinear coupling problem (b) is essentially reduced to the control of two partial errors of a different nature addressed in Proposition 2 and Proposition 3. The first error, which is dominated in the statement of Proposition 2 represents the crucial part of the proof of the main results. It directly compares the nonlinear process  $X^\varepsilon$  with its linear inhomogeneous Ornstein-Uhlenbeck approximation for short times. While there are very recent short-time couplings for SDEs with different (nonlinear) drift under a Brownian driver (see Eberle and Zimmer [34]), to our knowledge the literature on respective pure jump counterparts is virtually nonexistent. In order to obtain short-time coupling between the linear and the nonlinear vector field, we use Plancherel’s theorem, and appropriate differential inequalities for the characteristic function of a strongly localized version of  $X^\varepsilon$  for Blumenthal-Gettoor index  $\alpha > 3/2$ . To the best of our effort it seems hard to derive with this technique the correct (exponential) integrability of the tails of the characteristic function – even in the linear, scalar Gaussian case – and at the same time it is unclear how to relax this condition. The same sort of technical difficulties concerning the Fourier approach arises in condition (a) p. 345 of [36]. The second error consists of the total variation distance between the short time linear inhomogeneous Ornstein-Uhlenbeck (Freidlin-Wentzell first order) approximation under linear and nonlinear initial conditions. A slight extension of Theorem 3.1 in [42] provides a stable local limit theorem on the short-range behavior, which allows for an appropriate coupling in the proof of Proposition 3.

The difficulty of the nonlinear case studied in this article can be informally understood as follows. In the linear case  $b(x) = -Qx$ , it is well-known that by the variation-of-constants formula  $X_t^\varepsilon$  can be written as the sum of the deterministic matrix exponential dynamics plus the respective stochastic convolution. Since the total variation distance is well-behaved under deterministic and mutually independent components, it can be dominated without too much effort in the linear case. This program was carried out in [10]. In the nonlinear, additive noise case  $X_t^\varepsilon$  can be written analogously, but it exhibits an additional error term. That is,  $X^\varepsilon$  is given as the sum of the nonlinear deterministic dynamics, its stochastic (nonlinear) convolution with the noise and an additional random term representing the (implicit) nonlinear residual of the noise, which is neither deterministic nor independent from the noise convolution and therefore not easily dominated in total variation. Beyond that, the aforementioned random residual term turns out to be a challenge since there is no analogue of Slutsky’s lemma for the total variation distance even in the case of smooth densities. For the sake of completeness and since we are not aware of a reference literature, a counterexample is given in Subsubsection 1.3.5. On a more abstract level, the additional difficulties encountered are illustrated for the Wasserstein upper bounds of the total variation which require additional density gradient estimates (see Theorem 2.1 in [23]).

Our results cover the important examples of overdamped gradient systems, such as the Fermi-Pasta-Ulam-Tsingou potential, perturbed by pure jump Lévy processes with Blumenthal-Gettoor index  $\alpha > 3/2$  in the sense of Definition 1.3 and 1.4, such as symmetric

$\alpha$ -stable processes, symmetric tempered  $\alpha$ -stable processes in Rosiński [62] and the symmetric Lamperti- $\alpha$ -stable process [22]. If – in addition – the limiting distributions turns out to be rotationally invariant, the existence of a cutoff profile is shown to be equivalent to a computational linear algebra eigenvector problem first established in [6] for the easier situation of the Wasserstein distance. This characterization is given as a specific orthogonality condition of the (generalized) eigenvectors of the linearization  $-Db(0)$  of  $-b$  in the stable state 0. It allows to carry over several results from the linear case under the Wasserstein distance in [6], to the case of a nonlinear vector field  $-b$  and the total variation distance. In physics terminology, our results can be restated that the existence of a cutoff profile is equivalent to the absence of non-normal growth effects in  $-Db(0)$  in the case of rotationally invariant limiting distributions in the nonlinear setting.

For a complete comparison of the different settings and results and in order to avoid a lengthy introduction, we refer to the following self-explanatory Table 1.1.

Settings	[8]	[9]	[10]	[6]	this article
<b>Dimension</b>	scalar	multivariate	multivariate	multivariate	<b>multivariate</b>
<b>Vector field</b>	nonlinear	nonlinear	linear	linear	<b>nonlinear</b>
<b>Fixed point stability</b>	strong coercivity	strong coercivity	neg. real parts of the eigenvalues	neg. real parts of the eigenvalues	<b>strong coercivity</b>
<b>Noise process</b>	Brownian motion	Brownian motion	pure jump Lévy	pure jump Lévy	<b>pure jump Lévy</b>
<b>Noise process degeneracy</b>	no	no	no	yes	<b>no</b>
<b>Restrictions on the noise</b>	none	none	finite log-moment + Hypothesis (H)	finite moment of order $\beta > 0$	<b>finite moment of order <math>\beta &gt; 0</math>, + strongly locally layered stable <math>\alpha \in (3/2, 2)</math></b>
<b>Limiting distribution</b>	explicitly known	abstract, expansions in $\varepsilon$ known [69, 57]	characteristic function known	d.n.a. due to shift linearity of Wasserstein dist.	<b>completely abstract</b>
Results	[8]	[9]	[10]	[6]	this article
<b>Distance</b>	total variation	total variation	total variation	rescaled Wasserstein	<b>total variation</b>
<b>Window cutoff</b>	yes	yes	yes	yes	<b>yes</b>
<b>Profile cutoff</b>	yes	dynamical characterization	dynamical characterization	dynamical characterization + normal growth characterization (general case)	<b>dynamical characterization + normal growth characterization (rot. inv. case)</b>
<b>Short time coupling</b>	Girsanov + Pinsker's inequality	Girsanov + Hellinger's inequality	Fourier inversion	does not apply	<b>Plancherel isometry of <math>L^2</math></b>

Table 1: Schematic overview

The nonlinear Wasserstein setting with results in the spirit of [6] are studied in the paper [7].

In the manuscript we prove several results of interest in its own right which to our knowledge have not been present in the literature: (1) In Theorem 1 we generalize the strong ergodicity result Theorem 4.1 in [60] from moments  $\beta \geq 2$  to any  $\beta > 0$ . The proof is given in Subsection D. (2) In Definition 1.3 we introduce the class of locally layered stable process, which are precisely the class of processes for which the short-range behavior in Theorem 3.1 in [42] remains valid. (3) In Proposition 7 we give an elementary proof of the local  $\beta$ -Hölder continuity of the characteristic exponents in case of  $\beta \in (0, 1]$ -moments in Subsection C.1. (4) We also provide a complete overview of the behavior of the estimates of matrix exponentials and related flows for an asymmetric matrix in Appendix A, since we are not aware of a reference in the literature.

The manuscript is organized in two large sections and an extended Appendix. The

first section lays out the setting, the main results formulated as Theorem 2 and 3, the examples and the skeleton of concluding steps in the proof of the main results, which boils down to the proofs of Proposition 1, 2, 3 and 4. The respective results are proven in the (correspondingly ordered) Subsection 2.1, 2.3, 2.2 and 2.4, respectively. Subsection 2.1 shows the cutoff result for the linear inhomogeneous Ornstein-Uhlenbeck process. Subsections 2.3 and 2.4 yield the coupling of the inhomogeneous Ornstein-Uhlenbeck and the nonlinear short-time coupling, which exhibits the core difficulties. The Appendix is divided in Section A, B, C and D in which several auxiliary results are shown as a by-product in its own right. Section A provides all necessary fine results on the deterministic dynamics. Appendix B yields a quantitative estimates of the Freidlin-Wentzell first order approximation. Appendix C gives several auxiliary technical results, some of which we have not been aware in the literature, such as the local  $\beta$ -Hölder-continuity of a Lévy process in the presence of arbitrary  $\beta$ -moments. Appendix D yields the proof of Theorem 1, which implies the exponential ergodicity of  $X^\varepsilon$  towards  $\mu^\varepsilon$ , which extends a result by [60] to the case of an arbitrary positive finite moment.

## 1.2 The setting and the main results

### 1.2.1 The deterministic dynamics $\varphi^x$

Let  $b \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$  be a vector field with  $b(0) = 0$  satisfying the following coercivity condition.

**Hypothesis 1** (Coercivity).

Assume that there exists a positive constant  $\delta$  such that

$$\langle b(x) - b(y), x - y \rangle \geq \delta |x - y|^2 \quad \text{for all } x, y \in \mathbb{R}^d, \quad (1.2)$$

where  $|\cdot|$  and  $\langle \cdot, \cdot \rangle$  denote the Euclidean norm and the standard inner product on  $\mathbb{R}^d$ , respectively.

In this manuscript we are interested in the stochastically perturbed analogue of the dynamical system given as the global solution flow  $(\varphi_t^x)_{t \geq 0}$  of the ordinary differential equation

$$\begin{cases} d\varphi_t^x = -b(\varphi_t^x)dt & \text{for any } t \geq 0, \\ \varphi_0^x = x \in \mathbb{R}^d. \end{cases} \quad (1.3)$$

It is well-known that Hypothesis 1 implies the well-posedness of (1.3), see for instance Subsection 2.1 in [9]. Furthermore, in our setting inequality (1.2) is equivalent to

$$\langle Db(x)y, y \rangle \geq \delta |y|^2 \quad \text{for all } x, y \in \mathbb{R}^d,$$

where  $Db(x)$  denotes the derivative of the vector field  $b$  at the point  $x$ . Moreover, since  $b(0) = 0$ , we have

$$\frac{d}{dt} |\varphi_t^x|^2 = -2 \langle \varphi_t^x, b(\varphi_t^x) \rangle \leq -2\delta |\varphi_t^x|^2 \quad \text{for any } t \geq 0.$$

As a consequence  $|\varphi_t^x| \leq e^{-\delta t} |x|$  for any  $t \geq 0$  and  $x \in \mathbb{R}^d$ , i.e. 0 is an asymptotically exponentially stable fixed point of (1.3). For our purposes, however, we need the precise description of the convergence to 0 in terms of the spectral decomposition of  $-Db(0)$ . This is the purpose of the following lemma which characterizes the asymptotics of  $\varphi_t^x$  as  $t$  tends to  $\infty$  and slightly refines the classical and well-known result by Hartman-Grobman [38, 39] under Hypothesis 1. This lemma turns out to be crucial for the precise shape of the cutoff time and time window.

**Lemma 1.1** (Hartman-Grobman).

Consider  $(\varphi_t^x)_{t \geq 0}$  defined by (1.3) under Hypothesis 1. Then for any  $x \in \mathbb{R}^d \setminus \{0\}$  there exist

- (i) positive constants  $\lambda := \lambda_x, \tau := \tau_x, \ell := \ell_x, m := m_x, \ell, m \in \{1, \dots, d\}$ ,
- (ii) angles  $\theta^1 := \theta_x^1, \dots, \theta^m := \theta_x^m \in [0, 2\pi)$ , where all angles  $\theta^k \in (0, 2\pi)$  come in pairs  $(\theta^{j^*}, \theta^{j^*+1}) = (\theta^{j^*}, 2\pi - \theta^{j^*})$  and
- (iii) linearly independent vectors  $v^1 := v_x^1, \dots, v^m := v_x^m$  in  $\mathbb{C}^d$  satisfying  $(v^{j^*}, v^{j^*+1}) = (v^{j^*}, \bar{v}^{j^*})$  whenever  $(\theta^{j^*}, \theta^{j^*+1}) = (\theta^{j^*}, 2\pi - \theta^{j^*})$ ,

such that

$$\lim_{t \rightarrow \infty} \left| \frac{e^{\lambda t}}{t^{\ell-1}} \cdot \varphi_{t+\tau}^x - \sum_{k=1}^m e^{i\theta_k t} v^k \right| = 0. \tag{1.4}$$

Moreover,

$$0 < \liminf_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{it\theta^k} v^k \right| \leq \limsup_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{it\theta^k} v^k \right| \leq \sum_{k=1}^m |v^k|. \tag{1.5}$$

The proof of this result is given in Lemma B.2 of [9].

**Remark 1.2.** For  $x \in \mathbb{R}^d, x \neq 0$ ,  $\lambda_x$  corresponds to a real part of some eigenvalue of  $Db(0)$  and  $\{v^k, k = 1, \dots, m\}$  are elements of the Jordan decomposition of  $Db(0)$  according to the flag of eigenspaces (along increasing real parts of the corresponding eigenvalues) containing  $x$ . For any generic choice of  $x$ ,  $\lambda_x$  corresponds to the smallest real part of the eigenvalues of  $Db(0)$ .

**1.2.2 The stochastic perturbation  $\varepsilon dL$**

On a given probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  consider a Lévy process  $L = (L_t)_{t \geq 0}$  with values in  $\mathbb{R}^d$ , i.e. a stochastic process with càdlàg paths, independent and stationary increments and issued from 0. Its marginals are determined by the celebrated Lévy-Khintchin formula

$$\mathbb{E}[e^{i\langle u, L_t \rangle}] = e^{t\Psi(u)} \quad \text{for any } u \in \mathbb{R}^d,$$

with the characteristic exponent

$$\Psi(u) = i\langle a, u \rangle - \frac{1}{2}\langle u, \Sigma u \rangle + \int_{\mathbb{R}^d} \left( e^{i\langle u, z \rangle} - 1 - i\langle u, z \rangle \mathbf{1}_{B_1(0)}(z) \right) \nu(dz),$$

where  $B_1(0) = \{x \in \mathbb{R}^d \mid |x| < 1\}$ ,  $a \in \mathbb{R}^d, \Sigma \in \mathbb{R}^{d \times d}$  is a non-negative definite square matrix and  $\nu : \mathcal{B}(\mathbb{R}^d) \rightarrow [0, \infty]$  is a  $\sigma$ -finite Borel measure satisfying

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (1 \wedge |z|^2) \nu(dz) < \infty.$$

Let  $(\mathcal{F}_t)_{t \geq 0}$  be the enhanced natural filtration of  $L$  satisfying the usual conditions of Protter [61].

The stochastic analogue of the dynamical system (1.3) is described by the following stochastic differential equation. For  $\varepsilon > 0$ , we consider

$$\begin{cases} dX_t^\varepsilon = -b(X_t^\varepsilon)dt + \varepsilon dL_t & \text{for } t \geq 0, \\ X_0^\varepsilon = x, \end{cases} \tag{1.6}$$

which under Hypothesis 1 has a unique strong solution  $X^{\varepsilon, x} = (X_t^{\varepsilon, x})_{t \geq 0}$ . Such strong solution satisfies the strong Markov property with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ , see for instance p. 1026 in [77] and the references therein.

**1.2.3 Exponential ergodicity and regularity of the limiting distributions  $\mu^\varepsilon$**

**a) Hypotheses on the Lévy measure:** The existence of invariant measures is known to be true for systems with as little as logarithmic moments [44], however we need exponential ergodicity in the total variation distance, which typically needs some (arbitrarily low) finite moments and regularity of the transition kernel for the Lévy measure, see for instance [47]. Both requirements are met by the class of Lévy measures defined below.

The cutoff results we have in mind can be understood as asymptotically precise small noise formulations of an exponential ergodicity result in total variation distance. Such results typically need some kind of finite positive moments. We refer to a more detailed discussion directly after Theorem 1. To our knowledge – apart from dimension  $d = 1$  in [47] – there are not exponential ergodicity results available in the literature with weaker moment hypotheses.

**Hypothesis 2** (Moment condition).

We assume

$$\int_{|z|>1} |z|^\beta \nu(dz) < \infty \quad \text{for some } \beta > 0.$$

Since we consider a smooth exponentially stable dynamical system with a small random perturbation, it is natural to apply a linearization procedure, which makes it necessary to compare  $X^{\varepsilon,x}$  with a suitable linearized process  $Y^\varepsilon(x)$ . As they have different drift terms, this comparison can hold only for short times. In addition, as explained in the introduction,  $X^{\varepsilon,x}$  can be understood as  $Y^\varepsilon(x)$  plus some short time error term, which turns out to be not of independent nature and therefore hard to treat in the total variation, since the analogous statement of Slutsky’s lemma (for instance [45], Section 13.2, Theorem 13.18) for the total variation distance is false in general. We are not aware of this result in the literature and hence provide a counterexample in Subsection 1.3.5. The resulting difficulty is overcome by a short-time local limit theorem. Such a result has been given in Theorem 3.1 in [42] and requires some kind of regularization in terms of a sufficiently steep pole of the Lévy measure at the origin. With this reasoning in mind it comes not as a surprise that our results are shown for a specific class of Lévy processes with such a property. In what follows, we assume that the Lévy process  $L$  has no Gaussian component and its Lévy measure belongs to the following class.

**Definition 1.3** (Locally layered stable Lévy measure).

Let  $\nu$  be a Lévy measure on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . Then  $\nu$  is called a **locally layered stable Lévy measure** with parameters  $(\nu_0, \nu_\infty, \Lambda, q, c_0, \alpha)$  if the following is satisfied. There exist  $\sigma$ -finite Borel measures  $\nu_0$  and  $\nu_\infty$  such that  $\nu = \nu_0 + \nu_\infty$ , where  $\nu_\infty$  is a finite measure with support contained in  $\{|z| > 1\}$  and

$$\nu_0(A) = \int_{\mathbb{S}^{d-1}} \Lambda(d\theta) \int_0^1 \mathbf{1}_A(r\theta) q(r, \theta) dr \quad \text{for any } A \in \mathcal{B}(\mathbb{R}^d), 0 \notin \bar{A},$$

where  $\Lambda$  is a finite positive measure on the unit sphere in  $\mathbb{R}^d$ ,  $\mathbb{S}^{d-1} = \{|z| = 1\}$ , and  $q : (0, 1] \times \mathbb{S}^{d-1} \rightarrow (0, \infty)$  is a locally integrable function for which there exist a positive function  $c_0$  in  $L^1(\Lambda)$  and a parameter  $\alpha \in (0, 2)$  such that

$$|r^{1+\alpha} q(r, \theta) - c_0(\theta)| \rightarrow 0, \quad \text{as } r \rightarrow 0$$

for  $\Lambda$ -almost all  $\theta \in \mathbb{S}^{d-1}$ . A pure jump Lévy process with a locally layered stable Lévy measure is called a **locally layered stable Lévy process**.

This notion generalizes naturally the notion of a layered stable Lévy measure (and the respective Lévy process) introduced in Definition 2.1 of [42] to all Lévy measures for which Theorem 3.1 (Short-range behavior) remains valid under Hypothesis 2. They include more general tail measures  $\nu_\infty$  than layered stable Lévy measures given in [42], such as tempered stable Lévy measures defined in [62] and Lamperti stable Lévy measures [22]. The following more restrictive notion is tailor-made to strengthen the result of Theorem 3.1 in [42] to the convergence in the total variation distance which turns out to be crucial in the proof of Proposition 3. In addition, in Theorem 4 in Appendix D we extend Theorem 4.1 of [60] and show that under Hypothesis 2 the system (1.6) is strongly ergodic under the total variation distance.

**Definition 1.4** (Strongly locally layered stable Lévy measure).

Let  $\nu$  be a locally layered stable Lévy measure with parameters  $(\nu_0, \nu_\infty, \Lambda, q, c_0, \alpha)$ . If, in addition, we have the small jump symmetry

$$q(r, \theta) = q(r, -\theta) \quad \text{for any } r \in (0, 1), \theta \in \mathbb{S}^{d-1}, \quad (1.7)$$

the uniform convergence

$$\sup_{\theta \in \mathbb{S}^{d-1}} |r^{1+\alpha}q(r, \theta) - c_0(\theta)| \rightarrow 0, \quad \text{as } r \rightarrow 0, \quad (1.8)$$

and the gradient estimate

$$|\nabla \log q(r, \theta)| \leq C_1 r^{-1} \quad \text{for some } C_1 > 0 \text{ and all } r \in (0, 1), \quad (1.9)$$

we call  $\nu$  a **strongly locally layered stable Lévy measure** with the parameters  $(\nu_0, \nu_\infty, \Lambda, q, c_0, \alpha)$ . A pure jump Lévy process with a strongly locally layered stable Lévy measure is called a **strongly locally layered stable Lévy process**.

**Remark 1.5.** Examples of such processes are symmetric  $\alpha$ -stable Lévy processes (see [4, 65]), symmetric tempered  $\alpha$ -stable process [62] and symmetric Lamperti  $\alpha$ -stable processes.

**Hypothesis 3** (Regularity).

We assume that the Lévy process  $L$  has no Gaussian component and its Lévy measure  $\nu$  is strongly locally layered stable with parameters  $(\nu_0, \nu_\infty, \Lambda, q, c_0, \alpha)$ .

In the sequel, we define sufficient conditions for an abrupt convergence of  $X_t^{\varepsilon, x}$  to its unique limiting distribution  $\mu^\varepsilon$  as  $\varepsilon \rightarrow 0$  in the total variation distance.

**b) The total variation distance  $\|\cdot\|_{\text{TV}}$ :** Before we introduce the concept of cutoff formally, we recall the notion of the total variation distance. Given two probability measures  $\mathbb{P}$  and  $\mathbb{Q}$  which are defined on the same measurable space  $(\Omega, \mathcal{F})$ , denote the total variation distance between  $\mathbb{P}$  and  $\mathbb{Q}$  as follows

$$\|\mathbb{P} - \mathbb{Q}\|_{\text{TV}} := \sup_{A \in \mathcal{F}} |\mathbb{P}(A) - \mathbb{Q}(A)|.$$

For simplicity, in the case of two random vectors  $X$  and  $Y$  defined on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  we use the following notation for its total variation distance

$$\|X - Y\|_{\text{TV}} := \|\mathcal{L}(X) - \mathcal{L}(Y)\|_{\text{TV}},$$

where  $\mathcal{L}(X)$  and  $\mathcal{L}(Y)$  denote the law under  $\mathbb{P}$  of the random vectors  $X$  and  $Y$ , respectively. For the sake of intuitive reasoning and in a conscious abuse of notation we write  $\|X - \mu_Y\|_{\text{TV}}$  instead of  $\|X - Y\|_{\text{TV}}$ , where  $\mu_Y$  is the distribution of the random vector  $Y$ .



For a complete understanding of the total variation distance, we refer to Chapter 2 of the monograph of Kulik [48] and the references therein.

**c) Exponential ergodicity with smooth limiting measure.**

As we mentioned before, we are interested on the cutoff under the total variation distance, which is a rather robust distance for continuous distributions and rather sensitive for discrete distributions. It is therefore natural to assume the following additional hypothesis which with the help of Hypothesis 3 yields smooth densities for the finite time marginals and the limiting distribution of (1.6).

**Hypothesis 4** (Equator condition [71]).

Let  $\nu$  satisfy Hypothesis 3. The support of the measure  $\Lambda$  is not contained in any proper subspace of  $\mathbb{R}^d$  intersected with  $\mathbb{S}^{d-1}$ . Furthermore, we assume

$$c_0 := \operatorname{ess\,inf}_{\theta \in \mathbb{S}^{d-1}} c_0(\theta) > 0, \tag{1.10}$$

where the essential infimum is understood with respect to the spectral measure  $\Lambda$  of  $\nu$ .

The equator condition is motivated by the definition given in Simon [71], p.4. It provides a non-degeneracy condition on the support of  $\Lambda$  on  $\mathbb{S}^{d-1}$ .

**Remark 1.6.** It is not hard to see that Hypothesis 4 (1.10) implies

$$\inf_{\bar{v} \in \mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \cos^2(\langle \bar{v}, \theta \rangle) \Lambda(d\theta) > 0, \quad \text{where} \quad \cos(\langle \bar{v}, \theta \rangle) = \langle \bar{v}, \theta \rangle. \tag{1.11}$$

The following lemma links Definition 1.4 and Hypothesis 4 to the celebrated Orey-Masuda regularity condition, which is used in the proof of Proposition 2.

**Lemma 1.7** (Orey-Masuda’s cone condition).

Let  $\nu$  be a strongly locally layered stable Lévy measure on  $\mathbb{R}^d$  with the parameters  $(\nu_0, \nu_\infty, \Lambda, q, c_0, \alpha)$  for  $\alpha \in (0, 2)$ . Under Hypothesis 4 there exist positive constants  $c_{\triangleleft}$  and  $C_{\triangleleft}$  such that for all  $v \in \mathbb{R}^d$  with  $|v| > C_{\triangleleft}$  we have

$$\int_{|\langle v, z \rangle| \leq 1} |\langle v, z \rangle|^2 \nu(dz) \geq \int_{|\langle v, z \rangle| \leq 1} |\langle v, z \rangle|^2 \nu_0(dz) \geq c_{\triangleleft} |v|^\alpha.$$

*Proof.* Observe

$$\begin{aligned} \int_{|\langle v, z \rangle| \leq 1} |\langle v, z \rangle|^2 \nu(dz) &\geq \int_{|\langle v, z \rangle| \leq 1, |z| \leq 1} |\langle v, z \rangle|^2 \nu(dz) \\ &\geq |v|^2 \int_{\mathbb{S}^{d-1}} \int_0^1 r^2 \langle \bar{v}, \theta \rangle^2 \mathbf{1}\{r|v| |\cos(\langle \bar{v}, \theta \rangle)| \leq 1\} q(r, \theta) dr \Lambda(d\theta), \end{aligned}$$

where  $\bar{v} = v/|v|$ ,  $r = |z|$  and  $\theta = z/r$ . By (1.8) and (1.10) there exists  $r_0 > 0$  (without loss of generality  $r_0 \leq 1$ ) such that

$$q(r, \theta) \geq \frac{c_0}{2r^{1+\alpha}} \quad \text{for any } r \in (0, r_0).$$

Consequently, for  $|v| > 1/r_0 \geq 1$  we have

$$\begin{aligned} |v|^2 \int_{\mathbb{S}^{d-1}} \int_0^1 r^2 \langle \bar{v}, \theta \rangle^2 \mathbf{1}\left\{|\cos(\langle \bar{v}, \theta \rangle)| \leq \frac{1}{r|v|}\right\} q(r, \theta) dr \Lambda(d\theta) \\ &\geq |v|^2 \int_{\mathbb{S}^{d-1}} \int_0^{1/|v|} r^2 \cos^2(\langle \bar{v}, \theta \rangle) q(r, \theta) dr \Lambda(d\theta) \\ &\geq \left(\frac{c_0}{2(2-\alpha)} \inf_{\bar{v} \in \mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \cos^2(\langle \bar{v}, \theta \rangle) \Lambda(d\theta)\right) |v|^\alpha, \end{aligned}$$

which combined with Hypothesis 4 finishes the proof. □

The following result is a slight generalization of Theorem 4.1 in [60] and guarantees that under Hypotheses 1, 2, 3 and 4 the system (1.6) is strongly ergodic under the total variation distance.

**Theorem 1.** *Assume Hypotheses 1, 2, 3 and 4 for  $\alpha \in (0, 2)$  and  $\beta > 0$ . Then for any  $\varepsilon > 0$ , there exists a unique invariant distribution  $\mu^\varepsilon$  and positive constants  $C_\varepsilon, \theta_\varepsilon$  such that for all  $x \in \mathbb{R}^d$ , the law of the unique strong solution  $X^{\varepsilon,x}$  of (1.6) satisfies*

$$\|X_t^{\varepsilon,x} - \mu^\varepsilon\|_{\text{TV}} \leq C_\varepsilon e^{-\theta_\varepsilon t} (1 + |x|^{1 \wedge \beta}) \quad \text{for any } t \geq 0.$$

The proof is a direct corollary of Theorem 4 given in Appendix D. The tracking of the dependence  $\varepsilon \mapsto (\theta_\varepsilon, C_\varepsilon)$  is typically hard to follow through the discretization procedure laid out by Meyn and Tweedie [56]. In the special case of finite variation, the backtracking of  $\varepsilon$  can be carried out partially, we refer to [47].

We recall that in dimension  $d = 1$ , a classical result by Kulik (see Proposition 0.1 in [47]) implies that the solution of (1.6) enjoys exponential ergodicity without assumption (1.9) and consequently Theorem 1 holds for general locally layered stable Lévy measures in this case. Very recently, [58] contains exponential ergodicity by control theoretic methods for multidimensional compound Poisson noise with finite variance. For higher dimensions, we use the sufficient conditions including (1.9) in [60] and our generalizations of their results given in Appendix D. We point out that for the special case of symmetric  $\alpha$ -stable Lévy processes, assumption (1.9) is automatically satisfied and [76] yields exponential ergodicity in any dimension.

#### 1.2.4 The main results: window cutoff (Thm. 2) and profile cutoff (Thm. 3)

Following [12] and the references therein, there are three notions of cutoff phenomenon with increasing strength. The most restrictive notion is called profile cutoff which provides the precise asymptotic shape of the collapse for the total variation distance. Profile cutoff implies a weaker concept which is called window cutoff that states abrupt convergence within a precise time interval but losing the precise profile. Window cutoff is generalized further to the notion of cutoff in which we retain the abrupt convergence along time scale which corresponds to the center of the interval, however, without a quantification of the error.

**Definition 1.8.** *For any  $\varepsilon > 0$  and  $x \in \mathbb{R}^d$ , let  $X^{\varepsilon,x}$  be the solution of (1.6) with a unique limiting distribution  $\mu^\varepsilon$ . We say that for  $x \in \mathbb{R}^d$  the family  $(X^{\varepsilon,x})_{\varepsilon \in (0,1]}$  exhibits*

- a) a **cutoff phenomenon** at the time scale  $(t_\varepsilon^x)_{\varepsilon \in (0,1]}$ , where  $t_\varepsilon^x \rightarrow \infty$ , as  $\varepsilon \rightarrow 0$ , if it satisfies

$$\lim_{\varepsilon \rightarrow 0} \left\| X_{\delta \cdot t_\varepsilon^x}^{\varepsilon,x} - \mu^\varepsilon \right\|_{\text{TV}} = \begin{cases} 1 & \text{if } \delta \in (0, 1), \\ 0 & \text{if } \delta \in (1, \infty). \end{cases}$$

- b) a **window cutoff phenomenon** at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)_{\varepsilon \in (0,1]}$ , where  $t_\varepsilon^x \rightarrow \infty$  and  $w_\varepsilon^x/t_\varepsilon^x \rightarrow 0$ , as  $\varepsilon \rightarrow 0$ , if it satisfies

$$\lim_{\rho \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \left\| X_{t_\varepsilon^x + \rho \cdot w_\varepsilon^x}^{\varepsilon,x} - \mu^\varepsilon \right\|_{\text{TV}} = 1 \quad \text{and} \quad \lim_{\rho \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \left\| X_{t_\varepsilon^x + \rho \cdot w_\varepsilon^x}^{\varepsilon,x} - \mu^\varepsilon \right\|_{\text{TV}} = 0.$$

- c) a **profile cutoff phenomenon** at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)_{\varepsilon \in (0,1]}$  with the profile function  $G_x$ , where  $t_\varepsilon^x \rightarrow \infty$  and  $w_\varepsilon^x/t_\varepsilon^x \rightarrow 0$ , as  $\varepsilon \rightarrow 0$ , if the limit

$$G_x(\rho) := \lim_{\varepsilon \rightarrow 0} \left\| X_{t_\varepsilon^x + \rho \cdot w_\varepsilon^x}^{\varepsilon,x} - \mu^\varepsilon \right\|_{\text{TV}}$$

The cutoff phenomenon for nonlinear Langevin systems with small noise

is well-defined for all  $\rho \in \mathbb{R}$  and  $G_x$  satisfies

$$\lim_{\rho \rightarrow -\infty} G_x(\rho) = 1 \quad \text{and} \quad \lim_{\rho \rightarrow \infty} G_x(\rho) = 0.$$

The cut-off time scale  $t_\varepsilon^x$  is sometimes referred to as the center of the cutoff window and  $w_\varepsilon^x$  as its width. As mentioned above iii) implies ii) and ii) implies i).

The first main result of this study reads as follows.

**Theorem 2** (Generic window cutoff phenomenon).

Assume Hypotheses 1, 2, 3 and 4 are satisfied for some  $\alpha \in (3/2, 2)$  and  $\beta > 0$ . For any  $\varepsilon > 0$  and  $x \in \mathbb{R}^d \setminus \{0\}$ , let  $X^{\varepsilon, x}$  be the unique strong solution of (1.6) with a unique limiting distribution  $\mu^\varepsilon$ . Then the family  $(X^{\varepsilon, x})_{\varepsilon \in (0, 1]}$  exhibits a window cutoff phenomenon as  $\varepsilon \rightarrow 0$  at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$  given by

$$t_\varepsilon^x = \frac{1}{\lambda_x} \ln(1/\varepsilon) + \frac{\ell_x - 1}{\lambda_x} \ln(\ln(1/\varepsilon)) \quad \text{and} \quad w_\varepsilon^x = \frac{1}{\lambda_x} + o_\varepsilon(1), \quad (1.12)$$

where  $\lambda_x > 0$  and  $\ell_x \in \{1, \dots, d-1\}$  are the constants appearing in the Hartman-Grobman decomposition of Lemma 1.1.

Note that  $x = 0$  in Theorem 2 is essential.

**Remark 1.9.** For  $x = 0$ , there is no cutoff phenomenon since the linearization vanishes and intuitively cannot compete with the ergodicity. For details see Remark 2.2. For a complete discussion of the easier case of the Wasserstein distance, we refer to Section 3.2 in [6].

Assume the hypotheses of Theorem 2 are satisfied for some  $x \in \mathbb{R}^d \setminus \{0\}$ . Let  $v(t, x) = \sum_{k=1}^m e^{i\theta_x^k t} v_x^k$  and  $\lambda_x, \ell_x, \theta_x^1, \dots, \theta_x^m$  and  $v_x^1, \dots, v_x^m$  given in Lemma 1.1. We define the  $\omega$ -limit set for the dynamics of  $(v(t, x))_{t \geq 0}$  by

$$\omega(x) := \{v \in \mathbb{R}^d : \text{there exists a sequence } (t_j) \rightarrow \infty \text{ and } \lim_{j \rightarrow \infty} v(t_j, x) = v\}, \quad (1.13)$$

which due to the left-hand side of (1.5) does not include the null vector, i.e.  $0 \notin \omega(x)$ .

**Remark 1.10.** Note that  $\omega(x) \neq \emptyset$ . Indeed, a Cantor diagonal argument for any limiting sequence in (1.4) yields the existence of a subsequence  $(t_j)_{j \in \mathbb{N}}$  with  $t_j \rightarrow \infty$ , as  $j \rightarrow \infty$ , such that for any  $k = 1, \dots, m$  the limit  $\lim_{j \rightarrow \infty} e^{it_j \theta^k} = \vartheta_k$  exist. Moreover,  $|\vartheta_k| = 1$  for all  $k$ . Since  $v^1, \dots, v^m$  are linearly independent vector in  $\mathbb{C}^d$ , we deduce  $v = \sum_{j=1}^m \vartheta_j v^j \in \omega(x)$ .

In an abuse of notation let  $Z_\infty$  denote a parametrization of the unique invariant distribution of the Ornstein-Uhlenbeck process

$$dZ_t = -Db(0)Z_t dt + dL_t.$$

We have the following characterization of profile cutoff.

**Theorem 3** (A dynamical characterization of a profile cutoff phenomenon).

Assume the hypotheses of Theorem 2 are satisfied for some  $x \in \mathbb{R}^d \setminus \{0\}$ . Recall the  $\omega$ -limit set  $\omega(x)$  given in (1.13). Then the family  $(X^{\varepsilon, x})_{\varepsilon \in (0, 1]}$  exhibits a profile cutoff phenomenon as  $\varepsilon \rightarrow 0$  at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$  given by Theorem 2 with profile function

$$G_x(\rho) = \left\| \left( e^{-\rho} \cdot \frac{e^{-\lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} v + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} \quad \text{for any } \rho \in \mathbb{R}, v \in \omega(x)$$

if and only if for any  $a > 0$  the map

$$\omega(x) \ni v \mapsto \|(av + Z_\infty) - Z_\infty\|_{\text{TV}} \quad \text{is constant.} \quad (1.14)$$

Observe that  $\omega(x) = \{v_x\}$  immediately implies profile cutoff by the preceding theorem. The latter, indeed, is satisfied in the subsequent case of a gradient potential.

The following special case of gradient systems is particularly of interest in applications, such as for instance the Fermi-Ulam-Pasta-Tsingou-potential treated in Subsection 1.3.2.

**Corollary 1.11.** *Let the assumptions of Theorem 2 be satisfied and assume  $b(x) = \nabla\mathcal{V}(x)$ ,  $x \in \mathbb{R}^d$ , for a potential function  $\mathcal{V} : \mathbb{R}^d \rightarrow [0, \infty)$ . Then the family  $(X^{\varepsilon,x})_{\varepsilon \in (0,1]}$  exhibits a profile cutoff as  $\varepsilon \rightarrow 0$  at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$  given by*

$$t_\varepsilon^x = \frac{1}{\lambda_x} \ln(1/\varepsilon) \quad \text{and} \quad w_\varepsilon^x = \frac{1}{\lambda_x} + o_\varepsilon(1),$$

where  $\lambda_x > 0$  and  $\tau^x$  are the positive constants in the Hartman-Grobman decomposition of Lemma 1.1 such that

$$\lim_{t \rightarrow \infty} e^{\lambda_x t} \varphi_{t+\tau^x}^x = v^x \neq 0 \tag{1.15}$$

and the profile function is given by

$$G_x(\rho) = \|(e^{-\rho} \cdot e^{-\lambda_x \tau^x} v^x + Z_\infty) - Z_\infty\|_{\text{TV}}, \quad \rho \in \mathbb{R}.$$

**Remark 1.12.** *Note that the dependence of  $\lambda_x$  of  $x$  can be complicated, however, it is rather weak in the following qualitative sense:  $\lambda_x = \lambda$  for Lebesgue almost every  $x \in \mathbb{R}^d$ , where  $\lambda$  is the smallest eigenvalue of the positive definite symmetric matrix  $D^2\mathcal{V}(0)$ .*

We give a more general sufficient conditions for the existence of a cutoff profile in terms of a symmetry condition.

**Corollary 1.13.** *Assume the hypotheses of Theorem 2 are satisfied for some  $x \in \mathbb{R}^d \setminus \{0\}$ . If there exists an invertible  $d \times d$ -square matrix  $M$  such that the distribution of  $MZ_\infty$  is rotationally invariant and the image set satisfies  $M\omega(x) \subset \{|z| = r\}$  for some positive  $r = r_x$ , then the family  $(X^{\varepsilon,x})_{\varepsilon \in (0,1]}$  exhibits a profile cutoff phenomenon as  $\varepsilon \rightarrow 0$  at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$ .*

In the Gaussian case we have the following picture.

**Remark 1.14.** *For the non-degenerate Gaussian case we refer to Lemma A.2 in [8]. There, the law of  $Z_\infty$  is  $\mathcal{N}(0, \Sigma)$ , where  $\Sigma$  satisfies*

$$Db(0)\Sigma + \Sigma Db(0)^* = I_d.$$

*The choice of  $M = \Sigma^{-1/2}$  yields that  $MZ_\infty \stackrel{d}{=} \mathcal{N}(0, I_d)$  is rotationally invariant. Hence the sphere condition  $M\omega(x) \subset \{|z| = r\}$  for some  $r = r_x > 0$  is equivalent to the profile cutoff, see Corollary 2.11 in [8]. However, in the generic Lévy case, no symmetry on the law of  $Z_\infty$  can be expected. Note that we always find an invertible bi-measurable map  $\mathcal{T} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that the push-forward  $\mathcal{T}(Z_\infty)$  is rotationally invariant (for instance  $\mathcal{N}(0, I_d)$ ), however, it is highly nonlinear and irregular, and therefore the proof of Corollary 1.13 breaks down.*

A sufficient condition for the hypotheses of Corollary 1.13 to be satisfied can be given in terms of the following density condition on the invariant limiting measure of the Ornstein-Uhlenbeck process  $Z$ .

**Corollary 1.15** (Geometric profile characterization under rotational invariant  $Z_\infty$ ). *Assume the hypotheses of Theorem 2 are satisfied for some  $x \in \mathbb{R}^d \setminus \{0\}$ . If in addition, the law of  $Z_\infty$  is rotationally invariant and its density  $f \in C^1(\mathbb{R}^d, (0, \infty))$  is unimodal in*

the sense that  $f(z) = g(|z|)$  for some function  $g \in C^1((0, \infty), (0, \infty))$  with  $g'(s) < 0$  for all  $s > 0$  and  $g' \in L^1(\mathbb{R}^d)$ . Then the image set satisfies  $\omega(x) \subset \{|z| = r\}$  for some  $r = r_x > 0$  if and only if the family  $(X^{\varepsilon, x})_{\varepsilon \in (0, 1]}$  exhibits a profile cutoff as  $\varepsilon \rightarrow 0$  at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$ .

In case of a pure jump Lévy noise  $L$  the sufficient condition of Corollary 1.13 can be almost characterized (up to a non-resonance condition) in terms of the following normal growth condition, which is discussed in detail in [6].

**Remark 1.16** (Generic normal growth profile characterization). *In the sequel, we characterize when the function*

$$\omega(x) \ni u \mapsto |u|$$

is constant for the generic case of the setting in Corollary 1.15. We enumerate  $v^1, \dots, v^m$  given in Lemma 1.1 as follows. Without loss of generality we assume that  $\theta^1 = 0$ . Otherwise we take  $v^1 = 0$  and eliminate it from the sum  $\sum_{k=1}^m e^{i\theta^k t} v^k$ . Without loss of generality let  $m = 2n + 1$  for some  $n \in \mathbb{N}$ . We assume that  $v^k$  and  $v^{k+1} = \bar{v}^k$  are complex conjugate for all even number  $k \in \{2, \dots, m\}$ . For  $k \in \{2, \dots, m\}$  we write  $v^k = \hat{v}^k + i\check{v}^k$  where  $\hat{v}^k, \check{v}^k \in \mathbb{R}^d$ .

1. If the real parts and the imaginary parts of the (complex) vectors  $v^2, v^4, \dots, v^{2n}$  in the Hartman-Grobman Lemma 1.1 form an orthogonal family and  $|\operatorname{Re}(v^{2k})| = |\operatorname{Im}(v^{2k})|$  for all  $k$ . Then Lemma E.1 in [6] implies that  $\omega(x) \subset \{|z| = r\}$  for some  $r = r_x > 0$  and hence Corollary 1.13 yields a profile cutoff.
2. Assume the angles  $\theta^2, \theta^4, \dots, \theta^{2n}$  given in the Hartman-Grobman Lemma 1.1 are rationally independent from  $2\pi$ . If  $\omega(x) \subset \{|z| = r\}$  for some  $r = r_x > 0$  then Lemma E.2 implies that the real parts and the imaginary parts of the (complex) vectors  $v^2, v^4, \dots, v^{2n}$  in the Hartman-Grobman Lemma 1.1 form an orthogonal family and  $|\operatorname{Re}(v^{2k})| = |\operatorname{Im}(v^{2k})|$  for all  $k$ .

*Proof of Corollary 1.13:* We apply the characterization given in Theorem 3. Let  $v_1, v_2 \in \omega(x)$  and  $a > 0$ . For  $M$  given in the statement, we have  $|Mv_1| = |Mv_2| = r$ . Then there exists an orthogonal matrix  $\mathcal{O}$  such that  $\mathcal{O}(Mv_1) = Mv_2$ . Theorem 5.2 of [28] and  $\mathcal{O}, M$  being invertible implies

$$\begin{aligned} \|(av_1 + Z_\infty) - Z_\infty\|_{\text{TV}} &= \|(aMv_1 + MZ_\infty) - MZ_\infty\|_{\text{TV}} \\ &= \|(a\mathcal{O}(Mv_1) + \mathcal{O}MZ_\infty) - \mathcal{O}MZ_\infty\|_{\text{TV}}. \end{aligned}$$

Since  $\mathcal{O}(Mv_1) = Mv_2$  and  $\mathcal{O}$  is orthogonal, the rotational invariance of  $MZ_\infty$  implies

$$\|(a\mathcal{O}(Mv_1) + \mathcal{O}MZ_\infty) - \mathcal{O}MZ_\infty\|_{\text{TV}} = \|(aMv_2 + MZ_\infty) - MZ_\infty\|_{\text{TV}}.$$

Again, Theorem 5.2 of [28] yields

$$\|(aMv_1 + MZ_\infty) - MZ_\infty\|_{\text{TV}} = \|(av_2 + Z_\infty) - Z_\infty\|_{\text{TV}}.$$

Combining the preceding equalities we obtain

$$\|(av_1 + Z_\infty) - Z_\infty\|_{\text{TV}} = \|(av_2 + Z_\infty) - Z_\infty\|_{\text{TV}}$$

for any  $v_1, v_2 \in \omega(x)$  and  $a > 0$  which yields (1.14) and hence the desired profile cutoff.  $\square$

*Proof of Corollary 1.15:* By Corollary 1.13 ( $M = I_d$ ) it is enough to prove the converse implication. Since the family  $(X^{\varepsilon, x})_{\varepsilon \in (0, 1]}$  exhibits a profile cutoff as  $\varepsilon \rightarrow 0$  at the

enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$ , Theorem 3 implies for all  $a > 0$  that the following mapping  $v \in \omega(x) \mapsto \|(av + Z_\infty) - Z_\infty\|_{TV}$  is constant. Since the law of  $Z_\infty$  is rotationally invariant, we have

$$\|(av + Z_\infty) - Z_\infty\|_{TV} = \|(a|v|e_1 + Z_\infty) - Z_\infty\|_{TV},$$

where  $e_1 = (1, 0, \dots, 0)^*$ . By Lemma C.5 in Appendix C we have that  $\omega(x) \ni v \mapsto a|v|$  is constant. That is to say,  $\omega(x) \subset \{|z| = r_x\}$  for some  $r_x > 0$ . This finishes the proof.  $\square$

### 1.3 Examples

#### 1.3.1 More general linear dynamics

When the vector field is given by  $b(x) = Qx$ ,  $x \in \mathbb{R}^d$  for a general deterministic  $d \times d$  matrix  $Q$  whose eigenvalues have positive real parts, the cutoff phenomenon is completely discussed in [10], Theorem 2.3 under Hypothesis (H), which is covered by Hypothesis 4. It is well-known that such linear systems are more general than linear systems satisfying Hypothesis 1. For instance, the classical linear oscillator with friction  $\gamma > 0$  has negative real parts in  $(-\infty, -\gamma/2]$  but fails to be coercive, [6].

The case of pure Brownian motion is covered in detail in Section 3.3 in [9]. For degenerate driving noise processes  $L$  and general cutoff results in the Wasserstein distance, we refer to [6]. There, complex systems of linear oscillators in a thermal bath are covered.

#### 1.3.2 Gradient systems: Fermi-Ulam-Pasta-Tsingou

In the sequel, we consider the generalized Fermi-Ulam-Pasta-Tsingou potential [35, 26]

$$\mathcal{V}(x) = |Ax|^2/2 + |Bx|^4/4 + \eta(x), \quad x \in \mathbb{R}^d, \quad (1.16)$$

where  $A$  and  $B$  are  $d \times d$  deterministic matrices satisfying for some  $\delta_1 > 0$

$$\langle Ax, x \rangle \geq \delta_1|x|^2 \quad \text{and} \quad \langle Bx, x \rangle \geq 0 \quad \text{for all } x \in \mathbb{R}^d \quad (1.17)$$

and some  $\eta : \mathbb{R}^d \rightarrow \mathbb{R}$  with  $\eta \in C_b^2$ ,  $\nabla\eta(0) = 0$ , and for  $H_\eta$  being the Hessian of  $\eta$

$$\langle H_\eta(x)y, y \rangle \geq -\delta_2|y|^2$$

for all  $x, y \in \mathbb{R}^d$  and some  $\delta_2 < \delta_1$ . Note that  $\eta$  needs not be convex. We set  $b(x) = \nabla\mathcal{V}(x)$ ,  $x \in \mathbb{R}^d$ . Then for all  $x \in \mathbb{R}^d$  we have

$$b(x) = A^*Ax + \langle Bx, Bx \rangle B^*Bx + \nabla\eta(x) \quad (1.18)$$

and satisfies Hypothesis 1. Indeed, the Jacobian of  $b$  at  $x$  is given by

$$Db(x) = A^*A + 3\langle Bx, Bx \rangle B^*B + H_\eta(x),$$

where  $H_\eta(x)$  denote the Hessian matrix at  $x$ . By (1.17) we obtain for any  $x, y \in \mathbb{R}^d$

$$\langle y, Db(x)y \rangle = \langle Ay, Ay \rangle + 3\langle Bx, Bx \rangle \langle By, By \rangle + \langle H_\eta(x)y, y \rangle \geq \delta|y|^2,$$

where  $\delta = \delta_1 - \delta_2 > 0$ . Hence the vector field  $b = \nabla\mathcal{V}$  satisfies Hypothesis 1. We consider the solution of (1.3) with vector field  $b = \nabla\mathcal{V}$ . Note that in this case  $|\omega(x)| = 1$ , where  $\omega(x)$  is given in (1.13).

Note that generically equation (1.6) does not have an explicitly known solution for the Kolmogorov forward equation of the densities, not even in simplest case of  $d = 1$ ,  $L = W$ ,

a standard Wiener process,  $A = B = 1$  and  $\eta \equiv 0$ . While for any dimension  $d$ ,  $L = W$  a standard Brownian motion the invariant density is well-known to be proportional to  $\exp(-2\mathcal{V}(x)/\varepsilon^2)$ , for a complete discussion, see for instance Section 2.2 in [70]. For the case of dimension  $d = 1$ , nonlinear  $b$  satisfying Hypothesis 1,  $L = W$  a standard Brownian motion the authors prove profile cutoff for (1.6) in [8]. For higher dimensions, window cutoff is established in this case and the existence of profile cutoff is characterized, we refer to [9]. We remark that the authors strongly use the hypo-ellipticity property and the resulting regularization by the generator of the Brownian diffusion.

For a strongly locally layered stable noise  $L$  satisfying Hypotheses 2, 3 and 4, Corollary 1.11 implies for  $\alpha > 3/2$  the presence of a cutoff profile. In particular, the system exhibits cutoff in the sense of equation (18.3) in Chapter 18 of the monograph [53] as follows

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{T}_{\text{mix}}^{x,\varepsilon}(\eta)}{\mathcal{T}_{\text{mix}}^{x,\varepsilon}(1-\eta)} = 1$$

for any  $\eta \in (0, 1)$ , where the mixing time is given by

$$\mathcal{T}_{\text{mix}}^{x,\varepsilon}(\eta) = \inf\{t \geq 0 : \|X_t^\varepsilon(x) - \mu^\varepsilon\|_{\text{TV}} \leq \eta\}.$$

### 1.3.3 Profile vs Window cutoff for nonlinear oscillations

In the sequel we analyze a class of nonlinear oscillators for which the existence of a cutoff profile is studied in detail. We consider the nonlinear system (1.6) in  $\mathbb{R}^2$ , where  $b : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is given by

$$b(x_1, x_2) = \begin{pmatrix} \eta x_2 + \partial_1 \mathcal{H}(x_1, x_2) \\ -\eta x_1 + \partial_2 \mathcal{H}(x_1, x_2) \end{pmatrix},$$

for some  $\eta \in \mathbb{R}$ ,

$$\mathcal{H}(x_1, x_2) = \delta_1 x_1^2 + \delta_2 x_2^2 + \mathcal{G}(x_1, x_2)$$

for any  $x_1, x_2 \in \mathbb{R}$  and some positive constant  $\delta_1, \delta_2$ , and  $\mathcal{G} \in C^2(\mathbb{R}^2, \mathbb{R})$ . Assume that  $b(0, 0) = (0, 0)^*$ . We verify that  $b$  is a non-gradient vector field. The Jacobian matrix of  $b$  is given by

$$Db(x_1, x_2) = \begin{pmatrix} 2\delta_1 + \partial_{11}\mathcal{G}(x_1, x_2) & \eta + \partial_{12}\mathcal{G}(x_1, x_2) \\ -\eta + \partial_{12}\mathcal{G}(x_1, x_2) & 2\delta_2 + \partial_{22}\mathcal{G}(x_1, x_2) \end{pmatrix}. \quad (1.19)$$

Since for any  $\eta \neq 0$  the Jacobian  $Db$  matrix is asymmetric, there is no  $C^2$ -function  $\mathcal{V} : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that  $b(x_1, x_2) = \nabla \mathcal{V}(x_1, x_2)$  and consequently  $b$  is non-gradient. Under the assumption that

$$x_1^2 \partial_{11}^2 \mathcal{G}(u_1, u_2) + x_2^2 \partial_{22}^2 \mathcal{G}(u_1, u_2) + 2x_1 x_2 \partial_{12}^2 \mathcal{G}(u_1, u_2) \geq -\delta_3 (x_1^2 + x_2^2)$$

for some  $\delta_3 < 2 \min\{\delta_1, \delta_2\}$  and any  $x_1, x_2, u_1, u_2 \in \mathbb{R}$ , the vector field  $b$  satisfies Hypothesis 1

$$\begin{aligned} (x_1, x_2) Db(u_1, u_2) (x_1, x_2)^* &= 2\delta_1 x_1^2 + 2\delta_2 x_2^2 + x_1^2 \partial_{11}^2 \mathcal{G}(u_1, u_2) \\ &\quad + x_2^2 \partial_{22}^2 \mathcal{G}(u_1, u_2) + 2x_1 x_2 \partial_{12}^2 \mathcal{G}(u_1, u_2) \\ &\geq (2\delta_1 - \delta_3) x_1^2 + (2\delta_2 - \delta_3) x_2^2 = \delta (x_1^2 + x_2^2). \end{aligned}$$

For a rotationally invariant  $\alpha$ -stable noise  $L$  in  $\mathbb{R}^2$  with  $\alpha > 3/2$ , Theorem 2 yields window cutoff. In the sequel, we study the presence of a cutoff profile. We claim that  $Z_\infty$  is

rotationally invariant. Indeed, there is  $K_\alpha > 0$  such that the characteristic function of  $Z_\infty$  reads

$$z \mapsto \exp\left(-K_\alpha \int_0^\infty |e^{-Db(0,0)t} z|^\alpha dt\right) = \exp\left(-K_\alpha |z|^\alpha \int_0^\infty e^{-2\delta_1 \alpha t} dt\right) = \exp\left(-\frac{K_\alpha}{2\delta_1 \alpha} |z|^\alpha\right).$$

For  $a := \partial_{11}\mathcal{G}(0, 0) = \partial_{22}\mathcal{G}(0, 0)$  and  $\partial_{12}\mathcal{G}(0, 0) = 0$  we have

$$Db(0, 0) = \begin{pmatrix} 2\delta_1 + a & \eta \\ -\eta & 2\delta_2 + a \end{pmatrix}. \tag{1.20}$$

Assume a negative discriminant  $\Delta := (2\delta_2 - 2\delta_1)^2 - 4\eta^2 < 0$  and  $\delta_1 + \delta_2 + a > 0$ . Then the complex eigenvectors associated to the eigenvalues

$$\lambda_1 = \delta_1 + \delta_2 + a + \frac{\sqrt{\Delta}}{2} \quad \text{and} \quad \lambda_2 = \delta_1 + \delta_2 + a - \frac{\sqrt{\Delta}}{2}$$

are given by

$$v_1 = \left(1, \frac{2(\delta_2 - \delta_1) + \sqrt{\Delta}}{2\eta}\right) \quad \text{and} \quad v_2 = \left(1, \frac{2(\delta_2 - \delta_1) - \sqrt{\Delta}}{2\eta}\right).$$

The respective family real and imaginary part vectors are given

$$\hat{v}_1 = \hat{v}_2 = \left(1, \frac{\delta_2 - \delta_1}{\eta}\right) \quad \text{and} \quad \check{v}_1 = -\check{v}_2 = \left(0, \frac{\sqrt{|\Delta|}}{2\eta}\right).$$

**1. Nonlinear nongradient system with a cutoff profile:** For  $\delta_1 = \delta_2$  we obtain

$$e^{(2\delta_1+a)t}|e^{-Db(0,0)t}x| = |\mathcal{O}(\eta t)x| = |x|$$

for all  $t \geq 0$ , where  $\mathcal{O}(\eta t)$  is an orthogonal matrix. Therefore, whenever  $2\delta_1 + a > 0$ , Corollary 1.13 for  $M = I_2$  yields profile cutoff.

**2. Nonlinear counterexample to a cutoff profile:** Note that  $\hat{v}_2$  is orthogonal to  $\check{v}_2$  if and only if  $\delta_1 = \delta_2$ . Define  $\theta_2 = \arg(\lambda_2)$ . For  $\delta_1 < \delta_2$ , Corollary 1.15 and Remark 1.16 for  $\theta_2 \notin \mathbb{Q} \cdot \pi$  yield the absence of a cutoff profile. For further examples in the linear case for the Wasserstein distance we refer to [6].

### 1.3.4 The shape of cutoff profiles: Gaussian vs $\alpha$ -stable

Since  $Z_\infty$  is the limiting distribution as  $t \rightarrow \infty$  of the Ornstein-Uhlenbeck process  $(Z_t)_{t \geq 0}$ , Lemma 1.7 implies that  $Z_\infty$  has a  $C^\infty$  density  $f_\infty$ . In the sequel we study the unidimensional case. Theorem 53.1 in [65] yields that  $f_\infty$  is unimodal with mode  $\mathfrak{m}$ , that is to say, it is increasing on  $(-\infty, \mathfrak{m})$  and decreasing on  $(\mathfrak{m}, +\infty)$  and hence  $f_\infty(\mathfrak{m}) > 0$ . In the sequel, we determine the asymptotic behavior of the profile function  $\rho \mapsto \|(e^{-\rho} e^{-\lambda_x \tau_x} v^x + Z_\infty) - Z_\infty\|_{\text{TV}}$  in zero and at infinity for some special cases. The density  $f_\infty$  of  $Z_\infty$  is explicitly accessible only in a limited number of cases.

We start with the asymptotics for  $\rho \ll -1$ . Without loss of generality we assume that  $f_\infty$  is smooth. Then for any  $z > 0$  there exists  $\mathfrak{m}_z \in (\mathfrak{m} - z, \mathfrak{m})$  such that

$$\begin{aligned} 1 - \|(z + Z_\infty) - Z_\infty\|_{\text{TV}} &= \mathbb{P}(Z_\infty \leq \mathfrak{m}_z) + \mathbb{P}(Z_\infty \geq \mathfrak{m}_z + z) \\ &= \int_{-\infty}^{\mathfrak{m}_z} f_\infty(u) du + \int_{\mathfrak{m}_z+z}^{\infty} f_\infty(u) du. \end{aligned} \tag{1.21}$$



Assume that  $f_\infty(u) > 0$  for all  $u \in \mathbb{R}$ . By Scheffé’s lemma for densities, see Lemma 3.3.1 in [68], the left-hand side of the preceding equality tends to zero as  $z \rightarrow \infty$ . Hence the right-hand side implies  $m_z \rightarrow -\infty$  and  $m_z + z \rightarrow \infty$ , as  $z \rightarrow \infty$ . We have

$$1 - \|(e^{-\rho}e^{-\lambda_x\tau_x}v^x + Z_\infty) - Z_\infty\|_{\text{TV}} = F_\infty(m(e^{-\rho}e^{-\lambda_x\tau_x}|v^x|)) + (1 - F_\infty(m(e^{-\rho}e^{-\lambda_x\tau_x}|v^x|) + e^{-\rho}e^{-\lambda_x\tau_x}|v^x|)),$$

which reduces in the symmetric case to

$$1 - \|(e^{-\rho}e^{-\lambda_x\tau_x}v^x + Z_\infty) - Z_\infty\|_{\text{TV}} = 2(1 - F_\infty(e^{-\rho}e^{-\lambda_x\tau_x}|v^x|/2)),$$

where  $F_\infty$  is the cumulative function of  $Z_\infty$ .

We compare the prototypical shapes of the tails of the profile functions.

I) For the *symmetric  $\alpha$ -stable* process  $L$  we obtain the exponential profile function

$$1 - \|(e^{-(\rho+\lambda_x\tau_x)}v^x + Z_\infty) - Z_\infty\|_{\text{TV}} \sim \frac{2^{\alpha+1}C_\alpha e^{(\rho+\lambda_x\tau_x)\alpha}}{|v^x|^\alpha} \propto e^{\rho\alpha}, \quad \text{as } \rho \rightarrow -\infty,$$

where  $C_\alpha$  is an explicit constant.

II) The asymptotically doubly exponential shape of the profile for the case of *Gaussian tails*  $F_\infty$  is discussed in Remark 2.3 in [5] which reads in our setting as follows

$$2(1 - F_\infty(e^{-(\rho+\lambda_x\tau_x)}|v^x|/2)) \sim \frac{4}{\sqrt{2\pi}|v^x|} \exp(-\exp(-2(\rho + \lambda_x\tau_x))|v^x|^2/8 + \rho + \lambda_x\tau_x) \tag{1.22}$$

for  $\rho \rightarrow -\infty$ . In particular, (1.22) yields the doubly exponential asymptotic behavior ( $\rho \rightarrow -\infty$ ) for the leading term

$$\frac{2\sqrt{2}}{\sqrt{\pi}|v^x|} \exp(-\exp(-2(\rho + \lambda_x\tau_x))|v^x|^2/8) \propto \exp(-K_x \exp(-2\rho))$$

for some positive  $K_x$ .

We continue with the asymptotics at zero of the profile function and show

$$\frac{\|(z + Z_\infty) - Z_\infty\|_{\text{TV}}}{|z|} \rightarrow f_\infty(m), \quad \text{as } z \rightarrow 0,$$

where  $m = \frac{a}{b'(0)}$ , where  $b$  is the vector field of (1.3) and  $(a, 0, \nu)$  is the characteristic triplet of  $L$ . By (1.21) we have

$$\|(z + Z_\infty) - Z_\infty\|_{\text{TV}} = \int_{m_z}^{m_z+z} f_\infty(u)du,$$

where  $m_z \in (m - z, m)$ . Since  $m_z \rightarrow m$  as  $z \rightarrow 0$  and  $m$  is a Lebesgue point of  $f_\infty$ , it follows

$$\frac{\|(z + Z_\infty) - Z_\infty\|_{\text{TV}}}{z} \rightarrow f_\infty(m), \quad \text{as } z \rightarrow 0.$$

As a consequence of the preceding limit we obtain

$$\lim_{\rho \rightarrow \infty} \frac{\|(e^{-(\rho+\lambda_x\tau_x)}v^x + Z_\infty) - Z_\infty\|_{\text{TV}}}{e^{-(\rho+\lambda_x\tau_x)}|v^x|} = f_\infty(m).$$

In particular, as  $\rho \rightarrow \infty$ , the profile is asymptotically proportional to the respective Wasserstein profile [6].

### 1.3.5 Counterexample to Slutsky’s lemma in total variation distance

The following example is the main motivation for Hypothesis 3. It is given for completeness since we are not aware of a reference in the literature. It is based on private communication with professors M. Jara (IMPA) and R. Imbuzeiro Oliveira (IMPA).

**Lemma 1.17.** *Let  $(U_n)_{n \in \mathbb{N}}$  be a sequence of random variable with the discrete uniform distribution supported on the set  $\{j/n : j = 1, \dots, n\}$ . Let  $(R_n)_{n \in \mathbb{N}}$  be a sequence of random variables independent of  $(U_n)_{n \in \mathbb{N}}$  with the continuous uniform distribution supported on  $[0, a_n]$ , where  $(a_n)_{n \in \mathbb{N}}$  is any sequence of positive numbers such that  $n \cdot a_n \rightarrow 0$  and  $a_n \rightarrow 0$ , as  $n \rightarrow \infty$ . For each  $n \in \mathbb{N}$ , we define  $X_n = U_n + R_n$  and  $Y_n = -R_n$ . Then we have:*

1.  $X_n$  and  $Y_n$  are absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}$ .
2.  $\lim_{n \rightarrow \infty} U_n \stackrel{d}{=} U$ , where  $U$  is (continuously) uniformly distributed on  $[0, 1]$ .
3.  $Y_n \rightarrow 0$ , as  $n \rightarrow \infty$  in probability.
4.  $\lim_{n \rightarrow \infty} \|X_n - U\|_{\text{TV}} = 0$ .
5.  $\|(X_n + Y_n) - U\|_{\text{TV}} = 1$  for all  $n \in \mathbb{N}$ .

*Proof.* Items (1), (2), (3) and (5) are straightforward. In the sequel we verify (4). Since  $U_n$  and  $R_n$  are independent, the convolution formula yields that the density of  $X_n$ ,  $f_n$ , is given by

$$\mathbb{R} \ni z \mapsto f_n(z) = \frac{1}{n} \sum_{j=1}^n g(z - j/n), \quad \text{where } g(z) = (1/a_n) \mathbf{1}_{[0, a_n]}(z).$$

First, for  $z \leq 0$ , it follows that  $f_n(z) = 0$  for all  $n \in \mathbb{N}$ . Next, for  $z > 1$  there exists  $n_0 = n_0(z) \in \mathbb{N}$  such that  $0 < a_n < z - 1$  for all  $n \geq n_0$ . Hence, for all  $j = 1, \dots, n$  we have  $0 < a_n < z - j/n$  for all  $n \geq n_0$ . Consequently,  $f_n(z) = 0$  for all  $n \geq n_0$ . We continue with the case  $z \in (0, 1]$ . Then there exists  $n_1 := n_1(z) \in \mathbb{N}$  such that  $z/2 < z - a_n < z$  for all  $n \geq n_1$ . Then we have for all  $n \geq n_1$

$$f_n(z) = \frac{1}{n} \sum_{j=\lceil n(z-a_n) \rceil}^{\lfloor nz \rfloor} g(z - j/n) = \frac{na_n}{na_n} + \frac{C_n(z)}{na_n} = 1 + \frac{C_n(z)}{na_n},$$

where  $0 < C_n(z) \leq 2$ . Since  $na_n \rightarrow \infty$ ,  $n \rightarrow \infty$ , we have  $f_n(z) \rightarrow 1$ , as  $n \rightarrow \infty$ . In summary, it is shown for all  $z \in \mathbb{R}$  that  $f_n(z) \rightarrow \mathbf{1}_{(0,1]}(z)$ , as  $n \rightarrow \infty$ . Scheffé’s lemma for densities implies  $\|X_n - U\|_{\text{TV}} \rightarrow 0$ , as  $n \rightarrow \infty$ .  $\square$

## 1.4 Global steps of the proofs of Theorem 2 and Theorem 3

The fundamental idea of the proofs of Theorem 2 and Theorem 3 is to carry out a quantitative asymptotic expansion in  $\varepsilon$  by probabilistic methods. It turns out that the hyperbolic contracting nature of the underlying deterministic dynamics  $\varphi^x$  can be used to show that the correct first order expansion of  $X^{\varepsilon, x}$  of the sense of Freidlin-Wentzell [37] Chapter 2.2 given by the inhomogeneous Ornstein-Uhlenbeck defined in (1.26) provides an asymptotic description of  $X^{\varepsilon, x}$  which is effective for time scales beyond the cutoff time scale.

### 1.4.1 Freidlin-Wentzell first order expansion

It is not hard to see that for any  $\eta > 0$  and  $t \geq 0$  the law of large numbers implies

$$\mathbb{P} \left( \sup_{0 \leq s \leq t} |X_s^{\varepsilon, x} - \varphi_s^x| \geq \eta \right) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \tag{1.23}$$

In the sequel, we analyze the asymptotic fluctuations of  $X_t^{\varepsilon,x} - \varphi_t^x$ . Let

$$Z_t^{\varepsilon,x} := \frac{X_t^{\varepsilon,x} - \varphi_t^x}{\varepsilon}, \quad t \geq 0.$$

Then the process  $(Z_t^{\varepsilon,x})_{t \geq 0}$  is the unique strong solution of the stochastic differential equation

$$\begin{cases} dZ_t^{\varepsilon,x} = -\frac{1}{\varepsilon} (b(X_t^{\varepsilon,x}) - b(\varphi_t^x)) dt + dL_t & \text{for any } t \geq 0, \\ Z_0^{\varepsilon,x} = 0. \end{cases}$$

The mean value theorem yields

$$dZ_t^{\varepsilon,x} = -\left( \int_0^1 Db(\varphi_t^x + \theta(X_t^{\varepsilon,x} - \varphi_t^x)) d\theta \right) Z_t^{\varepsilon,x} dt + dL_t \quad \text{for any } t \geq 0. \quad (1.24)$$

By construction,  $X_t^{\varepsilon,x} = \varphi_t^x + \varepsilon Z_t^{\varepsilon,x}$  for any  $t \geq 0$ . However, (1.24) has the same level of complexity as (1.6). Using (1.23) in (1.24) we derive the linear inhomogeneous approximation of (1.24) as follows. Let  $(Y_t^x)_{t \geq 0}$  be the unique strong solution of the linear inhomogeneous stochastic differential equation

$$\begin{cases} dY_t^x = -Db(\varphi_t^x) Y_t^x dt + dL_t & \text{for any } t \geq 0, \\ Y_0^x = 0. \end{cases} \quad (1.25)$$

Instead of (1.23) we claim the following stronger result, that is, the first order approximation in the sense of Section 2, Chapter 2 in [37]

$$\mathbb{P}(|X_t^{\varepsilon,x} - (\varphi_t^x + \varepsilon Y_t^x)| \geq \varepsilon^{3/2}) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0$$

for times  $t \gg t_\varepsilon^x$ , where  $t_\varepsilon^x$  is given in (1.12). For a concise quantification of the approximation, see Lemma B.1 in Appendix B. Next, we define the first order approximation

$$Y_t^\varepsilon(x) := \varphi_t^x + \varepsilon Y_t^x \quad \text{for any } t \geq 0. \quad (1.26)$$

It is not hard to see that for any  $\varepsilon$  there exists a limiting distribution  $\mu_*^\varepsilon$  such that for any  $x \in \mathbb{R}^d$ , the process  $(Y_t^\varepsilon(x))_{t \geq 0}$  converges to  $\mu_*^\varepsilon$  in the total variation distance as  $t$  tends to infinity. For further details see Lemma C.4 in Appendix C. Moreover, it is shown there that  $\mu_*^\varepsilon$  is the unique invariant distribution of the homogeneous Ornstein-Uhlenbeck process

$$dZ_t^\varepsilon = -Db(0) Z_t^\varepsilon dt + \varepsilon dL_t,$$

and has a  $C^\infty$ -density with respect to the Lebesgue measure on  $\mathbb{R}^d$ . Note that  $(Y_t^\varepsilon(x))_{t \geq 0}$  satisfies the inhomogeneous equation

$$dY_t^\varepsilon(x) = (-b(\varphi_t^x) + Db(\varphi_t^x)\varphi_t^x - Db(\varphi_t^x)Y_t^\varepsilon(x)) dt + \varepsilon dL_t, \quad Y_0^\varepsilon(x) = x. \quad (1.27)$$

Since we need to compare solutions of stochastic differential equations with different initial conditions, we introduce the following notation. Let  $\mathfrak{T}$  be a positive number and  $\xi$  be a given random vector on  $\mathbb{R}^d$ . We assume that  $\xi$  is  $\mathcal{F}_{\mathfrak{T}}$ -measurable for  $(\mathcal{F}_t)_{t \geq 0}$  defined in Subsection 1.2.2. Let  $(Y^{\varepsilon,x}(t; \mathfrak{T}, \xi))_{t \geq 0}$  be the unique strong solution of the stochastic differential equation

$$\begin{cases} dY^{\varepsilon,x}(t; \mathfrak{T}, \xi) = (-b(\varphi_{t+\mathfrak{T}}^x) + Db(\varphi_{t+\mathfrak{T}}^x)\varphi_{t+\mathfrak{T}}^x - Db(\varphi_{t+\mathfrak{T}}^x)Y^{\varepsilon,x}(t; \mathfrak{T}, \xi)) dt + \varepsilon dL_{t+\mathfrak{T}}, \\ Y^{\varepsilon,x}(0; \mathfrak{T}, \xi) = \xi. \end{cases} \quad (1.28)$$

Let  $\Delta_\varepsilon > 0$  (independently of  $x$ ) with  $\lim_{\varepsilon \rightarrow 0} \Delta_\varepsilon = 0$ . For any  $\rho \in \mathbb{R}$ , we define

$$T_\varepsilon^x := t_\varepsilon^x - \Delta_\varepsilon + \rho \cdot w_\varepsilon^x,$$

where  $t_\varepsilon^x$  and  $w_\varepsilon^x$  are given in Theorem 2. In what follows, we always take  $\mathfrak{T} = T_\varepsilon^x$ . Then  $T_\varepsilon^x > 0$  for  $0 < \varepsilon \ll 1$ .

**1.4.2 Key cutoff estimate**

The proofs of the main results Theorem 2 and Theorem 3 are based on the following fundamental inequality. On the one hand, note that for any  $\rho \in \mathbb{R}$  and  $\varepsilon$  small enough we have

$$\begin{aligned} \left\| X_{t_\varepsilon^x + \rho \cdot w_\varepsilon^x}^\varepsilon(x) - \mu^\varepsilon \right\|_{\text{TV}} &= \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - \mu^\varepsilon \right\|_{\text{TV}} \\ &\leq \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} \\ &\quad + \left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) \right\|_{\text{TV}} \\ &\quad + \left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) - \mu_*^\varepsilon \right\|_{\text{TV}} + \left\| \mu_*^\varepsilon - \mu^\varepsilon \right\|_{\text{TV}}. \end{aligned}$$

Conversely, we obtain

$$\begin{aligned} &\left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) - \mu_*^\varepsilon \right\|_{\text{TV}} \\ &\leq \left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} \\ &\quad + \left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) - X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} \\ &\quad + \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - \mu^\varepsilon \right\|_{\text{TV}} + \left\| \mu^\varepsilon - \mu_*^\varepsilon \right\|_{\text{TV}}. \end{aligned}$$

Note that  $Y^{\varepsilon,x}(\Delta_\varepsilon, T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) = Y^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x; 0, x)$ . Combining both preceding inequalities we deduce

$$\begin{aligned} &\left| \left\| X_{t_\varepsilon^x + \rho \cdot w_\varepsilon^x}^\varepsilon(x) - \mu^\varepsilon \right\|_{\text{TV}} - \left\| Y^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x; 0, x) - \mu_*^\varepsilon \right\|_{\text{TV}} \right| \\ &\leq \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} \tag{1.29} \\ &\quad + \left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) \right\|_{\text{TV}} + \left\| \mu_*^\varepsilon - \mu^\varepsilon \right\|_{\text{TV}} \\ &= E_1 + E_2 + E_3, \end{aligned}$$

where

$$\begin{aligned} E_1 &:= \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}}, \\ E_2 &:= \left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) \right\|_{\text{TV}}, \\ E_3 &:= \left\| \mu_*^\varepsilon - \mu^\varepsilon \right\|_{\text{TV}}. \end{aligned}$$

Roughly speaking, it turns out that the processes  $(X_t^\varepsilon(x))_{t \geq 0}$  and  $(Y_t^\varepsilon(x))_{t \geq 0}$  are close enough for time scales of order  $\mathcal{O}(\ln(1/\varepsilon))$  in order to carry out the following quantitative coupling procedure. Since  $(X_t^\varepsilon(x))_{t \geq 0}$  and  $(Y_t^\varepsilon(x))_{t \geq 0}$  have different (inhomogeneous) drifts, couplings which dominate the total variation distance typically only hold for short-time horizons. For an excellent introduction on the subject in the diffusive case we refer to [34]. Since the process  $(Y_t^\varepsilon(x))_{t \geq 0}$  is linear, the precise cutoff behavior (cutoff, window cutoff and profile cutoff) is derived from it in the spirit of [10]. However, it is inhomogeneous such that the results of [10] cannot be applied directly. They are adapted in Subsection 2.1. Recall that  $\mu_*^\varepsilon$  is the limiting distribution of the process  $(Y^\varepsilon(t; 0, x))_{t \geq 0}$ .

**Proposition 1** (Window and profile cutoff phenomenon for the first order approximation  $Y^\varepsilon$ ). *Assume Hypotheses 1, 2, 3 and 4 are satisfied for  $\alpha \in (0, 2)$ ,  $\beta > 0$  and  $x \in \mathbb{R}^d \setminus \{0\}$ . Let  $(Y^{\varepsilon,x})_{\varepsilon \in (0,1]}$  be the family of inhomogeneous Ornstein-Uhlenbeck processes given by  $Y^{\varepsilon,x} := (Y^\varepsilon(t; 0, x))_{t \geq 0}$  in (1.26).*

The cutoff phenomenon for nonlinear Langevin systems with small noise

1. Then  $(Y^{\varepsilon,x})_{\varepsilon \in (0,1]}$  exhibits a window cutoff phenomenon with respect to  $\mu_*^\varepsilon$  as  $\varepsilon \rightarrow 0$  at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$  given by

$$t_\varepsilon^x = \frac{1}{\lambda_x} \ln(1/\varepsilon) + \frac{\ell_x - 1}{\lambda_x} \ln(\ln(1/\varepsilon)) \quad \text{and} \quad w_\varepsilon^x = \frac{1}{\lambda_x} + o_\varepsilon(1), \quad (1.30)$$

where  $\lambda_x > 0$  and  $\ell_x \in \{1, \dots, d-1\}$  are the constants appearing in the Hartman-Grobman decomposition of Lemma 1.1.

2. Then  $(Y^{\varepsilon,x})_{\varepsilon \in (0,1]}$  exhibits a profile cutoff phenomenon with respect to  $\mu_*^\varepsilon$  as  $\varepsilon \rightarrow 0$  at the enhanced time scale  $(t_\varepsilon^x, w_\varepsilon^x)$  given by (1.30) with profile function

$$G_x(\rho) = \left\| \left( e^{-\rho} \cdot \frac{e^{-\lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} v + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} \quad \text{for any } \rho \in \mathbb{R}, v \in \omega(x),$$

where  $\tau_x$  is given in Lemma 1.1 and  $\omega(x)$  is defined in (1.13) if and only if for any  $a > 0$  the map

$$\omega(x) \ni v \mapsto \|(av + Z_\infty) - Z_\infty\|_{\text{TV}} \quad \text{is constant.}$$

The proof is given in Subsection 2.1 and relies on the Hartman-Grobman decomposition of Lemma 1.1. In what follows, we argue that the upper bound of inequality (1.29) tends to zero as  $\varepsilon \rightarrow 0$ . To be precise, we show the following.

**Proposition 2** (Error term  $E_1$ : the nonlinear short time coupling). *Assume Hypotheses 1, 2, 3 and 4 are satisfied for  $\alpha \in (3/2, 2)$  and  $\beta > 0$ . Let  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ . For any  $x \in \mathbb{R}^d$  it follows*

$$\lim_{\varepsilon \rightarrow 0} \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} = 0.$$

The complete proof can be found in Subsection 2.3 and it is based on the local limit theorem for strongly locally layered stable Lévy measures on the short-time scale  $\Delta_\varepsilon \rightarrow 0$ . The limitation of  $\alpha \in (3/2, 2)$  is due to the tail integrability of the characteristic function of  $X_{\Delta_\varepsilon}^\varepsilon(x)$ . It is of technical nature, but it seems difficult to remove.

**Proposition 3** (Error term  $E_2$ : the linear inhomogeneous coupling). *Assume Hypotheses 1, 2, 3 and 4 are satisfied for  $\alpha \in (0, 2)$  and  $\beta > 0$ . Let  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ . For any  $x \in \mathbb{R}^d$  it follows*

$$\lim_{\varepsilon \rightarrow 0} \left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) \right\|_{\text{TV}} = 0.$$

The proof is given Subsection 2.2 and relies on a version of the local limit theorem by [42] for strongly locally layered stable distributions and small times  $\Delta_\varepsilon$ .

We approximate the invariant distribution  $\mu^\varepsilon$  of  $X^\varepsilon(x)$  by the limiting distribution  $\mu_*^\varepsilon$  of the inhomogeneous Ornstein-Uhlenbeck  $Y^\varepsilon(x)$  in the total variation distance.

**Proposition 4** (Error term  $E_3$ : the equilibrium asymptotics). *Assume Hypotheses 1, 2, 3 and 4 are satisfied for  $\alpha \in (3/2, 2)$  and  $\beta > 0$ . It follows*

$$\lim_{\varepsilon \rightarrow 0} \|\mu_*^\varepsilon - \mu^\varepsilon\|_{\text{TV}} = 0.$$

The proof is given in Subsection 2.4.

**Proof of Theorem 2 and Theorem 3:** We apply Propositions 2, 3 and 4 to the key estimate (1.29) and obtain

$$\lim_{\varepsilon \rightarrow 0} \left\| \left\| X_{t_\varepsilon^x + \rho \cdot w_\varepsilon^x}^\varepsilon(x) - \mu^\varepsilon \right\|_{\text{TV}} - \left\| Y^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x; 0, x) - \mu_*^\varepsilon \right\|_{\text{TV}} \right| = 0.$$

Finally, Proposition 1 implies the main result in Theorem 2 and Theorem 3.  $\square$

## 2 The local results (Prop. 1 - 4) in the proofs of Theorem 2 and Theorem 3

### 2.1 Cutoff for the inhomogeneous linearization (Proposition 1)

#### 2.1.1 Cutoff linearization via Hartman-Grobman

By Lemma C.3 in Appendix C we see that  $\mu_*^\varepsilon$  is the distribution of  $\varepsilon Z_\infty$ , where  $Z_\infty$  is the unique invariant distribution of the homogeneous Ornstein-Uhlenbeck process  $Z = (Z_t)_{t \geq 0}$  given by

$$dZ_t = -Db(0)Z_t dt + dL_t.$$

As a consequence  $\mu_*^\varepsilon \sim \varepsilon Z_\infty$  for any  $\varepsilon \in (0, 1)$ . We start with the observation that  $Z_\infty$  is absolutely continuous. Indeed, let  $\zeta_t$  be the characteristic function of  $Z_t$  and  $\zeta_\infty$  be the characteristic function of  $Z_\infty$ . By Theorem 3.1 in Sato and Yamazato [66] we have for any  $t \geq 0$

$$|\zeta_t(\theta)| = \exp\left(\int_0^t \operatorname{Re}(\psi(e^{-Db(0)s}\theta)) ds\right), \quad \theta \in \mathbb{R}^d,$$

where

$$\operatorname{Re}(\psi(\vartheta)) = \int_{\mathbb{R}^d} (\cos(\langle \vartheta, u \rangle) - 1) \nu(du) \leq 0, \quad \vartheta \in \mathbb{R}^d.$$

Hence,  $|\zeta_\infty(\theta)| \leq |\zeta_t(\theta)|$  for all  $\theta \in \mathbb{R}^d$ . Then Item 3. in Section 4 of [10] implies that  $Z_\infty$  has a bounded  $C^\infty$ -density with respect to the Lebesgue measure on  $\mathbb{R}^d$ , where we take  $\kappa(v) = c_{\triangleleft} |v|^\alpha$  in their notation, and  $c_{\triangleleft}, \alpha$  being given in Lemma 1.7. In particular,  $Z_\infty$  is absolutely continuous on  $\mathbb{R}^d$ .

The following lemma reduces the cutoff phenomenon for the non-homogeneous linearization  $Y_t^\varepsilon(x)$  of  $X^{\varepsilon,x}$  to the homogeneous linearization  $Z$ .

**Lemma 2.1** (Elimination of the inhomogeneity in the cutoff linearization). *Let the hypotheses of Proposition 1 be satisfied for some  $x \neq 0$ ,  $\alpha \in (0, 2)$  and  $\beta > 0$ . We define*

$$d^{\varepsilon,x}(t) := \|Y_t^\varepsilon(x) - \varepsilon Z_\infty\|_{\text{TV}}$$

and

$$\tilde{D}^{\varepsilon,x}(t) := \left\| \left( \frac{(t - \tau_x)^{\ell_x - 1} e^{-\lambda_x(t - \tau_x)}}{\varepsilon} v(t - \tau_x, x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}, \quad t \geq \tau_x,$$

where  $v(t, x) = \sum_{k=1}^m e^{i\theta_x^k t} v_x^k$  and  $\lambda_x, \ell_x, \tau_x, \theta_x^1, \dots, \theta_x^m$  and  $v_x^1, \dots, v_x^m$  are the quantities given by the Hartman-Grobman decomposition in Lemma 1.1.

Then for any  $\rho \in \mathbb{R}$

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} d^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon) &= \limsup_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon) \quad \text{and} \\ \liminf_{\varepsilon \rightarrow 0} d^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon) &= \liminf_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon). \end{aligned}$$

*Proof of Lemma 2.1:* Let  $\varepsilon \in (0, 1)$ . We observe that  $Y^{\varepsilon,x}(t; 0, x) = Y_t^\varepsilon(x)$ ,  $t \geq 0$ , where  $Y_t^\varepsilon(x) = \varphi_t^x + \varepsilon Y_t^x$  and  $(Y_t^x)_{t \geq 0}$  is the unique strong solution of (1.25). Due to scale and (deterministic) shift invariance of the total variation distance given in part ii) of Lemma A.1 of [10], it follows for all  $t \geq 0$

$$\begin{aligned} d^{\varepsilon,x}(t) &= \|Y_t^\varepsilon(x) - \mu_*^\varepsilon\|_{\text{TV}} = \|Y_t^\varepsilon(x) - \varepsilon Z_\infty\|_{\text{TV}} \\ &\leq \|(\varphi_t^x + \varepsilon Y_t^x) - (\varphi_t^x + \varepsilon Z_\infty)\|_{\text{TV}} + \|(\varphi_t^x + \varepsilon Z_\infty) - \varepsilon Z_\infty\|_{\text{TV}} \\ &= \|Y_t^x - Z_\infty\|_{\text{TV}} + \underbrace{\|(\varphi_t^x/\varepsilon + Z_\infty) - Z_\infty\|_{\text{TV}}}_{=: D^{\varepsilon,x}(t)}. \end{aligned} \tag{2.1}$$

That is,  $d^{\varepsilon,x}(t) - D^{\varepsilon,x}(t) \leq \|Y_t^x - Z_\infty\|_{\text{TV}}$ . Analogously, we obtain

$$\begin{aligned} D^{\varepsilon,x}(t) &= \|(\varphi_t^x + \varepsilon Z_\infty) - \varepsilon Z_\infty\|_{\text{TV}} \\ &\leq \|(\varphi_t^x + \varepsilon Z_\infty) - (\varphi_t^x + \varepsilon Y_t^x)\|_{\text{TV}} + \|(\varphi_t^x + \varepsilon Y_t^x) - \varepsilon Z_\infty\|_{\text{TV}} \\ &= \|Y_t^x - Z_\infty\|_{\text{TV}} + d^{\varepsilon,x}(t), \end{aligned}$$

and deduce that

$$|d^{\varepsilon,x}(t) - D^{\varepsilon,x}(t)| \leq \|Y_t^x - Z_\infty\|_{\text{TV}} \quad \text{for all } t \geq 0. \quad (2.2)$$

**Remark 2.2.** Note that for  $x = 0$ ,  $\varphi_t^x = 0$  for any  $t \geq 0$  and consequently,  $D^{\varepsilon,x}(t) = 0$ . Since the right-hand side of inequality (2.2) does not depend on  $\varepsilon$  and tends to zero for  $t \rightarrow \infty$ , we have for any time scale  $(s_\varepsilon)_{\varepsilon \in (0,1)}$  such that  $s_\varepsilon \rightarrow \infty$ , as  $\varepsilon \rightarrow 0$ , the limit  $\lim_{\varepsilon \rightarrow 0} d^{\varepsilon,x}(s_\varepsilon) = 0$ . Hence the family  $(Y^{\varepsilon,x})$  does not exhibit a cutoff phenomenon for any time scale.

As a consequence we continue with  $x \neq 0$  and recall that

$$D^{\varepsilon,x}(t) = \|(\varphi_t^x/\varepsilon + Z_\infty) - Z_\infty\|_{\text{TV}}.$$

In addition, let

$$R^{\varepsilon,x}(t) := \left\| \left( \varphi_t^x/\varepsilon + Z_\infty \right) - \left( \frac{(t - \tau_x)^{\ell_x - 1} e^{-\lambda_x(t - \tau_x)}}{\varepsilon} v(t - \tau_x, x) + Z_\infty \right) \right\|_{\text{TV}}, \quad t \geq \tau_x.$$

By the triangle inequality it follows

$$\begin{aligned} D^{\varepsilon,x}(t) &\leq \left\| \left( \varphi_t^x/\varepsilon + Z_\infty \right) - \left( \frac{(t - \tau_x)^{\ell_x - 1} e^{-\lambda_x(t - \tau_x)}}{\varepsilon} v(t - \tau_x, x) + Z_\infty \right) \right\|_{\text{TV}} \\ &\quad + \left\| \left( \frac{(t - \tau_x)^{\ell_x - 1} e^{-\lambda_x(t - \tau_x)}}{\varepsilon} v(t - \tau_x, x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} = R^{\varepsilon,x}(t) + \tilde{D}^{\varepsilon,x}(t), \end{aligned}$$

and analogously  $\tilde{D}^{\varepsilon,x}(t) \leq R^{\varepsilon,x}(t) + D^{\varepsilon,x}(t)$  which yields

$$|D^{\varepsilon,x}(t) - \tilde{D}^{\varepsilon,x}(t)| \leq R^{\varepsilon,x}(t) \quad \text{for all } t \geq \tau_x. \quad (2.3)$$

Combining (2.2) and (2.3) we obtain

$$|d^{\varepsilon,x}(t) - \tilde{D}^{\varepsilon,x}(t)| \leq R^{\varepsilon,x}(t) + \|Y_t^x - Z_\infty\|_{\text{TV}} \quad \text{for all } t \geq \tau_x. \quad (2.4)$$

The limit (C.8) in Lemma C.4 in Appendix C shows that

$$\lim_{t \rightarrow \infty} \|Y_t^x - Z_\infty\|_{\text{TV}} = 0.$$

In particular, for any  $\rho \in \mathbb{R}$  we obtain

$$\lim_{\varepsilon \rightarrow 0} \left\| Y_{t_\varepsilon^x + \rho \cdot w_\varepsilon^x}^x - Z_\infty \right\|_{\text{TV}} = 0. \quad (2.5)$$

**Claim:** For any  $\rho \in \mathbb{R}$  we have

$$\lim_{\varepsilon \rightarrow 0} R^{\varepsilon,x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x) = 0. \quad (2.6)$$

First, we note that the scale and shift invariance of the total variation distance imply for all  $t \geq 0$

$$R^{\varepsilon,x}(t) = \left\| \left( \frac{(t - \tau_x)^{\ell_x - 1} e^{-\lambda_x(t - \tau_x)}}{\varepsilon} \left( \frac{e^{\lambda_x(t - \tau_x)}}{(t - \tau_x)^{\ell_x - 1}} \varphi_t^x - v(t - \tau_x, x) \right) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}.$$

Secondly, the Hartman-Grobman decomposition in Lemma 1.1 states

$$\lim_{\varepsilon \rightarrow 0} \left| \frac{e^{\lambda_x(t_\varepsilon + \rho \cdot w_\varepsilon - \tau_x)}}{(t_\varepsilon + \rho \cdot w_\varepsilon - \tau_x)^{\ell_x - 1}} \varphi_{t_\varepsilon + \rho \cdot w_\varepsilon}^x - v(t_\varepsilon + \rho \cdot w_\varepsilon - \tau_x, x) \right| = 0, \quad (2.7)$$

and the very definition of  $t_\varepsilon^x$  and  $w_\varepsilon^x$  yields

$$\lim_{\varepsilon \rightarrow 0} \frac{(t_\varepsilon^x + \rho \cdot w_\varepsilon^x - \tau_x)^{\ell_x - 1} e^{-\lambda_x(t_\varepsilon^x + \rho \cdot w_\varepsilon^x - \tau_x)}}{\varepsilon} = \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}}. \quad (2.8)$$

Combining (2.7), (2.8) and the absolute continuity of  $Z_\infty$  with the Scheffé lemma for densities implies that  $R^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x)$  tends to zero as  $\varepsilon \rightarrow 0$ . Joining (2.4), (2.5) and (2.6) yields that any cutoff phenomenon in the sense of Definition 1.8 can be read off from the simpler term  $\tilde{D}^{\varepsilon, x}$ .  $\square$

### 2.1.2 Window cutoff for the inhomogeneous O-U process (Proposition 1, Item (1))

*Proof of Proposition 1, Item (1):* By Lemma 2.1 it is enough to show the window cutoff phenomenon for  $\tilde{D}^{\varepsilon, x}$ . We observe that  $\lim_{t \rightarrow \infty} v(t, x)$  may not exist in general. Set

$$\bar{D}_\rho^x := \limsup_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x).$$

Then there exists a subsequence  $(t_{\varepsilon_j}^x + \rho \cdot w_{\varepsilon_j}^x)_{j \in \mathbb{N}}$  of  $(t_\varepsilon^x + \rho \cdot w_\varepsilon^x)_{\varepsilon \in (0, 1]}$  such that  $\varepsilon_j \rightarrow 0$  as  $j \rightarrow \infty$  and for which

$$\bar{D}_\rho^x = \lim_{j \rightarrow \infty} \tilde{D}^{\varepsilon_j, x}(t_{\varepsilon_j}^x + \rho \cdot w_{\varepsilon_j}^x).$$

Notice that the sequence  $(v(t_{\varepsilon_j}^x + \rho \cdot w_{\varepsilon_j}^x - \tau_x))_{j \in \mathbb{N}}$  is bounded by  $\sum_{k=1}^m |v^k|$ . Then the Bolzano-Weierstrass theorem yields the existence of a subsequence  $(\varepsilon_{j_n})_{n \in \mathbb{N}}$  of  $(\varepsilon_j)_{j \in \mathbb{N}}$  such that

$$\lim_{n \rightarrow \infty} v(\varepsilon_{j_n}, x) =: \hat{v}_\rho(x) \quad \text{exists.} \quad (2.9)$$

By construction  $\hat{v}_\rho(x) \in \omega(x)$ . For convenience, we set

$$C(n, x, \rho) := \frac{(t_{\varepsilon_{j_n}}^x + \rho \cdot w_{\varepsilon_{j_n}}^x - \tau_x)^{\ell_x - 1} e^{-\lambda_x(t_{\varepsilon_{j_n}}^x + \rho \cdot w_{\varepsilon_{j_n}}^x - \tau_x)}}{\varepsilon_{j_n}}. \quad (2.10)$$

Combining (2.8), (2.9) and (2.10), and using that the law of  $Z_\infty$  is absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}^d$ , Scheffé's lemma for densities implies

$$\begin{aligned} \bar{D}_\rho^x &= \lim_{n \rightarrow \infty} \tilde{D}^{\varepsilon_{j_n}, x}(t_{\varepsilon_{j_n}}^x + \rho \cdot w_{\varepsilon_{j_n}}^x) \\ &= \lim_{n \rightarrow \infty} \left\| \left( C(n, x, \rho) \cdot v(t_{\varepsilon_{j_n}}^x + \rho \cdot w_{\varepsilon_{j_n}}^x - \tau_x, x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} \\ &= \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \hat{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}. \end{aligned} \quad (2.11)$$

Analogously, we deduce

$$\underline{D}_\rho^x := \liminf_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x) = \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \check{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}, \quad (2.12)$$

where  $\check{v}_\rho(x) \in \omega(x)$ . Let  $\rho > 0$ . In the sequel we send  $\rho \rightarrow \infty$ . We observe that the upper limiting vector  $\hat{v}_\rho(x)$  depends on  $\rho$ , however, it is uniformly bounded by  $\sum_{k=1}^m |v^k|$ .



Hence,  $e^{-\rho}\hat{v}_\rho(x)$  tends to zero as  $\rho \rightarrow \infty$ . With the help of Scheffé’s lemma for densities we obtain

$$\lim_{\rho \rightarrow \infty} \bar{D}_\rho^x = 0. \tag{2.13}$$

For  $\rho < 0$ , by (1.5) we observe that  $|\check{v}_\rho(x)| \geq \liminf_{t \rightarrow \infty} |v(t, x)| > 0$ , where the right-hand side does not depend on  $\rho$ . Hence,  $e^{-\rho}|\check{v}_\rho(x)| \rightarrow \infty$  as  $\rho \rightarrow -\infty$ . A standard version of Scheffé’s lemma for densities with diverging drift (see Lemma A.3 in [10]) implies

$$\lim_{\rho \rightarrow -\infty} \underline{D}_\rho^x = 1. \tag{2.14}$$

Combining (2.13) and (2.14) shows the window cutoff limits for  $\tilde{D}^{\varepsilon, x}$  and hence the window cutoff phenomenon for the family  $(Y^\varepsilon(x))_{\varepsilon \in (0,1)}$ .  $\square$

**2.1.3 Profile cutoff for the inhomogeneous O-U process (Proposition 1, Item (2))**

*Proof of Proposition 1, Item (2):* By Lemma 2.1 it is enough to show the window cutoff phenomenon for  $\tilde{D}^{\varepsilon, x}$ . By (2.11) and (2.12) we have for any  $\rho \in \mathbb{R}$

$$\limsup_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x) = \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \hat{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}$$

and

$$\liminf_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x) = \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \check{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}},$$

where  $\hat{v}_\rho(x), \check{v}_\rho(x) \in \omega(x)$  defined in (1.13). The limit

$$\lim_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x) \text{ exists}$$

if and only if

$$\left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \hat{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} = \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \check{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}.$$

We start with the necessary condition for profile cutoff in Theorem 3. If for any  $a > 0$  the map

$$v \in \omega(x) \mapsto \left\| \left( av + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} \text{ is constant,}$$

then we have

$$\lim_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x) = \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} v + Z_\infty \right) - Z_\infty \right\|_{\text{TV}},$$

where  $v$  is any representative of  $\omega(x)$ . This yields the desired profile cutoff phenomenon for the family  $(Y^{\varepsilon, x})_{\varepsilon \in (0,1)}$ .

We continue with the sufficient condition for profile cutoff in Theorem 3. Let  $v \in \omega(x)$ , i.e. there exists a subsequence  $(t_j)_{j \in \mathbb{N}}$  such that

$$\lim_{j \rightarrow \infty} v(t_j, x) = v.$$

For any  $x \in \mathbb{R}^d$  and  $\rho \in \mathbb{R}$  consider the parametrization  $\varepsilon \mapsto t_\varepsilon^x + \rho \cdot w_\varepsilon^x - \tau_x$  and set  $t_j := t_{\varepsilon_j}^x + \rho \cdot w_{\varepsilon_j}^x - \tau_x$  for all  $j \in \mathbb{N}$ . Limit (2.8) and Scheffé’s lemma for densities imply

$$\lim_{j \rightarrow \infty} \tilde{D}^{\varepsilon_j, x}(t_{\varepsilon_j}^x + \rho \cdot w_{\varepsilon_j}^x) = \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} v + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}. \tag{2.15}$$

Since we are assuming profile cutoff, it follows

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \tilde{D}^{\varepsilon, x}(t_\varepsilon^x + \rho \cdot w_\varepsilon^x) &= \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \hat{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} \\ &= \left\| \left( \frac{e^{-\rho - \lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} \check{v}_\rho(x) + Z_\infty \right) - Z_\infty \right\|_{\text{TV}}, \end{aligned}$$

where  $\hat{v}_\rho(x), \check{v}_\rho(x) \in \omega(x)$ . That is, the function

$$v \in \omega(x) \mapsto \left\| \left( e^{-\rho} \cdot \frac{e^{-\lambda_x \tau_x}}{\lambda_x^{\ell_x - 1}} v + Z_\infty \right) - Z_\infty \right\|_{\text{TV}} \text{ is constant.} \quad \square$$

*Proof of Proposition 1:* Combining Lemma 2.1 with Subsubsection 2.1.2 and 2.1.3 yields the Item (1) and (2) of Proposition 1.  $\square$

### 2.2 Coupling for the inhomogeneous O-U processes (Proposition 3)

We keep the notation introduced in Subsection 1.4. Let  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ . For any  $\rho \in \mathbb{R}$  and  $x \neq 0$ , recall that  $T_\varepsilon^x = t_\varepsilon^x - \Delta_\varepsilon + \rho \cdot w_\varepsilon^x$ , where  $t_\varepsilon^x$  and  $w_\varepsilon^x$  are given in Theorem 2. For  $x = 0$  any time  $T_\varepsilon^x = O(|\ln(\varepsilon)|^2)$  can be taken (see Lemma 2.4). We show the following limit

$$\lim_{\varepsilon \rightarrow 0} \left\| Y^{\varepsilon, x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon, x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon, x}(T_\varepsilon^x; 0, x)) \right\|_{\text{TV}} = 0. \quad (2.16)$$

#### 2.2.1 Coupling by the local limit theorem for locally layered stable drivers

We recall that  $(\varphi_t^x)_{t \geq 0}$  is the solution of (1.3). By (1.28) and the variation of constants formula yields the explicit representation

$$\begin{aligned} Y^{\varepsilon, x}(t; T_\varepsilon^x, z) &= (\Phi_t^\varepsilon(x))^{-1} z \\ &+ (\Phi_t^\varepsilon(x))^{-1} \int_0^{\Delta_\varepsilon} \Phi_s^\varepsilon(x) (Db(\varphi_{T_\varepsilon^x+s}^x) \varphi_{T_\varepsilon^x+s}^x - b(\varphi_{T_\varepsilon^x+s}^x)) ds + \varepsilon U_\varepsilon^x, \end{aligned} \quad (2.17)$$

where  $(\Phi_t^\varepsilon(x))_{t \geq 0}$  is the solution of the matrix valued inhomogeneous differential equation

$$\frac{d}{dt} \Phi_t = \Phi_t Db(\varphi_{T_\varepsilon^x+t}^x), \quad \Phi_0 = I_d, \quad (2.18)$$

and

$$U_\varepsilon^x := (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \int_0^{\Delta_\varepsilon} \Phi_s^\varepsilon(x) dL_{T_\varepsilon^x+s}. \quad (2.19)$$

Since  $\varphi_{T_\varepsilon^x+t}^x \rightarrow 0$ , as  $\varepsilon \rightarrow 0$ ,  $U_\varepsilon^x$  resembles the respective homogeneous Ornstein-Uhlenbeck process. We claim that there exists a scale  $\gamma_\varepsilon$  (independent of  $x$ ) and a deterministic vector  $a_\varepsilon^x$  such that  $\gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x$  converges in total variation distance to an absolutely continuous random vector as  $\varepsilon \rightarrow 0$ . To be precise, we state it as Proposition 5 below.

**Remark 2.3.** Assume that the Lévy measure  $\nu$  is strongly locally layered stable in the sense of Definition 1.4 with parameters  $(\nu_0, \nu_\infty, \Lambda, q, c_0, \alpha)$ . Let  $\alpha \in (0, 2)$  and  $\beta > 0$ , where  $\alpha$  is given in Definition 1.4 and  $\beta$  is given in Hypothesis 2. It is not hard to adapt the proof of Theorem 3.1 in [42] to deduce that

$$(h^{-1/\alpha}(L_{sh} + sh\eta_{\alpha, \beta}) - sb_{\alpha, \beta})_{s \geq 0} \xrightarrow{d} S_\alpha(\Lambda_1) \text{ as } h \rightarrow 0, h > 0, \quad (2.20)$$

where  $S_\alpha(\Lambda_1)$  is a strictly  $\alpha$ -stable process with spectral density  $\Lambda_1(d\theta) = c_0(\theta)\Lambda(d\theta)$ . If in addition, we assume (1.7) and (1.8), then  $c_0$  is a symmetric function and therefore  $\Lambda_1$ ,

too. The vectors  $\eta_{\alpha,\beta}$  and  $\mathfrak{b}_{\alpha,\beta}$  are explicit and their formulas are given in the statement of Theorem 3.1 in [42]. To be precise, the authors in [42] state the stronger tail condition (3.3) on the Lévy measure  $\nu$ . However, in their proof of Theorem 3.1 in [42] which treats short-range behavior, it is only used to guarantee the following (according to their notation): for  $f$  being a bounded continuous function vanishing in a neighbourhood of the origin, and  $h > 0, \varepsilon > 0$ , that the iterated integral below is bounded independently of  $h$

$$\int_{\mathbb{S}^{d-1}} \Lambda(d\xi) \int_{\varepsilon}^{\infty} f(h^{-1/\alpha}\xi)q(r, \xi)dr.$$

In our setting of Definition 1.4, it is bounded by

$$|f|_{\infty} \int_{\mathbb{S}^{d-1}} \Lambda(d\xi) \int_{\varepsilon}^1 q(r, \xi)dr + |f|_{\infty} \nu_{\infty}(B_1^c(0)),$$

which is finite for any  $\varepsilon \in (0, 1)$ . As a consequence, (2.20) for  $L$  being strongly locally layered stable.

**Proposition 5** (Local limit theorem for the inhomogeneous O-U approximation).

Assume that  $\nu$  is a strongly locally layered stable Lévy measure in the sense of Definition 1.4 with parameters  $(\nu_0, \nu_{\infty}, \Lambda, q, c_0, \alpha)$ . Let  $\alpha \in (0, 2)$  and  $\beta > 0$ , where  $\alpha$  and  $\beta$  are given in Definition 1.4. Then for any  $K > 0$  we have

$$\lim_{\varepsilon \rightarrow 0} \sup_{|x| \leq K} \|(\gamma_{\varepsilon} U_{\varepsilon}^x + a_{\varepsilon}^x) - U\|_{\text{TV}} = 0,$$

where  $\gamma_{\varepsilon} := \Delta_{\varepsilon}^{-1/\alpha}$ , the random vector  $U$  has a symmetric  $\alpha$ -stable distribution with spectral density  $\Lambda_1(d\theta) = c_0(\theta)\Lambda(d\theta)$ , and the deterministic vector  $a_{\varepsilon}^x$  is given by

$$a_{\varepsilon}^x = \Delta_{\varepsilon}^{1-1/\alpha} \eta_{\alpha,\beta} - \mathfrak{b}_{\alpha,\beta} - \gamma_{\varepsilon} \left( \eta_{\alpha,\beta} - \frac{\mathfrak{b}_{\alpha,\beta}}{\Delta_{\varepsilon}^{1-1/\alpha}} \right) (\Phi_{\Delta_{\varepsilon}}^{\varepsilon}(x))^{-1} \int_0^{\Delta_{\varepsilon}} \Phi_s^{\varepsilon}(x) Db(\varphi_{T_{\varepsilon}^x+s}^x) s ds. \quad (2.21)$$

In particular, for  $x = 0$  we have  $\lim_{\varepsilon \rightarrow 0} \|(\gamma_{\varepsilon} U_{\varepsilon}^0 + a_{\varepsilon}^0) - U\|_{\text{TV}} = 0$ .

*Proof of Proposition 5:* By the continuity shown in Lemma C.2 in Appendix C we have for any  $\varepsilon > 0$  a point  $x_{\varepsilon} \in \mathbb{R}^d$  with  $|x_{\varepsilon}| \leq K$  such that

$$\sup_{|x| \leq K} \|(\gamma_{\varepsilon} U_{\varepsilon}^x + a_{\varepsilon}^x) - U\|_{\text{TV}} = \|(\gamma_{\varepsilon} U_{\varepsilon}^{x_{\varepsilon}} + a_{\varepsilon}^{x_{\varepsilon}}) - U\|_{\text{TV}}.$$

In the sequel, we show that the right-hand side tends to 0 as  $\varepsilon \rightarrow 0$ . For simplicity, we drop the  $\varepsilon$ -dependence of  $x_{\varepsilon}$  which is denoted by  $x$ . We stress that in the proof below the dependence of  $x$  only enters in terms of  $|x|$ , which is uniformly bounded by  $K$ .

We show the existence of the distributional limit  $\lim_{\varepsilon \rightarrow 0} (\gamma_{\varepsilon} U_{\varepsilon}^x + a_{\varepsilon}^x)$  for a suitable deterministic scale  $\gamma_{\varepsilon}$  such that  $\lim_{\varepsilon \rightarrow 0} \gamma_{\varepsilon} = \infty$  and a deterministic vector  $a_{\varepsilon}^x$ . By (2.19) and since the process  $L$  is additive, it is not hard to deduce that its characteristic function has the following shape

$$z \mapsto \mathbb{E} \left[ e^{i\langle z, U_{\varepsilon}^x \rangle} \right] = \exp \left( \int_0^{\Delta_{\varepsilon}} \psi \left( (\Phi_s^{\varepsilon}(x))^* ((\Phi_{\Delta_{\varepsilon}}^{\varepsilon}(x))^{-1})^* z \right) ds \right), \quad z \in \mathbb{R}^d.$$

The translation invariance of the Lebesgue integral in the preceding exponent implies that in distribution  $U_{\varepsilon}^x = (\Phi_{\Delta_{\varepsilon}}^{\varepsilon}(x))^{-1} \int_0^{\Delta_{\varepsilon}} \Phi_s^{\varepsilon}(x) dL_s$ . Integration by parts yields

$$\begin{aligned} U_{\varepsilon}^x &\stackrel{d}{=} L_{\Delta_{\varepsilon}} - (\Phi_{\Delta_{\varepsilon}}^{\varepsilon}(x))^{-1} \int_0^{\Delta_{\varepsilon}} \dot{\Phi}_s^{\varepsilon}(x) L_s ds \\ &= L_{\Delta_{\varepsilon}} - (\Phi_{\Delta_{\varepsilon}}^{\varepsilon}(x))^{-1} \int_0^{\Delta_{\varepsilon}} \Phi_s^{\varepsilon}(x) Db(\varphi_{T_{\varepsilon}^x+s}^x) L_s ds = J_1 - J_2, \end{aligned} \quad (2.22)$$

where

$$J_1 := L_{\Delta_\varepsilon} \quad \text{and} \quad J_2 := (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \int_0^{\Delta_\varepsilon} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x) L_s ds.$$

We start with the second term. Since  $|\varphi_t^x| \leq |x|$  for any  $t \geq 0$ , it follows that

$$\begin{aligned} & \left| J_2 + \int_0^{\Delta_\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x) \left( s \cdot \eta_{\alpha,\beta} - \frac{s}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right) ds \right| \\ & \leq \int_0^{\Delta_\varepsilon} |(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x)| |L_s + s \cdot \eta_{\alpha,\beta} - \frac{s}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta}| ds \\ & \leq C(|x|) \Delta_\varepsilon \sup_{s \in [0, \Delta_\varepsilon]} |L_s + s \cdot \eta_{\alpha,\beta} - \frac{s}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta}|, \\ & = C(|x|) \Delta_\varepsilon \sup_{s \in [0, 1]} |L_{s\Delta_\varepsilon} + s \cdot \Delta_\varepsilon \eta_{\alpha,\beta} - \frac{s}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta}|, \end{aligned} \tag{2.23}$$

where the last inequality follows from inequality (A.3) in Lemma A.3 in Appendix A. Since  $\gamma_\varepsilon = \Delta_\varepsilon^{-1/\alpha}$ , we obtain

$$\begin{aligned} & \gamma_\varepsilon \left| J_2 + \int_0^{\Delta_\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x) \left( s \cdot \eta_{\alpha,\beta} - \frac{s}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right) ds \right| \\ & \leq C(|x|) \Delta_\varepsilon \sup_{s \in [0, 1]} |\gamma_\varepsilon (L_{s\Delta_\varepsilon} + s \cdot \Delta_\varepsilon \eta_{\alpha,\beta}) - s \cdot \mathbf{b}_{\alpha,\beta}|. \end{aligned} \tag{2.24}$$

By Remark 2.3 we have

$$(\gamma_\varepsilon (L_{s\Delta_\varepsilon} + s \cdot \Delta_\varepsilon \eta_{\alpha,\beta}) - s \cdot \mathbf{b}_{\alpha,\beta})_{s \geq 0} \xrightarrow{d} \mathcal{S}_\alpha(\Lambda_1), \quad \varepsilon \rightarrow 0, \tag{2.25}$$

where  $\mathcal{S}_\alpha(\Lambda_1)$  is a symmetric  $\alpha$ -stable process with spectral measure  $\Lambda_1$ . It is well-known in the literature that the supremum norm is continuous with respect to the Skorokhod topology, see for instance Theorem 6.1 of [78]. Hence the continuous mapping theorem implies

$$\sup_{s \in [0, 1]} |\gamma_\varepsilon (L_{s\Delta_\varepsilon} + s \cdot \Delta_\varepsilon \eta_{\alpha,\beta}) - s \cdot \mathbf{b}_{\alpha,\beta}| \xrightarrow{d} \sup_{[0, 1]} |\mathcal{S}_\alpha(\Lambda_1)|, \quad \varepsilon \rightarrow 0.$$

Since  $\Delta_\varepsilon \rightarrow 0$ , Slutsky's lemma yields

$$\Delta_\varepsilon \cdot \sup_{s \in [0, 1]} |\gamma_\varepsilon (L_{s\Delta_\varepsilon} + s \cdot \Delta_\varepsilon \eta_{\alpha,\beta}) - s \cdot \mathbf{b}_{\alpha,\beta}| \xrightarrow{d} 0, \quad \varepsilon \rightarrow 0. \tag{2.26}$$

As a consequence the right-hand side of (2.24) tends to zero, as  $\varepsilon \rightarrow 0$ .

We continue with the first term  $J_1$ . Since  $J_1 = L_{\Delta_\varepsilon}$ , limit (2.25) implies

$$\lim_{\varepsilon \rightarrow 0} (\Delta_\varepsilon^{-1/\alpha} L_{\Delta_\varepsilon} + \Delta_\varepsilon^{1-1/\alpha} \eta_{\alpha,\beta} - \mathbf{b}_{\alpha,\beta}) \stackrel{d}{=} U, \tag{2.27}$$

where  $\eta_{\alpha,\beta}$  and  $\mathbf{b}_{\alpha,\beta}$  are deterministic vectors on  $\mathbb{R}^d$ , and  $U$  has a symmetric  $\alpha$ -stable distribution with spectral measure  $\Lambda_1$ . By (2.21) and (2.22) we obtain

$$\begin{aligned} \gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x &= \gamma_\varepsilon U_\varepsilon^x + \Delta_\varepsilon^{1-1/\alpha} \eta_{\alpha,\beta} - \mathbf{b}_{\alpha,\beta} \\ & \quad - \gamma_\varepsilon \int_0^{\Delta_\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x) \left( s \cdot \eta_{\alpha,\beta} - \frac{s}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right) ds \\ & \stackrel{d}{=} (\gamma_\varepsilon L_{\Delta_\varepsilon} + \Delta_\varepsilon^{1-1/\alpha} \eta_{\alpha,\beta} - \mathbf{b}_{\alpha,\beta}) \\ & \quad - \gamma_\varepsilon \left( J_2 + \int_0^{\Delta_\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x) \left( s \cdot \eta_{\alpha,\beta} - \frac{s}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right) ds \right). \end{aligned}$$

By (2.24), (2.26) and (2.27) we deduce with the help of Slutsky's lemma

$$\lim_{\varepsilon \rightarrow 0} (\gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x) \stackrel{d}{=} U, \tag{2.28}$$

where  $a_\varepsilon^x$  is given in (2.21). We stress that the dependence of  $x$  in the preceding limit only enters via  $C(|x|)$  in (2.24) and holds uniformly for  $|x| \leq K$ .

Finally, we strengthen the convergence in distribution in (2.28) to the convergence in total variation distance, using the regularity of the densities and showing their convergence in  $L^1(\mathbb{R}^d)$ . This can be carried out using the Fourier inversion formula of the explicit characteristic function of the linear process  $\gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x$  and the Orey-Masuda condition in Lemma 1.7, analogously as in the proof of Lemma C.4 in Appendix C. Since this procedure is spelt out in full detail in Lemma C.4 for the limit  $\lim_{t \rightarrow \infty} Y_t^\varepsilon(x) \stackrel{d}{=} Z_\infty$  established in Lemma C.3 in Appendix C we refrain from repeating it here.  $\square$

### 2.2.2 Proof of Proposition 3

In this subsection we establish an upper bound of

$$\left\| Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)) \right\|_{\text{TV}}$$

with the help of Proposition 5, which tends to zero as  $\varepsilon \rightarrow 0$ .

*Proof of Proposition 3:* For short, let  $z = X_{T_\varepsilon^x}^\varepsilon(x)$  and  $\tilde{z} = Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)$ . The shift and scale invariance of the total variation distance and representation (2.17) yield

$$\begin{aligned} & \|Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, z) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, \tilde{z})\|_{\text{TV}} \\ &= \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + \gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + \gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x \right) \right\|_{\text{TV}} =: I_1, \end{aligned} \tag{2.29}$$

where  $a_\varepsilon^x$  is given in (2.21) and  $\gamma_\varepsilon$  being given in Proposition 5. The triangle inequality yields

$$\begin{aligned} I_1 &\leq \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + \gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + U \right) \right\|_{\text{TV}} \\ &\quad + \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + U \right) \right\|_{\text{TV}} \\ &\quad + \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + \gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x \right) \right\|_{\text{TV}}, \end{aligned}$$

where  $U$  has a  $S_\alpha(A_1)$  distribution given in Proposition 5. The independence of the increments of  $L$  yields that  $(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z$  and  $(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z}$  are independent of  $U_\varepsilon^x$  and  $U$ , respectively. Then the cancellation property of independent shifts in the total variation distance given in Item ii) of Lemma A.2 of [10] yields

$$I_1 \leq 2 \left\| \left( \gamma_\varepsilon U_\varepsilon^x + a_\varepsilon^x \right) - U \right\|_{\text{TV}} + \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + U \right) \right\|_{\text{TV}}. \tag{2.30}$$

We prove that the right-hand side of the preceding inequality tends to zero as  $\varepsilon \rightarrow 0$ . By Proposition 5 it remains to prove that

$$\lim_{\varepsilon \rightarrow 0} \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + U \right) \right\|_{\text{TV}} = 0.$$

Let  $\mathbb{P}_\varepsilon^x(du, d\tilde{u})$  denote the joint probability measure  $\mathbb{P} \left( X_{T_\varepsilon^x}^\varepsilon(x) \in du, Y^{\varepsilon,x}(T_\varepsilon^x; 0, x) \in d\tilde{u} \right)$  and keep the notation  $z = X_{T_\varepsilon^x}^\varepsilon(x)$  and  $\tilde{z} = Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)$ . Since  $z$  and  $\tilde{z}$  are nondegenerate

and mutually dependent random variables the *shift property* for the total variation distance cannot be applied directly. Nevertheless, the Markov property and the shift invariance allow to disintegrate  $\mathbb{P}_\varepsilon^x$  as follows

$$\begin{aligned} & \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + U \right) \right\|_{\text{TV}} \\ & \leq \int A_1^\varepsilon(\zeta, \tilde{\zeta}) \mathbb{P}_\varepsilon^x(d\zeta, d\tilde{\zeta}) = \int A_2^\varepsilon(\zeta, \tilde{\zeta}) \mathbb{P}_\varepsilon^x(d\zeta, d\tilde{\zeta}), \end{aligned} \quad (2.31)$$

where

$$\begin{aligned} A_1^\varepsilon(\zeta, \tilde{\zeta}) & := \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \zeta + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{\zeta} + U \right) \right\|_{\text{TV}}, \\ A_2^\varepsilon(\zeta, \tilde{\zeta}) & := \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} (\zeta - \tilde{\zeta}) + U \right) - U \right\|_{\text{TV}}. \end{aligned}$$

We continue with the following split. For any  $\eta > 0$  we consider

$$A_2^\varepsilon(\zeta, \tilde{\zeta}) = A_2^\varepsilon(\zeta, \tilde{\zeta}) \mathbf{1}\{\gamma_\varepsilon|\zeta - \tilde{\zeta}| > \eta\varepsilon\} + A_2^\varepsilon(\zeta, \tilde{\zeta}) \mathbf{1}\{\gamma_\varepsilon|\zeta - \tilde{\zeta}| \leq \eta\varepsilon\}. \quad (2.32)$$

We start with the second term on the right-hand side of (2.32). Since the shift operator is continuous at 0 in  $L^1(\mathbb{R}^d)$  for any  $\rho > 0$ , there exists  $\eta = \eta(\rho) > 0$  such that

$$\|(h + U) - U\|_{\text{TV}} < \rho \quad \text{whenever} \quad |h| \leq \sqrt{d}\eta. \quad (2.33)$$

By Lemma A.3 in Appendix A we obtain  $|(\Phi_{\Delta_\varepsilon}^\varepsilon)^{-1}(x)| \leq \sqrt{d}$  for any  $\varepsilon \in (0, 1]$  and  $x \in \mathbb{R}^d$ , where  $|\cdot|$  denote the standard matrix 2-norm which in abuse of notation we also denote by  $|\cdot|$ . By Hypothesis 1 the event  $\{\gamma_\varepsilon|\zeta - \tilde{\zeta}| \leq \eta\varepsilon\}$  implies

$$\left| \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} (\zeta - \tilde{\zeta}) \right| \leq \sqrt{d}\eta \quad \text{for any } \varepsilon.$$

The preceding estimate implies

$$\int A_2^\varepsilon(\zeta, \tilde{\zeta}) \mathbf{1}\{\gamma_\varepsilon|\zeta - \tilde{\zeta}| \leq \eta\varepsilon\} \mathbb{P}_\varepsilon^x(d\zeta, d\tilde{\zeta}) \leq \rho \mathbb{P}_\varepsilon^x(\gamma_\varepsilon|\zeta - \tilde{\zeta}| \leq \eta\varepsilon) \leq \rho. \quad (2.34)$$

On the other hand, for any  $\eta > 0$  we have

$$\int A_2^\varepsilon(\zeta, \tilde{\zeta}) \mathbf{1}\{\gamma_\varepsilon|\zeta - \tilde{\zeta}| > \eta\varepsilon\} \mathbb{P}_\varepsilon^x(d\zeta, d\tilde{\zeta}) \leq \mathbb{P}_\varepsilon^x(\gamma_\varepsilon|\zeta - \tilde{\zeta}| > \eta\varepsilon). \quad (2.35)$$

Combining (2.31)-(2.35) we obtain

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + U \right) \right\|_{\text{TV}} \\ \leq \limsup_{\varepsilon \rightarrow 0} \mathbb{P}(\gamma_\varepsilon|z - \tilde{z}| > \eta\varepsilon) + \rho \end{aligned}$$

for any  $\rho > 0$ . Note that  $\eta$  depends on  $\rho$ . Sending  $\rho \rightarrow 0$  we obtain

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \left\| \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z + U \right) - \left( \frac{\gamma_\varepsilon}{\varepsilon} (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \tilde{z} + U \right) \right\|_{\text{TV}} \\ \leq \limsup_{\rho \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \mathbb{P}(\gamma_\varepsilon|z - \tilde{z}| > \eta\varepsilon). \end{aligned} \quad (2.36)$$

By Lemma 2.4 stated below we obtain that the upper bound on the right-hand side of (2.36) tends to zero,  $\varepsilon \rightarrow 0$ , which together with inequality (2.29), (2.30) and Proposition 5 implies (2.16).  $\square$

**Lemma 2.4.** Let  $\gamma_\varepsilon = \Delta_\varepsilon^{-1/\alpha}$ , where  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ . Then it follows

$$\limsup_{\varepsilon \rightarrow 0} \mathbb{P}(\gamma_\varepsilon |z - \tilde{z}| > \eta\varepsilon) = 0 \quad \text{for any } \eta > 0.$$

*Proof.* Let  $\eta > 0$  and recall  $\gamma_\varepsilon = \Delta_\varepsilon^{-1/\alpha}$  for some  $\alpha \in (0, 2)$  as in Proposition 5. Then we observe

$$\mathbb{P}(|z - \tilde{z}| > \eta\varepsilon/\gamma_\varepsilon) = \mathbb{P}(|X_{T_\varepsilon^x}^\varepsilon(x) - Y^{\varepsilon,x}(T_\varepsilon^x; 0, x)| > \eta\varepsilon\Delta_\varepsilon^{1/\alpha}) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \quad (2.37)$$

Since  $T_\varepsilon^x = O(|\ln(\varepsilon)|)$ , Proposition B.1 in Appendix B yields

$$\limsup_{\varepsilon \rightarrow 0} \mathbb{P}(\gamma_\varepsilon |z - \tilde{z}| > \eta\varepsilon) = 0 \quad \text{for any } \eta > 0. \quad \square$$

### 2.3 Nonlinear short-time coupling (Proposition 2)

We keep the notation introduced in Subsection 1.4. Let  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ . For any  $\rho \in \mathbb{R}$ , recall that  $T_\varepsilon^x = t_\varepsilon^x - \Delta_\varepsilon + \rho \cdot w_\varepsilon^x$ , where  $t_\varepsilon^x$  and  $w_\varepsilon^x$  are given in Theorem 2. We show the following:

$$\lim_{\varepsilon \rightarrow 0} \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} = 0.$$

Recall that  $(\varphi_t^x)_{t \geq 0}$  is the solution of (1.3). By (1.28) the variation of constants formula yields the explicit representation

$$\begin{aligned} Y^{\varepsilon,x}(t; T_\varepsilon^x, z) &= (\Phi_t^\varepsilon(x))^{-1} z \\ &+ (\Phi_t^\varepsilon(x))^{-1} \int_0^{\Delta_\varepsilon} \Phi_s^\varepsilon(x) (Db(\varphi_{T_\varepsilon^x+s}^x) \varphi_{T_\varepsilon^x+s}^x - b(\varphi_{T_\varepsilon^x+s}^x)) ds + \varepsilon U_\varepsilon^x, \end{aligned} \quad (2.38)$$

where  $(\Phi_t^\varepsilon(x))_{t \geq 0}$  is the solution of the matrix valued inhomogeneous differential equation given in (2.18) and the random vector  $U_\varepsilon^x$  is defined by (2.19). For any  $z \in \mathbb{R}^d$  we consider the unique strong solution  $(Z_t^\varepsilon(z))_{t \geq 0}$  of

$$dZ_t^\varepsilon = -Db(0)Z_t^\varepsilon dt + \varepsilon dL_t \quad \text{with } Z_0^\varepsilon = z.$$

The variation of constant formula yields the representation

$$Z_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) = \Psi_{\Delta_\varepsilon}^{-1} X_{T_\varepsilon^x}^\varepsilon(x) + \varepsilon \tilde{U}_\varepsilon^x, \quad \text{where } \tilde{U}_\varepsilon^x = (\Psi_{\Delta_\varepsilon}^{-1}) \int_0^{\Delta_\varepsilon} \Psi_s dL_{T_\varepsilon^x+s}$$

and  $\Psi_t = e^{Db(0)t}$ ,  $t \in \mathbb{R}$ . It is easily seen that  $\Psi_t^{-1} = \Psi_{-t}$ . We start with the estimate

$$\begin{aligned} &\left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} \\ &\leq \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Z_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} + \left\| Z_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} \\ &= G_1 + G_2, \end{aligned} \quad (2.39)$$

where

$$\begin{aligned} G_1 &:= \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Z_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}}, \\ G_2 &:= \left\| Z_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; T_\varepsilon^x, X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}}. \end{aligned}$$

**2.3.1 Step 1: Domination of the error term  $G_2$**

In this subsection we estimate the second term on the right-hand side of (2.39).

**Lemma 2.5.**  $G_2 \rightarrow 0$  as  $\varepsilon \rightarrow 0$ .

*Proof of Lemma 2.5:* Let

$$\Gamma_\varepsilon^x := (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \int_0^{\Delta_\varepsilon} \Phi_s^\varepsilon(x) (b(\varphi_{T_\varepsilon^x+s}^x) - Db(\varphi_{T_\varepsilon^x+s}^x)\varphi_{T_\varepsilon^x+s}^x) ds. \tag{2.40}$$

By disintegration combined with the translation and scale invariance of the total variation distance, we obtain

$$\begin{aligned} G_2 &= \left\| \left( \Psi_{\Delta_\varepsilon}^{-1} X_{T_\varepsilon^x}^\varepsilon(x) + \varepsilon \tilde{U}_\varepsilon^x \right) - \left( (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} X_{T_\varepsilon^x}^\varepsilon(x) - \Gamma_\varepsilon^x + \varepsilon U_\varepsilon^x \right) \right\|_{\text{TV}} \\ &\leq \int_{\mathbb{R}^d} \left\| \left( \Psi_{\Delta_\varepsilon}^{-1} z + \varepsilon \tilde{U}_\varepsilon^x \right) - \left( (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} z - \Gamma_\varepsilon^x + \varepsilon U_\varepsilon^x \right) \right\|_{\text{TV}} \mathbb{P}_\varepsilon^x(dz) \\ &= \int_{\mathbb{R}^d} \left\| \left( \frac{(\Psi_{\Delta_\varepsilon}^{-1} - (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1})z + \Gamma_\varepsilon^x}{\varepsilon} + \tilde{U}_\varepsilon^x \right) - U_\varepsilon^x \right\|_{\text{TV}} \mathbb{P}_\varepsilon^x(dz), \end{aligned}$$

where  $\mathbb{P}_\varepsilon^x(dz) := \mathbb{P}(X_{T_\varepsilon^x}^\varepsilon(x) \in dz)$ . By Proposition 5 there exists a random variable  $U \stackrel{d}{=} \mathcal{S}_\alpha(\Lambda_1)$  and the deterministic vector  $a_\varepsilon^x \in \mathbb{R}^d$  defined in (2.21) such that

$$\left\| (\Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^x) - U \right\|_{\text{TV}} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \tag{2.41}$$

Repeating the same argument of Proposition 5, there exist a random variable  $\tilde{U} \stackrel{d}{=} \mathcal{S}_\alpha(\Lambda_1)$  and a deterministic vector  $a_\varepsilon^0 \in \mathbb{R}^d$  given by

$$a_\varepsilon^0 = \Delta_\varepsilon^{1-1/\alpha} \eta_{\alpha,\beta} - \mathbf{b}_{\alpha,\beta} - \gamma_\varepsilon \left( \eta_{\alpha,\beta} - \frac{\mathbf{b}_{\alpha,\beta}}{\Delta_\varepsilon^{1-1/\alpha}} \right) (\Psi_{\Delta_\varepsilon})^{-1} \int_0^{\Delta_\varepsilon} \Psi_s Db(0) ds$$

such that

$$\left\| (\Delta_\varepsilon^{-1/\alpha} \tilde{U}_\varepsilon^x + a_\varepsilon^0) - \tilde{U} \right\|_{\text{TV}} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \tag{2.42}$$

We define the deterministic function

$$g_\varepsilon^x(z) = \frac{((\Psi_{\Delta_\varepsilon})^{-1} - (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1})z + \Gamma_\varepsilon^x}{\varepsilon}, \quad z \in \mathbb{R}^d \tag{2.43}$$

and the pivotal terms

$$\begin{aligned} B_0^\varepsilon(z) &:= \left\| \left( g_\varepsilon^x(z) + \tilde{U}_\varepsilon^x \right) - U_\varepsilon^x \right\|_{\text{TV}}, \\ B_1^\varepsilon(z) &:= \left\| \left( \Delta_\varepsilon^{-1/\alpha} g_\varepsilon^x(z) + \Delta_\varepsilon^{-1/\alpha} \tilde{U}_\varepsilon^x + a_\varepsilon^0 \right) - \left( \Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^0 \right) \right\|_{\text{TV}}, \\ B_2^\varepsilon(z) &:= \left\| \left( \Delta_\varepsilon^{-1/\alpha} g_\varepsilon^x(z) + \Delta_\varepsilon^{-1/\alpha} \tilde{U}_\varepsilon^x + a_\varepsilon^0 \right) - \left( \Delta_\varepsilon^{-1/\alpha} g_\varepsilon^x(z) + \tilde{U} \right) \right\|_{\text{TV}}, \\ B_3^\varepsilon(z) &:= \left\| \left( \Delta_\varepsilon^{-1/\alpha} g_\varepsilon^x(z) + \tilde{U} \right) - U \right\|_{\text{TV}}, \\ B_4^\varepsilon(z) &:= \left\| U - \left( \Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^0 \right) \right\|_{\text{TV}}. \end{aligned}$$

The scale and shift invariance of the total variation distance combined with the triangle inequality yield

$$B_0^\varepsilon(z) = B_1^\varepsilon(z) \leq B_2^\varepsilon(z) + B_3^\varepsilon(z) + B_4^\varepsilon(z). \tag{2.44}$$



**Estimate of  $B_2^\varepsilon(z)$  in (2.44).** By the cancellation property of independent increments in the total variation distance we have

$$B_2^\varepsilon(z) \leq \left\| \left( \Delta_\varepsilon^{-1/\alpha} \tilde{U}_\varepsilon^x + a_\varepsilon^0 \right) - \tilde{U}^x \right\|_{\text{TV}} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0,$$

due to (2.42). As a consequence, we have

$$\int_{\mathbb{R}^d} B_2^\varepsilon(z) \mathbb{P}_\varepsilon^x(dz) \leq \left\| \left( \Delta_\varepsilon^{-1/\alpha} \tilde{U}_\varepsilon + a_\varepsilon^0 \right) - \tilde{U} \right\|_{\text{TV}} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

**Estimate of  $B_4^\varepsilon(z)$  in (2.44).** Analogously to  $B_2^\varepsilon(z)$  we have

$$\begin{aligned} B_4^\varepsilon(z) &= \left\| \left( \Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^0 \right) - U \right\|_{\text{TV}} = \left\| \left( \Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^x + a_\varepsilon^0 \right) - \left( U + a_\varepsilon^x \right) \right\|_{\text{TV}} \\ &\leq \left\| \left( \Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^x + a_\varepsilon^0 \right) - \left( U + a_\varepsilon^0 \right) \right\|_{\text{TV}} + \left\| \left( U + a_\varepsilon^0 \right) - \left( U + a_\varepsilon^x \right) \right\|_{\text{TV}} \\ &= \left\| \left( \Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^x \right) - U \right\|_{\text{TV}} + \left\| \left( a_\varepsilon^0 - a_\varepsilon^x + U \right) - U \right\|_{\text{TV}}. \end{aligned}$$

Due to (2.41) we obtain

$$\left\| \left( \Delta_\varepsilon^{-1/\alpha} U_\varepsilon^x + a_\varepsilon^x \right) - U \right\|_{\text{TV}} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

Proposition 5 yields

$$\begin{aligned} a_\varepsilon^x - a_\varepsilon^0 &= -\Delta_\varepsilon^{-1/\alpha} \left( \eta_{\alpha,\beta} - \frac{1}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right) \\ &\quad \cdot \int_0^{\Delta_\varepsilon} s \left( (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x) - (\Psi_{\Delta_\varepsilon})^{-1} \Psi_s Db(0) \right) ds, \end{aligned}$$

such that

$$\begin{aligned} |a_\varepsilon - a_\varepsilon^0| &\leq \Delta_\varepsilon^{1-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{1}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right| \int_0^{\Delta_\varepsilon} \left| (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) Db(\varphi_{T_\varepsilon^x+s}^x) - (\Psi_{\Delta_\varepsilon})^{-1} \Psi_s Db(0) \right| ds \\ &\leq \Delta_\varepsilon^{1-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{1}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right| \int_0^{\Delta_\varepsilon} \left| (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) \right| \left| Db(\varphi_{T_\varepsilon^x+s}^x) - Db(0) \right| ds \quad (2.45) \\ &\quad + \Delta_\varepsilon^{1-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{1}{\Delta_\varepsilon^{1-1/\alpha}} \mathbf{b}_{\alpha,\beta} \right| |Db(0)| \int_0^{\Delta_\varepsilon} \left| (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) - (\Psi_{\Delta_\varepsilon})^{-1} \Psi_s \right| ds. \end{aligned}$$

We start with the estimate of the first term on the right-hand side. By Lemma A.1 in Appendix A there exists a positive constant  $C(|x|)$  depending continuously on  $|x|$  such that

$$|\varphi_{T_\varepsilon^x}^x| \leq C(|x|)\varepsilon \quad \text{for all } \varepsilon \ll 1. \quad (2.46)$$

With the help of inequality (A.3) in Lemma A.3 in Appendix A, the mean value theorem and the fact that  $|\varphi_t^x| \leq |x|$ ,  $t \geq 0$ , we have

$$\begin{aligned} &\Delta_\varepsilon^{1-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{\mathbf{b}_{\alpha,\beta}}{\Delta_\varepsilon^{1-1/\alpha}} \right| \int_0^{\Delta_\varepsilon} \left| (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} \Phi_s^\varepsilon(x) \right| \left| Db(\varphi_{T_\varepsilon^x+s}^x) - Db(0) \right| ds \\ &\leq \sqrt{d} \Delta_\varepsilon^{1-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{\mathbf{b}_{\alpha,\beta}}{\Delta_\varepsilon^{1-1/\alpha}} \right| \int_0^{\Delta_\varepsilon} \left| Db(\varphi_{T_\varepsilon^x+s}^x) - Db(0) \right| ds \\ &\leq C(|x|, d) \varepsilon \Delta_\varepsilon^{2-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{\mathbf{b}_{\alpha,\beta}}{\Delta_\varepsilon^{1-1/\alpha}} \right| \\ &\leq C(|x|, d) \varepsilon \Delta_\varepsilon^{2-1/\alpha} |\eta_{\alpha,\beta}| + C(|x|, d) \varepsilon \Delta_\varepsilon |\mathbf{b}_{\alpha,\beta}|. \end{aligned}$$

Since  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ , both preceding terms on the right-hand side tend to zero as  $\varepsilon \rightarrow 0$ .

We continue with the second term on the right-hand side of (2.45). By Lemma A.3.v) in Appendix A we have for  $\varepsilon$  sufficiently small that

$$|\Phi_{\Delta_\varepsilon}^{-1}(x)\Phi_s(x) - \Psi_{\Delta_\varepsilon}^{-1}\Psi_s| \leq \frac{C_1(|x|)d^2}{2\delta} |\varphi_{T_\varepsilon^x}^x| e^{-\frac{\delta}{2}\Delta_\varepsilon} \sqrt{1 - e^{-4\delta(\Delta_\varepsilon - s)}} \leq C_1(|x|)\varepsilon, \quad s \in [0, \Delta_\varepsilon],$$

where  $C_1(|x|)$  is a constant that depends continuously on  $|x|$ . Then for small values of  $\Delta_\varepsilon$  we have

$$\begin{aligned} \Delta_\varepsilon^{1-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{\mathbf{b}_{\alpha,\beta}}{\Delta_\varepsilon^{1-1/\alpha}} \right| & \left| Db(0) \int_0^{\Delta_\varepsilon} |(\Phi_{\Delta_\varepsilon}^\varepsilon)^{-1}\Phi_s^\varepsilon - (\Psi_{\Delta_\varepsilon})^{-1}\Psi_s| ds \right. \\ & \leq |Db(0)|C_1(|x|)\varepsilon\Delta_\varepsilon^{2-1/\alpha} \left| \eta_{\alpha,\beta} - \frac{\mathbf{b}_{\alpha,\beta}}{\Delta_\varepsilon^{1-1/\alpha}} \right| \\ & \leq |Db(0)|C_1(|x|)\varepsilon\Delta_\varepsilon^{2-1/\alpha} |\eta_{\alpha,\beta}| + |Db(0)|C_1(|x|)\varepsilon\Delta_\varepsilon |\mathbf{b}_{\alpha,\beta}|. \end{aligned}$$

Since  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ , we have  $|a_\varepsilon - a_\varepsilon^0| \rightarrow 0$ , as  $\varepsilon \rightarrow 0$  and consequently by the Scheffé lemma for densities we obtain

$$\|(a_\varepsilon - a_\varepsilon^0 + U) - U\|_{TV} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

With the same reasoning we get

$$\int_{\mathbb{R}^d} B_4^\varepsilon(z) \mathbb{P}_\varepsilon^x(dz) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

**Estimate of  $B_3^\varepsilon(z)$  in (2.44).** The remainder of Step 1 is dedicated to show that

$$\int_{\mathbb{R}^d} B_3^\varepsilon(z) \mathbb{P}_\varepsilon^x(dz) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

For  $\vartheta \in (0, 1/4)$  we define  $r_\varepsilon := \varepsilon^{1-\vartheta}$  and estimate

$$\int_{\mathbb{R}^d} B_3^\varepsilon(z) \mathbb{P}_\varepsilon^x(dz) \leq \int_{|z| \leq r_\varepsilon} B_3^\varepsilon(z) \mathbb{P}_\varepsilon^x(dz) + \mathbb{P}(|X_{T_\varepsilon^x}^\varepsilon(x)| > r_\varepsilon).$$

By Lemma D.5 in Appendix D we have for the second term

$$\mathbb{P}(|X_{T_\varepsilon^x}^\varepsilon(x)| \geq r_\varepsilon) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

We continue with the first term of the right-hand side of the preceding inequality. Recall that

$$B_3^\varepsilon(z) = \left\| (\Delta_\varepsilon^{-1/\alpha} g_\varepsilon^x(z) + \tilde{U}) - U \right\|_{TV},$$

where  $U$  and  $\tilde{U}$  are  $\mathcal{S}_\alpha(\Lambda_1)$  distributed, and

$$g_\varepsilon^x(z) = \frac{((\Psi_{\Delta_\varepsilon})^{-1} - (\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1})z + \Gamma_\varepsilon^x}{\varepsilon}, \quad z \in \mathbb{R}^d,$$

where  $\Gamma_\varepsilon^x$  was defined in (2.40). By Lemma A.3.v) in Appendix A there exists a positive constant  $C(|x|)$  depending continuously on  $|x|$  such that

$$|(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} - \Psi_{\Delta_\varepsilon}^{-1}| \leq \frac{C(|x|)d^2}{2\delta} |\varphi_{T_\varepsilon^x}^x| e^{-\frac{\delta}{2}\Delta_\varepsilon} (1 - e^{-4\delta\Delta_\varepsilon}).$$

The preceding inequality combined with inequality (2.46) yields for  $\varepsilon$  sufficiently small

$$|(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} - (\Psi_{\Delta_\varepsilon})^{-1}| \leq C_1(|x|)\varepsilon\Delta_\varepsilon.$$

Therefore,

$$\begin{aligned} & \sup_{|z| \leq r_\varepsilon} \frac{\Delta_\varepsilon^{-1/\alpha} |(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1}z - (\Psi_{\Delta_\varepsilon})^{-1}z|}{\varepsilon} \\ & \leq \sup_{|z| \leq r_\varepsilon} \frac{\Delta_\varepsilon^{-1/\alpha} |(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1} - (\Psi_{\Delta_\varepsilon})^{-1}||z|}{\varepsilon} \leq C_1(|x|)\Delta_\varepsilon^{1-1/\alpha}r_\varepsilon. \end{aligned} \quad (2.47)$$

It remains to estimate

$$\begin{aligned} |\Gamma_\varepsilon^x| &= \left| \int_0^{\Delta_\varepsilon} [(\Phi_{\Delta_\varepsilon}^\varepsilon(x))^{-1}\Phi_s^\varepsilon(x) - (\Psi_{\Delta_\varepsilon})^{-1}\Psi_s] [b(\varphi_{T_\varepsilon^x+s}^x) - Db(\varphi_{T_\varepsilon^x+s}^x)\varphi_{T_\varepsilon^x+s}^x] ds \right| \\ &\leq C(|x|, d) \int_0^{\Delta_\varepsilon} |b(\varphi_{T_\varepsilon^x+s}^x) - Db(\varphi_{T_\varepsilon^x+s}^x)\varphi_{T_\varepsilon^x+s}^x| ds \\ &\leq C_1(|x|, d) \int_0^{\Delta_\varepsilon} |\varphi_{T_\varepsilon^x+s}^x|^2 ds \leq C\varepsilon^2\Delta_\varepsilon, \end{aligned}$$

where the last inequality follows from Lemma A.1 in Appendix A. As a consequence we have

$$\frac{\Delta_\varepsilon^{-1/\alpha} |\Gamma_\varepsilon^x|}{\varepsilon} \leq C_1(|x|, d)\varepsilon\Delta_\varepsilon^{1-1/\alpha}. \quad (2.48)$$

Finally we estimate

$$\int_{|z| \leq r_\varepsilon} B_3^\varepsilon(z) \mathbb{P}_\varepsilon^x(dz) \leq \sup_{|z| \leq r_\varepsilon} B_3^\varepsilon(z).$$

The continuity of the shift operator in  $L^1$  and the compactness of the Euclidean closed ball imply

$$\sup_{|z| \leq r_\varepsilon} B_3^\varepsilon(z) = B_3^\varepsilon(z_\varepsilon) \quad \text{for some } |z_\varepsilon| \leq r_\varepsilon.$$

Since  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ , the preceding inequality combined with estimates (2.47) and (2.48) yields

$$\int_{|z| \leq r_\varepsilon} B_3^\varepsilon(z) \mathbb{P}_\varepsilon^x(dz) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

This finishes the proof of Lemma 2.5. □

### 2.3.2 Step 2: Domination of the error term $G_1$ up to a term in distribution

In the sequel we treat the error term  $G_1$  in two consecutive steps (Step 2 and Step 3). By the end of Step 3 (Subsubsection 2.3.3) we obtain the desired result  $G_1 \rightarrow 0$  as  $\varepsilon \rightarrow 0$  by a suitable localization procedure combined with the Fourier inversion technique applied to the result of Step 2.

First note that by disintegration we have

$$\begin{aligned} G_1 &= \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) - Z_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon^x}^\varepsilon(x)) \right\|_{\text{TV}} \leq \int_{\mathbb{R}^d} \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \mathbb{P}_\varepsilon^x(dz) \\ &= \int_{|z| \leq r_\varepsilon} \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \mathbb{P}_\varepsilon^x(dz) + \int_{|z| > r_\varepsilon} \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \mathbb{P}_\varepsilon^x(dz) \quad (2.49) \\ &\leq \sup_{|z| \leq r_\varepsilon} \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} + \mathbb{P}(|X_{T_\varepsilon^x}^\varepsilon(x)| \geq r_\varepsilon), \end{aligned}$$

where  $\mathbb{P}_\varepsilon^x(dz) = \mathbb{P}(X_{T_\varepsilon^x}^\varepsilon(x) \in dz)$  and  $r_\varepsilon = \varepsilon^{1-\vartheta}$ ,  $\vartheta \in (0, 1/4)$ . By Lemma D.5 in Appendix D we have

$$\mathbb{P}(|X_{T_\varepsilon^x}^\varepsilon(x)| \geq r_\varepsilon) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \quad (2.50)$$

It remains to treat the first term on the right-hand side of (2.49). The variation of constant formula yields

$$X_t^\varepsilon(z) = \Psi_t^{-1}z + \Psi_t^{-1} \int_0^t \Psi_s \tilde{b}(X_s^\varepsilon(z)) ds + \varepsilon \Psi_t^{-1} \int_0^t \Psi_s dL_s, \quad (2.51)$$

and

$$Z_t^\varepsilon(z) = \Psi_t^{-1}z + \Psi_t^{-1} \int_0^t \Psi_s dL_s, \quad (2.52)$$

where  $\tilde{b}(x) = b(x) - Db(0)x$ ,  $x \in \mathbb{R}^d$  and  $\Psi_t = e^{Db(0)t}$ . We denote

$$U_\varepsilon := \Psi_{\Delta_\varepsilon}^{-1} \int_0^{\Delta_\varepsilon} \Psi_s dL_s \quad \text{and} \quad D_\varepsilon(z) := \Psi_{\Delta_\varepsilon}^{-1} \int_0^{\Delta_\varepsilon} \Psi_s \tilde{b}(X_s^\varepsilon(z)) ds,$$

such that

$$\frac{X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)}{\varepsilon \Delta_\varepsilon^{1/\alpha}} = \left( \frac{D_\varepsilon(z)}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + \frac{1}{\Delta_\varepsilon^{1/\alpha}} U_\varepsilon + a_\varepsilon^0 - U \right) - \left( \frac{1}{\Delta_\varepsilon^{1/\alpha}} U_\varepsilon + a_\varepsilon^0 - U \right),$$

where  $a_\varepsilon^0$  is given in Proposition 5 and  $U$  is  $\mathcal{S}_\alpha(\Lambda_1)$ -distributed. By Proposition 5 we have

$$\left\| U - \left( \frac{1}{\Delta_\varepsilon^{1/\alpha}} U_\varepsilon + a_\varepsilon^0 \right) \right\|_{\text{TV}} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \quad (2.53)$$

In this subsection we show the following.

**Lemma 2.6.** *Assume Hypotheses 1, 2, 3 and 4 are satisfied for  $\alpha \in (0, 2)$  and  $\beta > 0$ . Then*

$$\frac{D_\varepsilon(z)}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + \frac{1}{\Delta_\varepsilon^{1/\alpha}} U_\varepsilon + a_\varepsilon^0 \rightarrow U, \quad \text{as } \varepsilon \rightarrow 0. \quad (2.54)$$

This convergence is strengthened to the total variation distance in Step 3 below.

*Proof of Lemma 2.6:* By (2.53) and Slutsky's lemma we have the following statement:

$$\text{If } \frac{|D_\varepsilon(z)|}{\varepsilon \Delta_\varepsilon^{1/\alpha}} \xrightarrow{\mathbb{P}} 0 \quad \text{as } \varepsilon \rightarrow 0, \quad \text{then } \frac{D_\varepsilon(z)}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + \frac{1}{\Delta_\varepsilon^{1/\alpha}} U_\varepsilon + a_\varepsilon^0 \xrightarrow{d} U \quad \text{as } \varepsilon \rightarrow 0. \quad (2.55)$$

Consequently, the remainder of the proof is dedicated to the verification of

$$\frac{|D_\varepsilon(z)|}{\varepsilon \Delta_\varepsilon^{1/\alpha}} \xrightarrow{\mathbb{P}} 0, \quad \text{as } \varepsilon \rightarrow 0. \quad (2.56)$$

For  $\eta > 0$ ,  $r_\varepsilon = \varepsilon^{1-\vartheta}$ ,  $\vartheta \in (0, 1/4)$ , and  $|z| \leq r_\varepsilon$  we have

$$\begin{aligned} \mathbb{P}\left(\frac{|D_\varepsilon(z)|}{\varepsilon \Delta_\varepsilon^{1/\alpha}} \geq \eta\right) &\leq \mathbb{P}\left(\frac{1}{\varepsilon \Delta_\varepsilon^{1/\alpha}} \int_0^{\Delta_\varepsilon} |\tilde{b}(X_s^\varepsilon(z))| ds \geq \eta\right) \\ &\leq \mathbb{P}\left(\frac{1}{\varepsilon \Delta_\varepsilon^{1/\alpha}} \int_0^{\Delta_\varepsilon} |\tilde{b}(X_s^\varepsilon(z))| ds \geq \eta, \sup_{0 \leq s \leq \Delta_\varepsilon} |X_s^\varepsilon(z)| \leq 2r_\varepsilon\right) + \mathbb{P}\left(\sup_{0 \leq s \leq \Delta_\varepsilon} |X_s^\varepsilon(z)| > 2r_\varepsilon\right). \end{aligned} \quad (2.57)$$

We start with the first term of the preceding inequality. Since  $\tilde{b} \in \mathcal{C}^2$ , there are positive constants  $C, r$  such that

$$|\tilde{b}(y)| = |b(y) - Db(0)y| \leq C|y|^2 \quad \text{for any } |y| \leq r.$$

Bearing in mind that  $r_\varepsilon \rightarrow 0$  we have

$$\begin{aligned} \mathbb{P}\left(\frac{1}{\varepsilon\Delta_\varepsilon^{1/\alpha}} \int_0^{\Delta_\varepsilon} |\tilde{b}(X_s^\varepsilon(z))| ds \geq \eta, \sup_{0 \leq s \leq \Delta_\varepsilon} |X_s^\varepsilon(z)| \leq 2r_\varepsilon\right) \\ \leq \mathbb{P}\left(\frac{4C}{\varepsilon\Delta_\varepsilon^{1/\alpha}} \Delta_\varepsilon r_\varepsilon^2 \geq \eta, \sup_{0 \leq s \leq \Delta_\varepsilon} |X_s^\varepsilon(z)| \leq 2r_\varepsilon\right) = 0, \end{aligned}$$

for all  $\varepsilon$  small enough, since the choice  $r_\varepsilon = \varepsilon^{1-\vartheta}$ ,  $\vartheta \in (0, 1/4)$  and  $\Delta_\varepsilon = \varepsilon^{\frac{\alpha}{2}}$  yields

$$\frac{1}{\varepsilon\Delta_\varepsilon^{1/\alpha}} \Delta_\varepsilon r_\varepsilon^2 \leq \Delta_\varepsilon^{1-1/\alpha} \varepsilon^{1-2\vartheta} = \varepsilon^{\alpha/2} \varepsilon^{1/2-2\vartheta} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

It remains to treat the second term on the right-hand side of (2.57). More precisely we show

$$\mathbb{P}\left(\sup_{0 \leq s \leq \Delta_\varepsilon} |X_s^\varepsilon(z)| > 2r_\varepsilon\right) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \tag{2.58}$$

By Theorem 1 in [72], we have the following almost sure estimate

$$\sup_{0 \leq s \leq \Delta_\varepsilon} |X_s^\varepsilon(z)| \leq 6\sqrt{[X^\varepsilon(z)]_{\Delta_\varepsilon}} + 2 \int_0^{\Delta_\varepsilon} H_{s-}^\varepsilon(z) dX_s^\varepsilon(z),$$

where

$$H_s^\varepsilon(z) = \frac{X_{s-}^\varepsilon(z)}{\sqrt{\sup_{0 \leq u \leq s} |X_{u-}^\varepsilon(z)|^2 + [X^\varepsilon(z)]_{s-}}}$$

Recall that by the Lévy-Itô decomposition [65], Chapter 4, the driving noise process  $(L_t)_{t \geq 0}$  under Hypotheses 3 has the following representation as Poisson random integrals

$$L_t = \int_{|z| \leq 1} z \tilde{N}(dsdz) + \int_{|z| > 1} z N(ds dz),$$

where  $N$  is the Poisson random measure associated to the Lévy measure  $\nu$  on  $\mathbb{R}^d \setminus \{0\}$  and  $\tilde{N}$  is its compensated counterpart

$$\tilde{N}([a, b] \times A) = N([a, b] \times A) - (b - a)\nu(A), \quad a < b, \quad A \in \mathcal{B}(\mathbb{R}^d).$$

In particular, we have the representation of the quadratic variation of  $X^\varepsilon$  given by

$$[X^\varepsilon(z)]_t = [L]_t = \varepsilon^2 \int_0^t \int_{|u| \leq 1} |u|^2 N(dsdu).$$

Furthermore, we have

$$\begin{aligned} \int_0^t H_{s-}^\varepsilon(z) dX_s^\varepsilon(z) &= \int_0^t \langle H_{s-}^\varepsilon(z), -b(X_s^\varepsilon(z)) \rangle ds \\ &+ \int_0^t \int_{|u| \leq 1} \langle H_{s-}^\varepsilon(z), \varepsilon u \rangle \tilde{N}(dsdu) + \int_0^t \int_{|u| > 1} \langle H_{s-}^\varepsilon(z), \varepsilon u \rangle N(dsdu). \end{aligned}$$

Since  $b(0) = 0$ , Hypothesis 1 yields

$$\int_0^t \langle H_{s-}^\varepsilon(z), -b(X_s^\varepsilon(z)) \rangle ds \leq 0, \quad \text{a.s.}$$

Hence

$$\begin{aligned}
 \mathbb{P}\left(\sup_{0 \leq s \leq \Delta_\varepsilon} |X_s^\varepsilon(z)| > 2r_\varepsilon\right) &\leq \mathbb{P}\left(6\varepsilon \sqrt{\int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} |u|^2 N(dsdu)} \right. \\
 &\quad \left. + 2 \int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} \langle H_{s-}^\varepsilon(z), u \rangle \tilde{N}(dsdu) + 2 \int_0^{\Delta_\varepsilon} \int_{|u| > 1} \langle H_{s-}^\varepsilon(z), u \rangle N(dsdu) > 2r_\varepsilon\right) \\
 &\leq \mathbb{P}\left(\varepsilon^2 \int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} |u|^2 N(dsdu) > \frac{r_\varepsilon^2}{9^2}\right) + \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} \langle H_{s-}^\varepsilon(z), \varepsilon u \rangle \tilde{N}(dsdu) > \frac{1}{3}r_\varepsilon\right) \\
 &\quad + \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| > 1} \langle H_{s-}^\varepsilon(z), \varepsilon u \rangle N(dsdu) > \frac{1}{3}r_\varepsilon\right) \\
 &= \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} |u|^2 N(dsdu) > \frac{\varepsilon^{-2\vartheta}}{9^2}\right) + \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} \langle H_{s-}^\varepsilon(z), u \rangle \tilde{N}(dsdu) > \frac{1}{3}\varepsilon^{-\vartheta}\right) \\
 &\quad + \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| > 1} \langle H_{s-}^\varepsilon(z), u \rangle N(dsdu) > \frac{1}{3}\varepsilon^{-\vartheta}\right).
 \end{aligned}$$

We continue term by term. The first term on the right side of the preceding inequality satisfies

$$\mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} |u|^2 N(dsdu) > \frac{\varepsilon^{-2\vartheta}}{9^2}\right) \leq 9^2 \Delta_\varepsilon \varepsilon^{2\vartheta} \int_{|u| \leq 1} |u|^2 \nu(du) \rightarrow 0, \text{ as } \varepsilon \rightarrow 0. \quad (2.59)$$

The second term can be estimated as follows

$$\begin{aligned}
 \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|z| \leq 1} \langle H_{s-}^\varepsilon(z), z \rangle \tilde{N}(dsdz) > \frac{1}{3}\varepsilon^{-\vartheta}\right) &\leq 9\varepsilon^{2\vartheta} \mathbb{E}\left[\left(\int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} \langle H_{s-}^\varepsilon(z), u \rangle \tilde{N}(dsdu)\right)^2\right] \\
 &= 9\varepsilon^{2\vartheta} \mathbb{E}\left[\int_0^{\Delta_\varepsilon} \int_{|u| \leq 1} \langle H_{s-}^\varepsilon(z), u \rangle^2 \nu(du) ds\right] = 9\varepsilon^{2\vartheta} \Delta_\varepsilon \int_{|u| \leq 1} |u|^2 \nu(du) \rightarrow 0, \text{ as } \varepsilon \rightarrow 0.
 \end{aligned} \quad (2.60)$$

Finally, the third term is treated as follows. For  $\beta \geq 1$  we have

$$\begin{aligned}
 \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| > 1} \langle H_{s-}^\varepsilon(z), u \rangle N(dsdu) > \frac{1}{3}\varepsilon^{-\vartheta}\right) &\leq \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| > 1} |u| N(dsdu) > \frac{1}{3}\varepsilon^{-\vartheta}\right) \\
 &\leq 3\varepsilon^\vartheta \Delta_\varepsilon \int_{|u| > 1} |u| \nu(du) \rightarrow 0, \text{ as } \varepsilon \rightarrow 0.
 \end{aligned} \quad (2.61)$$

For  $\beta \in (0, 1)$  we use the subadditivity of the root for sums of nonnegative terms (see [63]), Markov's inequality and Hypothesis 2 and obtain

$$\begin{aligned}
 \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| > 1} \langle H_{s-}^\varepsilon(z), u \rangle N(dsdu) > \frac{1}{3}\varepsilon^{-\vartheta}\right) &\leq \mathbb{P}\left(\left(\int_0^{\Delta_\varepsilon} \int_{|u| > 1} |u| N(dsdu)\right)^\beta > \frac{1}{3^\beta}\varepsilon^{-\beta\vartheta}\right) \\
 &\leq \mathbb{P}\left(\int_0^{\Delta_\varepsilon} \int_{|u| > 1} |u|^\beta N(dsdu) > \frac{1}{3^\beta}\varepsilon^{-\beta\vartheta}\right) \leq 3^\beta \varepsilon^{\beta\vartheta} \Delta_\varepsilon \int_{|u| > 1} |u|^\beta \nu(du) \rightarrow 0, \text{ as } \varepsilon \rightarrow 0.
 \end{aligned} \quad (2.62)$$

This finishes the proof of Lemma 2.6. □

### 2.3.3 Step 3: Strengthening Step 2 to $\|\cdot\|_{TV}$ by localization for $\alpha > 3/2$

In this step we prove that  $G_1 \rightarrow 0$ ,  $\varepsilon \rightarrow 0$ . More precisely, we show for  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ ,  $\alpha \in (3/2, 2)$ ,  $\beta > 0$  the following limit

$$\lim_{\varepsilon \rightarrow 0} \left\| X_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon}^\varepsilon(x)) - Z_{\Delta_\varepsilon}^\varepsilon(X_{T_\varepsilon}^\varepsilon(x)) \right\|_{TV} = 0. \quad (2.63)$$

Note that the scale and shift invariance property of the total variation imply

$$\begin{aligned} & \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \\ &= \left\| \left( \frac{X_{\Delta_\varepsilon}^\varepsilon(z) - e^{-Db(0)\Delta_\varepsilon}z}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 \right) - \left( \frac{Z_{\Delta_\varepsilon}^\varepsilon(z) - e^{-Db(0)\Delta_\varepsilon}z}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 \right) \right\|_{\text{TV}} \\ &\leq \left\| \left( \frac{X_{\Delta_\varepsilon}^\varepsilon(z) - e^{-Db(0)\Delta_\varepsilon}z}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 \right) - U \right\|_{\text{TV}} - \left\| \left( \frac{Z_{\Delta_\varepsilon}^\varepsilon(z) - e^{-Db(0)\Delta_\varepsilon}z}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 \right) - U \right\|_{\text{TV}} \end{aligned} \tag{2.64}$$

By (2.53) it remains to prove the following result.

**Proposition 6** (Nonlinear local short-time coupling). *For  $\alpha \in (3/2, 2)$  and  $\beta > 0$  it follows*

$$\sup_{|z| \leq r_\varepsilon} \left\| \left( \frac{X_{\Delta_\varepsilon}^\varepsilon(z) - e^{-Db(0)\Delta_\varepsilon}z}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 \right) - U \right\|_{\text{TV}} \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0, \tag{2.65}$$

where  $a_\varepsilon^0$  is given in Proposition 5,  $U \stackrel{d}{=} S_\alpha(\Lambda_1)$  and  $r_\varepsilon = \varepsilon^{1-\vartheta}$ ,  $\vartheta \in (0, 1/4)$ , defined below (2.56).

The proof is given after the subsequent localization results. Note that (2.54) states exactly (2.65) in distribution in a slightly different notation. In order to strengthen the result to the total variation we apply the following consecutive localization procedures to bounded jumps and a bounded vector field. The proof relies on the Plancherel isometry and Fourier inversion. We stress that the following two lemmas are true in full generality, that is, for any  $\alpha \in (0, 2)$  and  $\beta > 0$ .

**Lemma 2.7** (Jump size localization). *Let*

$$\mathcal{T}_\varepsilon := \inf\{t > 0 \mid |\varepsilon(L_t - L_{t-})| > 1\}, \quad \varepsilon \in (0, 1). \tag{2.66}$$

Then for any  $z \in \mathbb{R}^d$ ,

$$X_t^{\varepsilon, \mathcal{T}_\varepsilon}(z) := X_{t \wedge \mathcal{T}_\varepsilon}^\varepsilon(z) \quad \text{and} \quad Z_t^{\varepsilon, \mathcal{T}_\varepsilon}(z) := Z_{t \wedge \mathcal{T}_\varepsilon}^\varepsilon(z),$$

we have

$$\left| \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} - \|X_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z)\|_{\text{TV}} \right| \leq 2\mathbb{P}(\mathcal{T}_\varepsilon \leq \Delta_\varepsilon).$$

In addition,  $\beta > 0$  and Hypothesis 2 imply, for  $\varepsilon$  small enough,

$$\left| \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} - \|X_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z)\|_{\text{TV}} \right| \leq 2\Delta_\varepsilon \varepsilon^\beta. \tag{2.67}$$

*Proof.* It is well-known that  $\mathcal{T}_\varepsilon$  and  $(X_t^{\varepsilon, \mathcal{T}_\varepsilon}(z))_{t \geq 0}$  are conditionally independent (as well as  $\mathcal{T}_\varepsilon$  and  $(Z_t^{\varepsilon, \mathcal{T}_\varepsilon}(z))_{t \geq 0}$ , respectively). Hence disintegration yields

$$\begin{aligned} & \left\| X_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) \right\|_{\text{TV}} \leq \int_0^\infty \mathbb{E} \left[ \left\| X_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) \right\|_{\text{TV}} \mid \mathcal{T}_\varepsilon = s \right] \mathbb{P}(\mathcal{T}_\varepsilon \in ds) \\ &= \int_0^\infty \left\| X_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^\varepsilon(z) \right\|_{\text{TV}} \mathbb{P}(\mathcal{T}_\varepsilon \in ds) \\ &= \int_0^{\Delta_\varepsilon} \left\| X_s^\varepsilon(z) - Z_s^\varepsilon(z) \right\|_{\text{TV}} \mathbb{P}(\mathcal{T}_\varepsilon \in ds) + \left\| X_{\Delta_\varepsilon}^\varepsilon(z) - Y_{\Delta_\varepsilon}^\varepsilon(z) \right\|_{\text{TV}} \int_{\Delta_\varepsilon}^\infty \mathbb{P}(\mathcal{T}_\varepsilon \in ds) \\ &\leq \mathbb{P}(\mathcal{T}_\varepsilon \leq \Delta_\varepsilon) + \left\| X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z) \right\|_{\text{TV}}. \end{aligned}$$

On the other hand, we notice

$$\begin{aligned} & \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \\ & \leq \|X_{\Delta_\varepsilon}^\varepsilon(z) - X_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z)\|_{\text{TV}} + \|X_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z)\|_{\text{TV}} + \|Z_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \\ & \leq 2\mathbb{P}(\mathcal{T}_\varepsilon \leq \Delta_\varepsilon) + \|X_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z)\|_{\text{TV}}. \end{aligned}$$

Consequently, it follows

$$\left| \|X_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z) - Z_{\Delta_\varepsilon \wedge \mathcal{T}_\varepsilon}^{\varepsilon, \mathcal{T}_\varepsilon}(z)\|_{\text{TV}} - \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \right| \leq 2\mathbb{P}(\mathcal{T}_\varepsilon \leq \Delta_\varepsilon). \quad (2.68)$$

Finally we calculate

$$\mathbb{P}(\mathcal{T}_\varepsilon \leq \Delta_\varepsilon) = 1 - \mathbb{P}(\mathcal{T}_\varepsilon > \Delta_\varepsilon) = 1 - e^{-\Delta_\varepsilon \nu(\frac{1}{\varepsilon} B_1^c(0))}. \quad (2.69)$$

Since  $\beta > 0$ , Hypothesis 2 implies  $\lim_{r \rightarrow \infty} r^\beta \nu(r B_1^c(0)) = 0$ , which yields

$$\limsup_{\varepsilon \rightarrow 0} \Delta_\varepsilon \nu\left(\frac{1}{\varepsilon} B_1^c(0)\right) \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon^\beta \Delta_\varepsilon = 0. \quad (2.70)$$

Combining (2.68)-(2.70) we obtain (2.67).  $\square$

Since  $\mathcal{T}_\varepsilon > \Delta_\varepsilon$  with high probability, we can assume without loss of generality the presence of only bounded jumps even in the total variation distance.

**Lemma 2.8** (Spatial localization).

Let  $h \in C_b^\infty(\mathbb{R}^d, [0, 1])$  be given by

$$h(\zeta) = \begin{cases} 1 & \text{for } |\zeta| \leq 1, \\ \in (0, 1) & \text{for } 1 < |\zeta| < 2, \\ 0 & \text{for } |\zeta| \geq 2. \end{cases} \quad (2.71)$$

Consider the following localized solutions

$$\begin{aligned} d\hat{X}_t^\varepsilon(z) &= -b(\hat{X}_t^\varepsilon(z))h(\hat{X}_t^\varepsilon(z))dt + \varepsilon dL_t, & \hat{X}_0^\varepsilon(z) &= z, \\ d\hat{Z}_t^\varepsilon(z) &= -Db(0)\hat{Z}_t^\varepsilon(z)h(\hat{Z}_t^\varepsilon(z))dt + \varepsilon dL_t, & \hat{Z}_0^\varepsilon(z) &= z, \end{aligned}$$

of  $(X_t^\varepsilon)_{t \geq 0}$  and  $(Z_t^\varepsilon)_{t \geq 0}$  defined in (2.51) and (2.52), respectively. Then for  $|z| < 1/2$  we have for all  $t \geq 0$

$$\left| \|X_t^\varepsilon(z) - Z_t^\varepsilon(z)\|_{\text{TV}} - \|\hat{X}_t^\varepsilon(z) - \hat{Z}_t^\varepsilon(z)\|_{\text{TV}} \right| \leq \mathbb{P}(\hat{\tau}^\varepsilon(z) < t) + \mathbb{P}(\hat{\sigma}^\varepsilon(z) < t), \quad (2.72)$$

where

$$\hat{\tau}^\varepsilon(z) = \inf\{s \geq 0 : |X_s^\varepsilon(z)| > 1\} \quad \text{and} \quad \hat{\sigma}^\varepsilon(z) = \inf\{s \geq 0 : |Z_s^\varepsilon(z)| > 1\}.$$

In particular, we have

$$\lim_{\varepsilon \rightarrow 0} \left| \|X_{\Delta_\varepsilon}^\varepsilon(z) - Z_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} - \|\hat{X}_{\Delta_\varepsilon}^\varepsilon(z) - \hat{Z}_{\Delta_\varepsilon}^\varepsilon(z)\|_{\text{TV}} \right| = 0. \quad (2.73)$$

*Proof.* By the triangle inequality and the coupling representation of the total variation distance, we have

$$\begin{aligned} \|X_t^\varepsilon(z) - Z_t^\varepsilon(z)\|_{\text{TV}} & \leq \|X_t^\varepsilon(z) - \hat{X}_t^\varepsilon(z)\|_{\text{TV}} + \|\hat{X}_t^\varepsilon(z) - \hat{Z}_t^\varepsilon(z)\|_{\text{TV}} + \|\hat{Z}_t^\varepsilon(z) - Z_t^\varepsilon(z)\|_{\text{TV}} \\ & \leq \mathbb{P}(X_t^\varepsilon(z) \neq \hat{X}_t^\varepsilon(z)) + \|\hat{X}_t^\varepsilon(z) - \hat{Z}_t^\varepsilon(z)\|_{\text{TV}} + \mathbb{P}(Z_t^\varepsilon(z) \neq \hat{Z}_t^\varepsilon(z)) \\ & \leq \mathbb{P}(\hat{\tau}^\varepsilon(z) < t) + \|\hat{X}_t^\varepsilon(z) - \hat{Z}_t^\varepsilon(z)\|_{\text{TV}} + \mathbb{P}(\hat{\sigma}^\varepsilon(z) < t). \end{aligned}$$



Exchanging the roles of  $X_t^\varepsilon(z)$  and  $Z_t^\varepsilon(z)$  with  $\hat{X}_t^\varepsilon(z)$  and  $\hat{Z}_t^\varepsilon(z)$  yields (2.72). Since  $\Delta_\varepsilon \rightarrow 0$ , we have

$$\mathbb{P}\left(\hat{\tau}^\varepsilon(z) < \Delta_\varepsilon\right) \leq \mathbb{P}\left(\sup_{s \in [0, \Delta_\varepsilon]} |X_s^\varepsilon(z)| > 1/2\right)$$

which tends to zero due to (2.58). Analogously, the same result holds true for the linear process  $(Z_s^\varepsilon(z))_{t \geq 0}$  and  $\sigma^\varepsilon(z)$ . This implies the desired result (2.73).  $\square$

**Remark 2.9.**

1. We stress the following intentional abuse of notation. Lemma 2.7 yields that it is enough to prove (2.65) for  $X^\varepsilon$  being replaced by  $X^{\varepsilon, \mathcal{T}^\varepsilon}$ . In other words, we may assume that  $X^\varepsilon$  has bounded jumps beforehand and consequently all polynomial moments finite.
2. In addition, Lemma 2.8 allows us to consider bounded vector fields in the spirit of Section 4 in [33]. That is to say, it is enough to prove (2.65) for  $X^\varepsilon$  being replaced by  $\mathfrak{X}_t^\varepsilon(z) = \hat{X}_t^{\varepsilon, \mathcal{T}^\varepsilon}(z)$ .
3. We emphasize that due to Hypothesis 3, in particular, (1.9) the process  $(\mathfrak{X}_t^\varepsilon(z))_{t \geq 0}$  is a strong Feller process with  $C_b^1$  density  $f_\varepsilon$ . Since  $|f_\varepsilon|_\infty < \infty$ , we have  $f_\varepsilon \in L^2(\mathbb{R}^d)$ . See Theorem 1.1 and Theorem 1.3 in [73] for details.

*Proof of Proposition 6:* For simplicity we keep the same notation except for the driving noise which we denote by  $\tilde{L}$ . We set

$$\tilde{L}_t = \int_0^t \int_{|u| \leq 1} u \tilde{N}(dsdu).$$

Note that since  $\Psi_t = e^{Db(0)t}$  we have

$$\frac{\mathfrak{X}_{\Delta_\varepsilon}^\varepsilon(z) - \Psi_{\Delta_\varepsilon}^{-1}z}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 = \frac{\mathfrak{D}_\varepsilon(z)}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + \frac{1}{\Delta_\varepsilon^{1/\alpha}} \mathfrak{U}_\varepsilon + a_\varepsilon^0, \tag{2.74}$$

where

$$\mathfrak{U}_\varepsilon := \Psi_{\Delta_\varepsilon}^{-1} \int_0^{\Delta_\varepsilon} \Psi_s d\tilde{L}_s, \quad \mathfrak{D}_\varepsilon(z) := \Psi_{\Delta_\varepsilon}^{-1} \int_0^{\Delta_\varepsilon} \Psi_s \tilde{b}(\mathfrak{X}_s^\varepsilon(z)) ds$$

and  $\tilde{b}(x) = b(x)h(x) - Db(0)x$ ,  $x \in \mathbb{R}^d$ , where  $h$  is given in (2.71). Note that the limit (2.56) is shown for  $X^\varepsilon$ . It is easily seen – going through the proof of Lemma 2.6 line by line – that with the help of (2.74) the limit (2.54) remains valid for  $X^\varepsilon$  being replaced by  $\mathfrak{X}^\varepsilon$ , i.e.,

$$\frac{\mathfrak{X}_{\Delta_\varepsilon}^\varepsilon(z) - \Psi_{\Delta_\varepsilon}^{-1}z}{\varepsilon \Delta_\varepsilon^{\frac{1}{\alpha}}} + a_\varepsilon^0 \xrightarrow{d} \mathcal{S}_\alpha(\Lambda_1), \quad \text{as } \varepsilon \rightarrow 0. \tag{2.75}$$

Recall

$$\mathfrak{X}_t^\varepsilon(z) = \Psi_t^{-1}z - \Psi_t^{-1} \int_0^t \Psi_s \tilde{b}(\mathfrak{X}_s^\varepsilon(z)) ds + \varepsilon \Psi_t^{-1} \int_0^t \Psi_s d\tilde{L}_s,$$

and set  $\mathcal{X}_t^\varepsilon(z) = \mathfrak{X}_t^\varepsilon(z) - \Psi_t^{-1}z$  which satisfies

$$\mathcal{X}_t^\varepsilon(z) = -\Psi_t^{-1} \int_0^t \Psi_s \tilde{b}(\mathcal{X}_s^\varepsilon(z) + \Psi_s^{-1}z) ds + \varepsilon \Psi_t^{-1} \int_0^t \Psi_s d\tilde{L}_s. \tag{2.76}$$

In the sequel, we strengthen the convergence of (2.75) to the convergence in the total variation distance. Since  $\tilde{L}$  has absolutely continuous marginals and  $\mathfrak{X}_{\Delta_\varepsilon}^\varepsilon$  is a continuous

push-forward of  $\tilde{L}$ , it retains the absolute continuity property. In addition, it is not hard to see that Lemma 1.7 yields a  $C^\infty$ -density for  $\mathcal{S}_\alpha(\Lambda_1)$ . Hence it is enough to prove

$$\int_{\mathbb{R}^d} |f_\varepsilon(u) - f_0(u)| du \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0,$$

where  $f_\varepsilon$  is the density of  $\mathcal{X}_{\Delta_\varepsilon}^\varepsilon(z)/(\varepsilon\Delta_\varepsilon^{\frac{1}{\alpha}}) + a_\varepsilon$  and  $f_0$  is the density of  $\mathcal{S}_\alpha(\Lambda_1)$ . By Scheffé's lemma for densities it is sufficient show that  $f_\varepsilon \rightarrow f_0$ , as  $\varepsilon \rightarrow 0$ , Lebesgue almost everywhere in  $\mathbb{R}^d$ . For this sake, it is sufficient to prove that

$$\int_{\mathbb{R}^d} |f_\varepsilon(u) - f_0(u)|^2 du \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

Since  $f_\varepsilon, f_0 \in L^2(\mathbb{R}^d)$ , by Plancherel's identity we have a positive constant  $C_\pi$  such that

$$\int_{\mathbb{R}^d} |f_\varepsilon(u) - f_0(u)|^2 du = C_\pi \int_{\mathbb{R}^d} |\hat{f}_\varepsilon(\theta) - \hat{f}_0(\theta)|^2 d\theta.$$

Since the weak convergence (2.75) implies that  $\hat{f}_\varepsilon \rightarrow \hat{f}_0$  uniformly on compacts, we have for any  $K > 0$

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} |\hat{f}_\varepsilon(\theta) - \hat{f}_0(\theta)|^2 d\theta &\leq \limsup_{\varepsilon \rightarrow 0} \int_{|\theta| > K} |\hat{f}_\varepsilon(\theta) - \hat{f}_0(\theta)|^2 d\theta \\ &\leq 2 \limsup_{\varepsilon \rightarrow 0} \int_{|\theta| > K} |\hat{f}_\varepsilon(\theta)|^2 d\theta + 2 \int_{|\theta| > K} |\hat{f}_0(\theta)|^2 d\theta. \end{aligned}$$

The exponential decay of  $\hat{f}_0$  yields  $\hat{f}_0 \in L^2(\mathbb{R}^d)$ . Sending  $K$  to infinity we deduce that

$$\limsup_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} |\hat{f}_\varepsilon(\theta) - \hat{f}_0(\theta)|^2 d\theta \leq 2 \limsup_{K \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \int_{|\theta| > K} |\hat{f}_\varepsilon(\theta)|^2 d\theta. \quad (2.77)$$

In order to conclude, it remains to show that the right-hand side of the preceding inequality is 0. Recall the differential version of (2.76)

$$d\mathcal{X}_t^\varepsilon(z) = -Db(0)\mathcal{X}_t^\varepsilon(z)dt - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z)dt + \varepsilon d\tilde{L}_t$$

with initial datum  $\mathcal{X}_0^\varepsilon(z) = 0$ . In the sequel, we calculate  $\phi_t(\theta) := \mathbb{E}[e^{i\langle \theta, \mathcal{X}_t^\varepsilon(z) \rangle}]$ . Itô's formula yields

$$\begin{aligned} \exp(i\langle \theta, \mathcal{X}_t^\varepsilon(z) \rangle) &= 1 + \int_0^t \exp(i\langle \theta, \mathcal{X}_s^\varepsilon(z) \rangle) i\langle \theta, -Db(0)\mathcal{X}_s^\varepsilon(z) - \tilde{b}(\mathcal{X}_s^\varepsilon(z) + \Psi_s^{-1}z) \rangle ds \\ &\quad + \int_0^t \int_{|z| \leq 1} \left( \exp(i\langle \theta, \mathcal{X}_{s-}^\varepsilon(z) + \varepsilon u \rangle) - \exp(i\langle \theta, \mathcal{X}_{s-}^\varepsilon(z) \rangle) \right) \tilde{N}(dsdu) \\ &\quad + \int_0^t \exp(i\langle \theta, \mathcal{X}_{s-}^\varepsilon(z) \rangle) \int_{|z| \leq 1} \left( \exp(i\langle \theta, \varepsilon u \rangle) - 1 - i\langle \theta, \varepsilon u \rangle \right) \nu(du) ds. \end{aligned}$$

Since the process  $\mathfrak{X}^\varepsilon$  has finite first moment, taking expectation and using Fubini's theorem we obtain

$$\begin{aligned} \phi_t(\theta) &= \mathbb{E} \left[ \exp(i\langle \theta, \mathcal{X}_t^\varepsilon(z) \rangle) \right] \\ &= 1 + \int_0^t \mathbb{E} \left[ \exp(i\langle \theta, \mathcal{X}_s^\varepsilon(z) \rangle) i\langle \theta, -Db(0)\mathcal{X}_s^\varepsilon(z) - \tilde{b}(\mathcal{X}_s^\varepsilon(z) + \Psi_s^{-1}z) \rangle \right] ds \\ &\quad + \int_0^t \mathbb{E} \left[ \exp(i\langle \theta, \mathcal{X}_{s-}^\varepsilon(z) \rangle) \int_{|z| \leq 1} \left( \exp(i\langle \theta, \varepsilon u \rangle) - 1 - i\langle \theta, \varepsilon u \rangle \right) \nu(du) \right] ds. \end{aligned}$$

Note that

$$\psi(\varepsilon\theta) = \int_{\mathbb{R}^d} \left( \exp(i\langle\theta, \varepsilon u\rangle) - 1 - i\langle\theta, \varepsilon u\rangle \right) \nu(du),$$

where  $\psi$  is the characteristic exponent of the Lévy measure  $\nu$ . We set  $\theta_\varepsilon := \theta/(\varepsilon\Delta_\varepsilon^{\frac{1}{\alpha}})$  for  $|\theta| \geq K$ . For the real and the imaginary part of  $\phi_t(\theta)$  we have the equalities

$$\begin{aligned} & \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \\ &= 1 - \int_0^t \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_s^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_s^\varepsilon(z) - \tilde{b}(\mathcal{X}_s^\varepsilon(z) + \Psi_s^{-1}z) \rangle \right] ds \\ & \quad + \operatorname{Re} \psi(\varepsilon\theta_\varepsilon) \int_0^t \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_s^\varepsilon(z) \rangle \right) \right] ds - \operatorname{Im} \psi(\varepsilon\theta_\varepsilon) \int_0^t \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_s^\varepsilon(z) \rangle \right) \right] ds \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \\ &= 1 + \int_0^t \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_s^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_s^\varepsilon(z) - \tilde{b}(\mathcal{X}_s^\varepsilon(z) + \Psi_s^{-1}z) \rangle \right] ds \\ & \quad + \operatorname{Re} \psi(\varepsilon\theta_\varepsilon) \int_0^t \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_s^\varepsilon(z) \rangle \right) \right] ds + \operatorname{Im} \psi(\varepsilon\theta_\varepsilon) \int_0^t \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_s^\varepsilon(z) \rangle \right) \right] ds. \end{aligned}$$

The chain rule for the respective differential forms reads as follows

$$\begin{aligned} & \frac{d}{dt} \left( \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \right)^2 = 2\mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \frac{d}{dt} \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \\ &= -2\mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \rangle \right] \\ & \quad + 2\operatorname{Re} \psi(\varepsilon\theta_\varepsilon) \left( \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \right)^2 \\ & \quad - 2\operatorname{Im} \psi(\varepsilon\theta_\varepsilon) \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right], \end{aligned}$$

with  $\left( \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_0^\varepsilon(z) \rangle \right) \right] \right)^2 = 1$  and

$$\begin{aligned} & \frac{d}{dt} \left( \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \right)^2 = 2\mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \frac{d}{dt} \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \\ &= 2\mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \rangle \right] \\ & \quad + 2\operatorname{Re} \psi(\varepsilon\theta_\varepsilon) \left( \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \right)^2 \\ & \quad + 2\operatorname{Im} \psi(\varepsilon\theta_\varepsilon) \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right], \end{aligned}$$

with  $\left( \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_0^\varepsilon(z) \rangle \right) \right] \right)^2 = 0$ . We sum up the preceding equations and obtain

$$\begin{aligned} & \frac{d}{dt} |\phi_t(\theta_\varepsilon)|^2 = 2\operatorname{Re} \psi(\varepsilon\theta_\varepsilon) |\phi_t(\theta_\varepsilon)|^2 \\ & \quad - 2\mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \rangle \right] \\ & \quad + 2\mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \rangle \right]. \end{aligned}$$

We start with the first term on the right-hand side. By  $|\theta| \geq K$  and Lemma 1.7 we have

for small  $\varepsilon$  (where the smallness of  $\varepsilon$  only depends of  $K$  and  $C_{\triangleleft}$ ) the estimate

$$\begin{aligned} \operatorname{Re} \psi(\varepsilon\theta_\varepsilon) &= \int_{|u| \leq 1} \left( \cos \left( \left\langle \frac{\theta}{\Delta_\varepsilon^{1/\alpha}}, u \right\rangle \right) - 1 \right) \nu(du) \\ &\leq \int_{\substack{|\langle \frac{\theta}{\Delta_\varepsilon^{1/\alpha}}, u \rangle| \leq 1 \\ |u| \leq 1}} \left( \cos \left( \left\langle \frac{\theta}{\Delta_\varepsilon^{1/\alpha}}, u \right\rangle \right) - 1 \right) \nu(du) \\ &\leq -\frac{2}{\pi^2} \int_{\substack{|\langle \frac{\theta}{\Delta_\varepsilon^{1/\alpha}}, u \rangle| \leq 1 \\ |u| \leq 1}} \left| \left\langle \frac{\theta}{\Delta_\varepsilon^{1/\alpha}}, u \right\rangle \right|^2 \nu(du) \leq -\frac{2}{\pi^2} c_{\triangleleft} \left| \frac{\theta}{\Delta_\varepsilon^{1/\alpha}} \right|^\alpha = -C \frac{|\theta|^\alpha}{\Delta_\varepsilon}. \end{aligned}$$

We continue with the second term. Recall that  $\alpha \in (3/2, 2)$ . By the Cauchy-Schwarz inequality and the classical Young inequality for  $p = p^* = 2$  we have

$$\begin{aligned} &\left| -2\mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \rangle \right] \right| \\ &\leq 2 \left| \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \right| |\varepsilon\theta_\varepsilon|^{3/4} \frac{|\theta_\varepsilon|^{1/4}}{\varepsilon^{3/4}} \mathbb{E} \left[ \left| -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \right| \right] \\ &\leq |\varepsilon\theta_\varepsilon|^{3/2} \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right]^2 + \frac{|\theta_\varepsilon|^{1/2}}{\varepsilon^{3/2}} \mathbb{E} \left[ \left| -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \right| \right]^2. \end{aligned}$$

For the third term we infer analogously

$$\begin{aligned} &\left| 2\mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right] \mathbb{E} \left[ \cos \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \langle \theta_\varepsilon, -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \rangle \right] \right| \\ &\leq |\varepsilon\theta_\varepsilon|^{3/2} \mathbb{E} \left[ \sin \left( \langle \theta_\varepsilon, \mathcal{X}_t^\varepsilon(z) \rangle \right) \right]^2 + \frac{|\theta_\varepsilon|^{1/2}}{\varepsilon^{3/2}} \mathbb{E} \left[ \left| -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \right| \right]^2. \end{aligned}$$

Since  $\alpha \in (3/2, 2)$  we obtain for sufficiently small  $\varepsilon$  and  $|\theta| \geq K$

$$\operatorname{Re} \psi(\varepsilon\theta_\varepsilon) + |\varepsilon\theta_\varepsilon|^{3/2} \leq -2C|\varepsilon\theta_\varepsilon|^\alpha + |\varepsilon\theta_\varepsilon|^{3/2} \leq -C \frac{|\theta|^\alpha}{\Delta_\varepsilon}.$$

Therefore for  $|\theta| \geq K$  and  $\varepsilon$  small enough we have the following differential inequality

$$\begin{aligned} \frac{d}{dt} |\phi_t(\theta_\varepsilon)|^2 &\leq \left( \operatorname{Re} \psi(\varepsilon\theta_\varepsilon) + |\varepsilon\theta_\varepsilon|^{3/2} \right) |\phi_t(\theta_\varepsilon)|^2 \\ &\quad + 2 \frac{|\theta_\varepsilon|^{1/2}}{\varepsilon^{3/2}} \mathbb{E} \left[ \left| -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \right| \right]^2 \\ &\leq -C \frac{|\theta|^\alpha}{\Delta_\varepsilon} |\phi_t(\theta_\varepsilon)|^2 + 2 \frac{|\theta_\varepsilon|^{1/2}}{\varepsilon^{3/2}} \mathbb{E} \left[ \left| -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \right| \right]^2. \end{aligned}$$

In the sequel, we dominate the term  $\mathbb{E} \left[ \left| -Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \right| \right]^2$ . Recall the definition of  $\tilde{b}(x) = b(x)h(x) - Db(0)x$ ,  $x \in \mathbb{R}^d$ , where  $h$  is given in (2.71). Note that

$$\begin{aligned} &-Db(0)\mathcal{X}_t^\varepsilon(z) - \tilde{b}(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \\ &= -Db(0)\mathcal{X}_t^\varepsilon(z) - b(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z)h(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) + Db(0)(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) \\ &= -b(\mathcal{X}_t^\varepsilon(z) + Db(0)\Psi_t^{-1}z)h(\mathcal{X}_t^\varepsilon(z) + \Psi_t^{-1}z) + Db(0)\Psi_t^{-1}z. \end{aligned}$$

Since  $\mathcal{X}_t^\varepsilon(z) = \mathfrak{X}_t^\varepsilon(z) - \Psi_t^{-1}z$ , (2.58) implies

$$\lim_{\varepsilon \rightarrow 0} \mathbb{P} \left( \sup_{s \in [0, \Delta_\varepsilon]} |\mathcal{X}_s^\varepsilon(z) + Db(0)\Psi_s^{-1}z| > 2r_\varepsilon \right) = \lim_{\varepsilon \rightarrow 0} \mathbb{P} \left( \sup_{s \in [0, \Delta_\varepsilon]} |\mathfrak{X}_s^\varepsilon(z)| > 2r_\varepsilon \right) = 0$$

for  $|z| \leq r_\varepsilon$  for sufficiently small  $\varepsilon$ . Taylor's theorem combined with the jump size and spatial localizations yields for  $\varepsilon$  sufficiently small

$$\sup_{s \in [0, \Delta_\varepsilon]} \mathbb{E} \left[ \left| -Db(0)\mathcal{X}_s^\varepsilon(z) - \tilde{b}(\mathcal{X}_s^\varepsilon(z) + \Psi_s^{-1}z)h(\mathcal{X}_s^\varepsilon(z) + \Psi_s^{-1}z) \right| \right]^2 \leq C_2 r_\varepsilon^2,$$

where  $C_2 > 0$  only depends  $\max_{|u| \leq 2} |b(u)|$ ,  $\max_{|u| \leq 2} |Db(u)|$  and  $\max_{|u| \leq 2} |D^2b(u)|$ . Hence the variation of constants formula yields

$$\begin{aligned} |\phi_{\Delta_\varepsilon}(\theta_\varepsilon)|^2 &\leq e^{-C\frac{|\theta|^\alpha}{\Delta_\varepsilon}\Delta_\varepsilon} 1 + C_2\varepsilon^{2(1-\vartheta)} \frac{|\theta_\varepsilon|^{1/2}}{\varepsilon^{3/2}} e^{-C\frac{|\theta|^\alpha}{\Delta_\varepsilon}\Delta_\varepsilon} \int_0^{\Delta_\varepsilon} e^{C\frac{|\theta|^\alpha}{\Delta_\varepsilon}s} ds \\ &= e^{-C|\theta|^\alpha} + C_3 \frac{\Delta_\varepsilon^{1-\frac{1}{2\alpha}}}{\varepsilon^{-2\vartheta}|\theta|^{\alpha-\frac{1}{2}}} \left(1 - e^{-C\frac{|\theta|^\alpha}{\Delta_\varepsilon}\Delta_\varepsilon}\right) \leq e^{-C|\theta|^\alpha} + C_3 \frac{\Delta_\varepsilon^{1-\frac{1}{2\alpha}}}{\varepsilon^{-2\vartheta}|\theta|^{\alpha-\frac{1}{2}}}. \end{aligned} \quad (2.78)$$

Note that the shift  $a_\varepsilon^0$  does not change the modulus of  $\hat{f}_\varepsilon(\theta)$  and hence the integrability in  $\theta$ . The parameter value  $\alpha \in (3/2, 2)$  implies that

$$\int_{|\theta|>K} |\hat{f}_\varepsilon(\theta)|^2 d\theta \leq \int_{|\theta|>K} e^{-C|\theta|^\alpha} d\theta + 2C_3 \frac{\Delta_\varepsilon^{1-\frac{1}{2\alpha}}}{\varepsilon^{-2\vartheta}} \int_{|\theta|>K} \frac{1}{|\theta|^{\alpha-\frac{1}{2}}} d\theta < \infty.$$

Since  $\Delta_\varepsilon = \varepsilon^{\frac{\alpha}{2}}$  we have the desired limit (2.77) for any  $|z| \leq r_\varepsilon$

$$\lim_{K \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \int_{|\theta|>K} |\hat{f}_\varepsilon(\theta)|^2 d\theta \leq \lim_{K \rightarrow \infty} \int_{|\theta|>K} e^{-C|\theta|^\alpha} d\theta = 0.$$

In order to see the uniformity we refer to the continuity of the map

$$z \mapsto \left\| \frac{X_{\Delta_\varepsilon}^\varepsilon(z) - \Psi_{\Delta_\varepsilon}^{-1}z}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 - U \right\|_{\text{TV}}.$$

That is, the supremum is taken at some value  $z_\varepsilon$  such that

$$\sup_{|z| \leq r_\varepsilon} \left\| \frac{X_{\Delta_\varepsilon}^\varepsilon(z) - \Psi_{\Delta_\varepsilon}^{-1}z}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 - U \right\|_{\text{TV}} = \left\| \frac{X_{\Delta_\varepsilon}^\varepsilon(z_\varepsilon) - \Psi_{\Delta_\varepsilon}^{-1}z_\varepsilon}{\varepsilon\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^0 - U \right\|_{\text{TV}}.$$

In the previous calculation the only property of  $z$  we use is that  $|z| \leq r_\varepsilon$ . Hence all previous results remain valid for  $z$  being replaced by  $z_\varepsilon$ . This finishes the proof of Proposition 6.  $\square$

*Proof of Proposition 2:* The proof consists of the domination of the error terms  $G_1$  and  $G_2$  in (2.39). The result of Subsubsection 2.3.1 is the convergence  $G_2 \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . The term  $G_1$  is estimated by inequality (2.49) whose right-hand side is dominated by the terms given in (2.50), Proposition 5 and Proposition 6, all of which tend to 0 as  $\varepsilon \rightarrow 0$ . This finishes the proof of Proposition 2.  $\square$

## 2.4 Inhomogeneous O-U approximation of the limiting distribution (Prop. 4)

*Proof of Proposition 4:* Let  $x_0 \in \mathbb{R}^d$  and  $t > 0$ . The triangle inequality yields

$$\begin{aligned} \|\mu^\varepsilon - \mu_*^\varepsilon\|_{\text{TV}} &\leq \|\mu^\varepsilon - X_t^\varepsilon(x_0)\|_{\text{TV}} + \|X_t^\varepsilon(x_0) - Y^{\varepsilon,x_0}(t; 0, x_0)\|_{\text{TV}} + \|Y^{\varepsilon,x_0}(t; 0, x_0) - \mu_*^\varepsilon\|_{\text{TV}}. \end{aligned} \quad (2.79)$$

Here, we estimate the first-term of the right-hand side of inequality (2.79). By disintegration and the invariance property of  $\mu^\varepsilon$  it follows

$$\|\mu^\varepsilon - X_t^\varepsilon(x_0)\|_{\text{TV}} \leq \int_{\mathbb{R}^d} \|X_t^\varepsilon(u) - X_t^\varepsilon(x_0)\|_{\text{TV}} \mu^\varepsilon(du).$$

Let  $s_\varepsilon \gg t_\varepsilon^{x_0}$  for sufficiently small  $\varepsilon$ . The triangle inequality for the total variation distance implies

$$\begin{aligned} \int_{\mathbb{R}^d} \|X_{s_\varepsilon}^\varepsilon(u) - X_{s_\varepsilon}^\varepsilon(x_0)\|_{\text{TV}} \mu^\varepsilon(du) &\leq \int_{\mathbb{R}^d} \|X_{s_\varepsilon}^\varepsilon(u) - Y^{\varepsilon,u}(s_\varepsilon; 0, u)\|_{\text{TV}} \mu^\varepsilon(du) \\ &\quad + \int_{\mathbb{R}^d} \|Y^{\varepsilon,u}(s_\varepsilon; 0, u) - Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0)\|_{\text{TV}} \mu^\varepsilon(du) + \|Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0) - X_{s_\varepsilon}^\varepsilon(x_0)\|_{\text{TV}}. \end{aligned}$$

Since the total variation distance is bounded by one, we have for any  $K > 0$

$$\begin{aligned} & \int_{\mathbb{R}^d} \|X_{s_\varepsilon}^\varepsilon(u) - Y^{\varepsilon,u}(s_\varepsilon; 0, u)\|_{\text{TV}} \mu^\varepsilon(du) \\ & \leq \int_{|u| \leq K} \|X_{s_\varepsilon}^\varepsilon(u) - Y^{\varepsilon,u}(s_\varepsilon; 0, u)\|_{\text{TV}} \mu^\varepsilon(du) + \mu^\varepsilon(|u| \geq K) \end{aligned}$$

and

$$\begin{aligned} & \int_{\mathbb{R}^d} \|Y^{\varepsilon,u}(s_\varepsilon; 0, u) - Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0)\|_{\text{TV}} \mu^\varepsilon(du) \\ & \leq \int_{|u| \leq K} \|Y^{\varepsilon,u}(s_\varepsilon; 0, u) - Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0)\|_{\text{TV}} \mu^\varepsilon(du) + \mu^\varepsilon(|u| \geq K). \end{aligned}$$

Combining the preceding inequalities with inequality (2.79) we obtain

$$\|\mu^\varepsilon - \mu_*^\varepsilon\|_{\text{TV}} \leq I_1 + 2I_2 + 2I_3 + I_4 + I_5, \tag{2.80}$$

where

$$\begin{aligned} I_1 & := \|Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0) - \mu_*^\varepsilon\|_{\text{TV}}, & I_2 & := \|Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0) - X_{s_\varepsilon}^\varepsilon(x_0)\|_{\text{TV}}, \\ I_3 & := \mu^\varepsilon(|u| \geq K), & I_4 & := \int_{|u| \leq K} \|X_{s_\varepsilon}^\varepsilon(u) - Y^{\varepsilon,u}(s_\varepsilon; 0, u)\|_{\text{TV}} \mu^\varepsilon(du) \quad \text{and} \\ I_5 & := \int_{|u| \leq K} \|Y^{\varepsilon,u}(s_\varepsilon; 0, u) - Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0)\|_{\text{TV}} \mu^\varepsilon(du), \end{aligned}$$

for any  $x_0 \in \mathbb{R}^d$  and  $s_\varepsilon \gg t_\varepsilon^{x_0}$  for sufficiently small  $\varepsilon > 0$ . The remainder of the proof consists of showing that each of the terms  $I_i \rightarrow 0$  as  $\varepsilon \rightarrow 0$ ,  $i = 1, \dots, 5$ .

**Estimates for  $I_1$  in (2.80).** Let  $x_0 \neq 0$ . By Proposition 1 we have

$$\lim_{\varepsilon \rightarrow 0} \|Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0) - \mu_*^\varepsilon\|_{\text{TV}} = 0 \quad \text{for any } s_\varepsilon \gg t_\varepsilon^{x_0}, \text{ as } \varepsilon \rightarrow 0. \tag{2.81}$$

By (2.1) and Lemma C.2 in Appendix C we have that for any  $K > 0$

$$\lim_{\varepsilon \rightarrow 0} \sup_{|x_0| \leq K} \|Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0) - \mu_*^\varepsilon\|_{\text{TV}} = 0 \quad \text{for any } s_\varepsilon \gg \ln(1/\varepsilon) \text{ as } \varepsilon \rightarrow 0.$$

**Estimates for  $I_2$  in (2.80).** We estimate the second term as follows

$$\begin{aligned} & \|X_{s_\varepsilon}^\varepsilon(x_0) - Y^{\varepsilon,x_0}(s_\varepsilon; 0, x_0)\|_{\text{TV}} \\ & \leq \|X_{\Delta_\varepsilon}^\varepsilon(X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0)) - Y^{\varepsilon,x_0}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0))\|_{\text{TV}} \\ & \quad + \|Y^{\varepsilon,x_0}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0)) - Y^{\varepsilon,x_0}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, Y^{\varepsilon,x_0}(s_\varepsilon - \Delta_\varepsilon; 0, x_0))\|_{\text{TV}}. \end{aligned} \tag{2.82}$$

By Proposition B.1 in Appendix B our estimates in the previous sections remain valid up to times of order  $\varepsilon^{-\vartheta}$  for some  $\vartheta > 0$ . In the sequel, we set  $s_\varepsilon := \ln^2(\varepsilon)$ .

We start with the second term on the right-hand side of (2.82) and lighten the notation. By Proposition B.1 it is not hard to see that Lemma 2.4 remains valid for

$$z = X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0) \quad \text{and} \quad \tilde{z} = Y^{\varepsilon,x_0}(s_\varepsilon - \Delta_\varepsilon; 0, x_0).$$

For the convenience of the reader, we restate it here.

**Lemma 2.10.** *Let  $\gamma_\varepsilon = \Delta_\varepsilon^{-1/\alpha}$ , where  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$ . Then*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{P}_\varepsilon(\gamma_\varepsilon |z - \tilde{z}| > \eta \varepsilon) = 0 \quad \text{for any } \eta > 0, \tag{2.83}$$

for the joint distribution  $\mathbb{P}_\varepsilon(du, d\tilde{u}) = \mathbb{P}(X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0) \in du, Y^{\varepsilon,x_0}(s_\varepsilon - \Delta_\varepsilon; 0, x_0) \in d\tilde{u})$ .

The cutoff phenomenon for nonlinear Langevin systems with small noise

By (2.29), (2.30), Proposition 5 and the application the preceding statement to (2.36) we deduce

$$\lim_{\varepsilon \rightarrow 0} \|Y^{\varepsilon, x_0}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0)) - Y^{\varepsilon, x_0}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, Y^{\varepsilon, x_0}(s_\varepsilon - \Delta_\varepsilon, 0, x_0))\|_{\text{TV}} = 0.$$

In the sequel, we continue with the first term on the right-hand side of (2.82). By Corollary D.5 in Appendix D we have for any  $\eta > 0$ ,  $\vartheta \in (0, 1)$  and  $K > 0$

$$\lim_{\varepsilon \rightarrow 0} \sup_{|x_0| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0)| \geq \eta r_\varepsilon) = 0.$$

Since  $s_\varepsilon - \Delta_\varepsilon \gg t_\varepsilon^{x_0}$ , it is straightforward to see that the limit (2.63) remains valid for  $T_\varepsilon^x$  being replaced by  $s_\varepsilon - \Delta_\varepsilon$ . Consequently, we have

$$\lim_{\varepsilon \rightarrow 0} \|X_{\Delta_\varepsilon}^\varepsilon(X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0)) - Y^{\varepsilon, x_0}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x_0))\|_{\text{TV}} = 0.$$

**Estimates for  $I_3$  in (2.80).** By Corollary D.4 in Appendix D we have for all  $\beta' \leq \beta \wedge 1$  a positive constant  $C$  such that

$$\mathbb{E}[|X_t^\varepsilon(x)|^{\beta'}] \leq C\varepsilon^{\beta'} + |\varphi_t^x|^{\beta'} \leq C\varepsilon^{\beta'} + e^{-\delta\beta't}|x|^{\beta'} \quad \text{for } t \geq 0, x \in \mathbb{R}^d.$$

Let  $n \in \mathbb{N}$ . Then

$$\mathbb{E}[|X_t^\varepsilon(x)|^{\beta'} \wedge n] \leq e^{-\delta\beta't}|x|^{\beta'} \wedge n + C\varepsilon^{\beta'} \wedge n \quad \text{for } t \geq 0, x \in \mathbb{R}^d.$$

Since  $\mu^\varepsilon$  is stationary, we estimate for all  $t \geq 0$

$$\begin{aligned} \int_{\mathbb{R}^d} (|u|^{\beta'} \wedge n) \mu^\varepsilon(du) &= \int_{\mathbb{R}^d} \mathbb{E}[|X_t^\varepsilon(u)|^{\beta'} \wedge n] \mu^\varepsilon(du) \\ &\leq \int_{\mathbb{R}^d} (e^{-\delta\beta't}|u|^{\beta'} \wedge n) \mu^\varepsilon(du) + C\varepsilon^{\beta'} \wedge n. \end{aligned}$$

By the dominated convergence theorem we infer

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}^d} (e^{-\delta\beta't}|u|^{\beta'} \wedge n) \mu^\varepsilon(du) = 0 \quad \text{for all } n \in \mathbb{N}, \varepsilon \in (0, 1].$$

Therefore, we have for all  $n \in \mathbb{N}$  such  $n > C\varepsilon$

$$\int_{\mathbb{R}^d} (|u|^{\beta'} \wedge n) \mu^\varepsilon(du) \leq C\varepsilon^{\beta'} \wedge n \leq C\varepsilon^{\beta'}.$$

By the monotone convergence theorem we obtain

$$\int_{\mathbb{R}^d} |u|^{\beta'} \mu^\varepsilon(du) \leq C\varepsilon^{\beta'}. \tag{2.84}$$

The Markov inequality and (2.84) imply

$$\mu^\varepsilon(|u| \geq K) \leq \frac{\int_{\mathbb{R}^d} |u|^{\beta'} \mu^\varepsilon(du)}{K^{\beta'}} \leq \frac{C\varepsilon^{\beta'}}{K^{\beta'}}.$$

**Estimates for  $I_5$  in (2.80).** We start with the triangle inequality

$$\begin{aligned} \int_{|u| \leq K} \|Y^{\varepsilon, u}(s_\varepsilon; 0, u) - Y^{\varepsilon, x_0}(s_\varepsilon; 0, x_0)\|_{\text{TV}} \mu^\varepsilon(du) \\ \leq \sup_{|u| \leq K} \|Y^{\varepsilon, u}(s_\varepsilon; 0, u) - \mu_*^\varepsilon\|_{\text{TV}} + \|\mu_*^\varepsilon - Y^{\varepsilon, x_0}(s_\varepsilon; 0, x_0)\|_{\text{TV}}. \end{aligned}$$

The second term of the preceding inequality is equal to  $I_1$  and tends to 0 as  $\varepsilon \rightarrow 0$ . By (2.1) and since in Lemma C.3, Item (1), in Appendix C it is shown that  $\mu_\varepsilon^*$  is the law of  $\varepsilon Z_\infty$  we have

$$\|Y^{\varepsilon,u}(s_\varepsilon; 0, u) - \mu_\varepsilon^*\|_{\text{TV}} \leq \|Y^{\varepsilon,u}(s_\varepsilon; 0, u) - Z_\infty\|_{\text{TV}} + \|\varphi_{s_\varepsilon}^u/\varepsilon + Z_\infty - Z_\infty\|_{\text{TV}}$$

for any  $u \in \mathbb{R}^d$ . We start with the first term. By Lemma C.4 in Appendix C we have

$$\lim_{\varepsilon \rightarrow 0} \sup_{|u| \leq K} \|Y^{\varepsilon,u}(s_\varepsilon; 0, u) - Z_\infty\|_{\text{TV}} = 0.$$

We treat the second term. Let  $\eta > 0$ . By the shift-continuity of  $L^1$  distance yields that there exists  $\rho := \rho(\eta) > 0$  such that

$$\|u + Z_\infty - Z_\infty\|_{\text{TV}} \leq \eta \quad \text{whenever } |u| \leq \rho.$$

Note that for  $|u| \leq K$  we have

$$\left| \frac{\varphi_{s_\varepsilon}^u}{\varepsilon} \right| \leq \frac{e^{-\delta s_\varepsilon} |u|}{\varepsilon} \leq \frac{e^{-\delta s_\varepsilon} K}{\varepsilon} < \rho \quad \text{for sufficiently small } \varepsilon.$$

Therefore

$$\limsup_{\varepsilon \rightarrow 0} \int_{|u| \leq K} \left\| \frac{\varphi_{s_\varepsilon}^u}{\varepsilon} + Z_\infty - Z_\infty \right\|_{\text{TV}} \mu^\varepsilon(du) \leq \eta,$$

and consequently

$$\lim_{\varepsilon \rightarrow 0} \int_{|u| \leq K} \left\| \frac{\varphi_{s_\varepsilon}^u}{\varepsilon} + Z_\infty - Z_\infty \right\|_{\text{TV}} \mu^\varepsilon(du) = 0.$$

**Estimates for  $I_4$  in (2.80).** Note that

$$\begin{aligned} & \|X_{s_\varepsilon}^\varepsilon(x) - Y^{\varepsilon,x}(s_\varepsilon; 0, x)\|_{\text{TV}} \\ & \leq \|X_{\Delta_\varepsilon}^\varepsilon(X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x))\|_{\text{TV}} \\ & \quad + \|Y^{\varepsilon,x}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x)) - Y^{\varepsilon,x}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, Y^{\varepsilon,x}(s_\varepsilon - \Delta_\varepsilon; 0, x))\|_{\text{TV}}. \end{aligned} \tag{2.85}$$

We start with the first term. Recall that  $r_\varepsilon = \varepsilon^{1-\vartheta}$  for  $\vartheta \in (0, 1/4)$ . By disintegration we have  $\mathbb{P}_\varepsilon^x(dz) = \mathbb{P}(X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(x) \in dz)$

$$\begin{aligned} & \int_{|u| \leq K} \|X_{\Delta_\varepsilon}^\varepsilon(X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u)) - Y^{\varepsilon,u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u))\|_{\text{TV}} \mu^\varepsilon(du) \\ & \leq \int_{|u| \leq K} \int_{|z| \leq 2r_\varepsilon} \|X_{\Delta_\varepsilon}^\varepsilon(z) - Y^{\varepsilon,u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z)\|_{\text{TV}} \mathbb{P}_\varepsilon^u(dz) \mu^\varepsilon(du) \\ & \quad + \sup_{|u| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u)| > 2r_\varepsilon) \\ & \leq \sup_{|u| \leq K} \sup_{|z| \leq 2r_\varepsilon} \|X_{\Delta_\varepsilon}^\varepsilon(z) - Y^{\varepsilon,u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z)\|_{\text{TV}} + \sup_{|u| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u)| > 2r_\varepsilon) \\ & = \|X_{\Delta_\varepsilon}^\varepsilon(z_\varepsilon) - Y^{\varepsilon,u_\varepsilon}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z_\varepsilon)\|_{\text{TV}} + \sup_{|u| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u)| > 2r_\varepsilon) \end{aligned}$$

for some  $|u_\varepsilon| \leq K$  and  $|z_\varepsilon| \leq 2r_\varepsilon$ . The right-hand side of the preceding inequality tends to zero, as  $\varepsilon \rightarrow 0$ . This is due to Proposition 2 and Corollary D.5 in Appendix D.

We continue with the second term on the right-hand side of (2.85). Let

$$\mathbb{P}_\varepsilon^u(dz, d\tilde{z}) = \mathbb{P}(X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u) \in dz, Y^{\varepsilon,u}(s_\varepsilon - \Delta_\varepsilon; 0, u) \in d\tilde{z}).$$



Using the shift continuity (2.33) we fix  $\rho$  and choose  $\eta > 0$  accordingly. Again, by disintegration we have

$$\begin{aligned} & \int_{|u| \leq K} \|Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u)) - Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, Y^{\varepsilon, u}(s_\varepsilon - \Delta_\varepsilon; 0, u))\|_{\text{TV}} \mu^\varepsilon(du) \\ & \leq \int_{|u| \leq K} \int_{\substack{|z - \tilde{z}| \leq \eta \varepsilon \Delta_\varepsilon^{1/\alpha}, \\ |z| \leq 2r_\varepsilon, |\tilde{z}| \leq 2r_\varepsilon}} \|Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z) - Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, \tilde{z})\|_{\text{TV}} \mathbb{P}_\varepsilon^u(dz, d\tilde{z}) \mu^\varepsilon(du) \\ & \quad + \sup_{|u| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u) - Y^{\varepsilon, u}(s_\varepsilon - \Delta_\varepsilon, u, 0)| > \eta \varepsilon \Delta_\varepsilon^{1/\alpha}) \\ & \quad + \sup_{|u| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u)| > 2r_\varepsilon) + \sup_{|u| \leq K} \mathbb{P}(|Y^{\varepsilon, u}(s_\varepsilon - \Delta_\varepsilon; 0, u)| > 2r_\varepsilon). \end{aligned} \tag{2.86}$$

where the first term on the right-hand side is estimated by

$$\sup_{|u| \leq K} \sup_{\substack{|z - \tilde{z}| \leq \eta \varepsilon \Delta_\varepsilon^{1/\alpha}, \\ |z| \leq 2r_\varepsilon, |\tilde{z}| \leq 2r_\varepsilon}} \|Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z) - Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, \tilde{z})\|_{\text{TV}}. \tag{2.87}$$

We prove that the right-hand sides of (2.86) and (2.87) tend to zero, as  $\varepsilon \rightarrow 0$ . Due to limit (2.83) it follows

$$\lim_{\varepsilon \rightarrow 0} \sup_{|u| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u) - Y^{\varepsilon, u}(s_\varepsilon - \Delta_\varepsilon, u, 0)| > \eta \varepsilon \Delta_\varepsilon^{1/\alpha}) = 0.$$

By Corollary D.5 in Appendix D and a straightforward adaptation for the linearization  $Y^{\varepsilon, u}$ , we have

$$\lim_{\varepsilon \rightarrow 0} \sup_{|u| \leq K} \mathbb{P}(|X_{s_\varepsilon - \Delta_\varepsilon}^\varepsilon(u)| > 2r_\varepsilon) = \lim_{\varepsilon \rightarrow 0} \sup_{|u| \leq K} \mathbb{P}(|Y^{\varepsilon, u}(s_\varepsilon - \Delta_\varepsilon; 0, u)| > 2r_\varepsilon) = 0.$$

We continue with the term (2.87)

$$\sup_{|u| \leq K} \sup_{\substack{|z - \tilde{z}| \leq \eta \varepsilon \Delta_\varepsilon^{1/\alpha}, \\ |z| \leq 2r_\varepsilon, |\tilde{z}| \leq 2r_\varepsilon}} \|Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z) - Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, \tilde{z})\|_{\text{TV}}.$$

By (2.38) we have

$$\begin{aligned} Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z) &= (\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1} z \\ &+ (\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1} \int_0^{\Delta_\varepsilon} \Phi_s^\varepsilon(u) (Db(\varphi_{s_\varepsilon - \Delta_\varepsilon + s}^u) \varphi_{s_\varepsilon - \Delta_\varepsilon + s}^u - b(\varphi_{s_\varepsilon - \Delta_\varepsilon + s}^u)) ds + \varepsilon U_\varepsilon. \end{aligned}$$

By the shift and scale invariance of the total variation distance we obtain

$$\begin{aligned} & \|Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z) - Y^{\varepsilon, u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, \tilde{z})\|_{\text{TV}} \\ &= \|(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1} z + \varepsilon U_\varepsilon - (\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1} \tilde{z} + \varepsilon U_\varepsilon\|_{\text{TV}} \\ &= \left\| \left( (\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1} (z - \tilde{z}) + \varepsilon U_\varepsilon \right) - \varepsilon U_\varepsilon \right\|_{\text{TV}} \\ &= \left\| \left( \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1} (z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + \frac{U_\varepsilon}{\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^u \right) - \left( \frac{U_\varepsilon}{\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^u \right) \right\|_{\text{TV}}. \end{aligned}$$

Hence

$$\begin{aligned} & \|Y^{\varepsilon,u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, z) - Y^{\varepsilon,u}(\Delta_\varepsilon; s_\varepsilon - \Delta_\varepsilon, \tilde{z})\|_{\text{TV}} \\ & \leq \left\| \left( \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1}(z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + \frac{U_\varepsilon}{\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^u \right) - \left( \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1}(z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + U \right) \right\|_{\text{TV}} \\ & \quad + \left\| \left( \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1}(z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + U \right) - U \right\|_{\text{TV}} + \left\| U - \left( \frac{U_\varepsilon}{\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^u \right) \right\|_{\text{TV}} \\ & = 2 \left\| U - \left( \frac{U_\varepsilon}{\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^u \right) \right\|_{\text{TV}} + \left\| \left( \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1}(z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + U \right) - U \right\|_{\text{TV}}, \end{aligned}$$

where  $U \stackrel{d}{=} \mathcal{S}_\alpha(\Lambda_1)$ . Proposition 5 yields

$$\lim_{\varepsilon \rightarrow 0} \sup_{|u| \leq K} \left\| U - \left( \frac{U_\varepsilon}{\Delta_\varepsilon^{1/\alpha}} + a_\varepsilon^u \right) \right\|_{\text{TV}} = 0.$$

It remains to show for  $\eta > 0$

$$\sup_{|u| \leq K} \sup_{\substack{|z - \tilde{z}| \leq \eta \varepsilon \Delta_\varepsilon^{1/\alpha}, \\ |z| \leq 2r_\varepsilon, |\tilde{z}| \leq 2r_\varepsilon}} \left\| \left( \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1}(z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} + U \right) - U \right\|_{\text{TV}}. \tag{2.88}$$

Recall that (2.33) implies that

$$\left| \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1}(z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} \right| \leq \sqrt{d} \eta$$

yields that (2.88) is bounded from above by  $\rho$ . Sending first  $\varepsilon \rightarrow 0$  and then  $\rho \rightarrow 0$  yields the limit of (2.88) equals 0.

By (A.3) we have for  $|z - \tilde{z}| \leq \eta \varepsilon \Delta_\varepsilon^{1/\alpha}$  that

$$\left| \frac{(\Phi_{\Delta_\varepsilon}^\varepsilon(u))^{-1}(z - \tilde{z})}{\varepsilon \Delta_\varepsilon^{1/\alpha}} \right| \leq \sqrt{d} \eta.$$

Hence  $I_4 \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . This completes the proof of Proposition 4. □

## A The deterministic dynamics

This section gathers all results concerning the deterministic fine dynamics of the solution  $\varphi_t^x$  of (1.3) under Hypothesis 1. The following lemma is of interest since it shows that the time scale  $t_\varepsilon^x$  yields an estimate on the deterministic dynamics with of order exactly  $\varepsilon$ .

**Lemma A.1.** *Let  $\Delta_\varepsilon > 0$  such that  $\lim_{\varepsilon \rightarrow 0} \Delta_\varepsilon = 0$ . Let  $\rho \in \mathbb{R}$  we define  $T_\varepsilon^x = t_\varepsilon^x - \Delta_\varepsilon + \rho \cdot w_\varepsilon^x$ , where  $t_\varepsilon^x$  and  $w_\varepsilon^x$  are given in Theorem 2. Then there exists a positive constant  $C(|x|, \rho)$  that depends continuously on  $|x|$  such that  $|\varphi_{T_\varepsilon^x}^x| \leq C(|x|, \rho)\varepsilon$ .*

*Proof.* By Lemma 1.1 we have

$$\lim_{t \rightarrow \infty} \left| \frac{e^{\lambda_x t}}{t^{\ell_x - 1}} \varphi_{\tau_x + t}^x - v(t, x) \right| = 0, \tag{A.1}$$

where  $v(t, x) = \sum_{k=1}^m e^{i\theta_x^k t} v_x^k$ . A straightforward calculation shows that

$$\lim_{\varepsilon \rightarrow 0} \frac{(T_\varepsilon^x)^{\ell_x - 1} e^{-\lambda_x T_\varepsilon^x}}{\varepsilon} = \lambda_x^{1 - \ell_x} e^{-\rho}. \tag{A.2}$$

Then the triangle inequality yields

$$\begin{aligned} |\varphi_{T_\varepsilon^x}^x| &\leq |\varphi_{T_\varepsilon^x}^x - (T_\varepsilon^x)^{\ell_x-1} e^{-\lambda_x T_\varepsilon^x} v(T_\varepsilon^x, x)| + (T_\varepsilon^x)^{\ell_x-1} e^{-\lambda_x T_\varepsilon^x} |v(T_\varepsilon^x, x)| \\ &= (T_\varepsilon^x)^{\ell_x-1} e^{-\lambda_x T_\varepsilon^x} \left| \frac{e^{\lambda_x T_\varepsilon^x} \varphi_{T_\varepsilon^x}^x}{(T_\varepsilon^x)^{\ell_x-1}} - v(T_\varepsilon^x, x) \right| + (T_\varepsilon^x)^{\ell_x-1} e^{-\lambda_x T_\varepsilon^x} |v(T_\varepsilon^x, x)| \leq C(|x|, \rho)\varepsilon, \end{aligned}$$

where the last inequality follows from limit (A.1) and limit (A.2). □

The following strong version of the Grönwall-Bellman lemma frequently used and given for completeness.

**Lemma A.2** (Grönwall-Bellman inequality).

Let  $T > 0$  be fixed. Let  $g : [0, T] \rightarrow \mathbb{R}$  be a  $C^1$ -function and  $h : [0, T] \rightarrow \mathbb{R}$  be continuous. If

$$\frac{d}{dt}g(t) \leq -ag(t) + h(t) \quad \text{for any } t \in [0, T],$$

where  $a \in \mathbb{R}$ , and the derivative at 0 and  $T$  are understanding as the right and left derivatives, respectively. Then

$$g(t) \leq e^{-at}g(0) + e^{-at} \int_0^t e^{as}h(s)ds \quad \text{for any } t \in [0, T].$$

Moreover, if  $a \neq 0$  we have

$$|g(t)| \leq e^{-at}|g(0)| + \frac{(1 - e^{-at})}{a} \max_{s \in [0, t]} |h(s)| \quad \text{for any } t \in [0, T].$$

For the proof, see for instance Theorem 1.3.3 page 15 of [59]. Due to the variation of constants formula, the proof of linear cutoff relies essentially on precise norm estimates of the homogeneous and inhomogeneous linear solution flow, which are gathered in the following lemma.

**Lemma A.3.** Let  $(\varphi_t^x)_{t \geq 0}$  be the solution of (1.6). We consider for any fixed  $T \geq 0$  the solution  $\Phi = (\Phi_t(x))_{t \geq 0}$  of the matrix differential equation

$$\frac{d}{dt}\Phi_t = \Phi_t Db(\varphi_{T+t}^x) \quad \text{with } \Phi_0 = I_d,$$

the solution  $\Psi = (\Psi_t)_{t \geq 0}$  of the matrix differential equation

$$\frac{d}{dt}\Psi_t = \Psi_t Db(0) \quad \text{with } \Psi_0 = I_d$$

and the standard matrix 2-norm  $|\cdot|$ . Then the following statements are valid for any  $0 \leq s \leq t$ .

i) It follows

$$|\Phi_s^*(x)(\Phi_t^{-1}(x))^*| \leq \sqrt{d}e^{-\delta(t-s)} \quad \text{and} \quad |\Psi_s^*(\Psi_t^{-1})^*| \leq \sqrt{d}e^{-\delta(t-s)}. \quad (\text{A.3})$$

ii) For  $C(|x|) = \max_{|u| \leq |x|} |Db(u)|$  we have

$$|\Phi_t^*(x)(\Phi_s^{-1}(x))^*| \leq \sqrt{d}e^{-C(|x|)(t-s)} \quad \text{and} \quad |\Psi_t^*(\Psi_s^{-1})^*| \leq \sqrt{d}e^{-|Db(0)|(t-s)}.$$

iii) Let  $c_1 = 1/\sqrt{d}$ ,  $c_2 = C(|x|)$ ,  $c_3 = \sqrt{d}$  and  $c_4 = \delta$ , where  $C(|x|)$  is the constant obtained in item ii). Then for all  $z \in \mathbb{R}^d$

$$c_1 e^{-c_2(t-s)}|z| \leq |\Phi_s^*(x)(\Phi_t^{-1}(x))^*z| \leq c_3 e^{-c_4(t-s)}|z|.$$

The cutoff phenomenon for nonlinear Langevin systems with small noise

iv) There exist positive constant  $\tilde{c}_1, \tilde{c}_2, \tilde{c}_3$  and  $\tilde{c}_4$  such that for all  $z \in \mathbb{R}^d$

$$\tilde{c}_1 e^{-\tilde{c}_2(t-s)} |z| \leq |\Psi_s^*(\Psi_t^{-1})^* z| \leq \tilde{c}_3 e^{-\tilde{c}_4(t-s)} |z|.$$

v) For  $C(|x|)$  given in item iii) we have

$$|\Phi_t^{-1}(x)\Phi_s(x) - \Psi_t^{-1}\Psi_s|^2 \leq \frac{C^2(|x|)d^3}{4\delta^2} |\varphi_T^x|^2 e^{-\delta t} (1 - e^{-4\delta(t-s)}). \quad (\text{A.4})$$

In particular,

$$|\Phi_t^{-1}(x)\Phi_s(x) - \Psi_t^{-1}\Psi_s|^2 \leq \frac{C^2(|x|)d^3}{4\delta^2} |\varphi_T^x|^2 e^{-\delta t}. \quad (\text{A.5})$$

*Proof.* Let  $t \geq s \geq 0$  be fixed.

*Proof of item i).* Define  $\Pi_t^s(x) := \Phi_t^{-1}(x)\Phi_s(x)$ . Note that

$$\frac{d}{dt} \Pi_t^s(x) = \frac{d}{dt} \Phi_t^{-1}(x)\Phi_s(x) = -Db(\varphi_{T+t}^x)\Phi_t^{-1}(x)\Phi_s(x) = -Db(\varphi_{T+t}^x)\Pi_t^s(x).$$

We denote by  $\Pi_t^s(x) = ((\Pi_t^s(x))_{i,j})_{i,j \in \{1, \dots, d\}}$ . Observe that

$$\begin{aligned} \frac{d}{dt} |\Pi_t^s(x)|^2 &= -2 \sum_{i,j=1}^d (\Pi_t^s(x))_{i,j} \sum_{k=1}^d (Db(\varphi_{T+t}^x))_{i,k} (\Pi_t^s(x))_{k,j} \\ &= -2 \sum_{j=1}^d \sum_{i,k=1}^d (\Pi_t^s(x))_{i,j} (Db(\varphi_{T+t}^x))_{i,k} (\Pi_t^s(x))_{k,j} \leq -2\delta |\Pi_t^s(x)|^2, \end{aligned}$$

where the last inequality follows from Hypothesis 1. Since  $\Pi_s^s(x) = I_d$ , we have from Lemma A.2 that  $|\Phi_t^{-1}(x)\Phi_s(x)|^2 \leq d e^{-2\delta(t-s)}$ .

*Proof of item ii).* Let  $\tilde{\Pi}_t^s(x) := \Phi_s^{-1}(x)\Phi_t(x)$ . Note that

$$\frac{d}{dt} \tilde{\Pi}_t^s(x) = \Phi_s^{-1}(x) \frac{d}{dt} \Phi_t(x) = \Phi_s^{-1}(x)\Phi_t(x) Db(\varphi_{T+t}^x) = \tilde{\Pi}_t^s(x) Db(\varphi_{T+t}^x).$$

Observe that  $|\langle \tilde{z}, Db(\varphi_{T+t}^x)\tilde{z} \rangle| \leq |Db(\varphi_{T+t}^x)| |\tilde{z}|^2$  for  $\tilde{z} \in \mathbb{R}^d$ . By Hypothesis 1 we obtain  $|\varphi_t^x| \leq |x| e^{-\delta t}$  combined with  $b \in \mathcal{C}^2$  implies  $|Db(\varphi_t^x)| \leq \max_{|u| \leq |x|} |Db(u)|$ . Let  $C(|x|) :=$

$\max_{|u| \leq |x|} |Db(u)|$ . Here, we denote by  $\tilde{\Pi}_t^s(x) = ((\tilde{\Pi}_t^s(x))_{i,j})_{i,j \in \{1, \dots, d\}}$ . Then we have

$$\begin{aligned} \frac{d}{dt} |\tilde{\Pi}_t^s(x)|^2 &= 2 \sum_{i,j=1}^d (\tilde{\Pi}_t^s(x))_{i,j} \sum_{k=1}^d (\tilde{\Pi}_t^s(x))_{i,k} (Db(\varphi_{T+t}^x))_{k,j} \\ &= 2 \sum_{i=1}^d \sum_{k,j=1}^d (\tilde{\Pi}_t^s(x))_{i,j} (Db(\varphi_{T+t}^x))_{k,j} (\tilde{\Pi}_t^s(x))_{i,k} \leq 2C |\tilde{\Pi}_t^s(x)|^2. \end{aligned}$$

Since  $\tilde{\Pi}_s^s = I_d$ , Lemma A.2 yields  $|\Phi_s^{-1}(x)\Phi_t(x)|^2 \leq d e^{2C(|x|)(t-s)}$ .

*Proof of item iii).* Let  $z \in \mathbb{R}^d$  be fixed. On the one hand, item i) yields

$$|\Phi_s^*(x)(\Phi_t^{-1}(x))^* z| \leq |\Phi_s^*(x)(\Phi_t^{-1}(x))^*| |z| \leq \sqrt{d} e^{-\delta(t-s)} |z|.$$

On the other hand, we have

$$|z| = |(\Phi_t(x))^*(\Phi_s^{-1}(x))^*\Phi_s^*(x)(\Phi_t^{-1}(x))^*z| \leq |(\Phi_t(x))^*(\Phi_s^{-1}(x))^*\Phi_s^*(x)(\Phi_t^{-1}(x))^*z| \leq \sqrt{d}e^{C(|x|)(t-s)}|\Phi_s^*(x)(\Phi_t^{-1}(x))^*z|,$$

where the last inequality follows from item ii). Consequently,

$$|\Phi_s^*(x)(\Phi_t^{-1}(x))^*z| \geq \frac{1}{\sqrt{d}}e^{-C(|x|)(t-s)}|z|.$$

*Proof of item iv).* It follows analogously from item i) and ii). We omit the details.

*Proof of item v).* Let  $\Delta_t^s(x) := \Phi_t^{-1}(x)\Phi_s(x) - \Psi_t^{-1}\Psi_s$ . Then

$$\begin{aligned} \frac{d}{dt}\Delta_t^s(x) &= \frac{d}{dt}\Phi_t^{-1}(x)\Phi_s(x) - \frac{d}{dt}\Psi_t^{-1}\Psi_s \\ &= -Db(\varphi_{T+t}^x)\Phi_t^{-1}(x)\Phi_s(x) + Db(0)\Psi_t^{-1}\Psi_s \\ &= -Db(\varphi_{T+t}^x)\Delta_t^s(x) + (Db(0) - Db(\varphi_{T+t}^x))\Psi_t^{-1}\Psi_s. \end{aligned}$$

Here we denote by  $\Delta_t^s(x) = ((\Delta_t^s(x))_{i,j})_{i,j \in \{1, \dots, d\}}$ . Note that

$$\begin{aligned} \frac{d}{dt}|\Delta_t^s(x)|^2 &= 2 \sum_{i,j=1}^d (\Delta_t^s(x))_{i,j} \frac{d}{dt}(\Delta_t^s(x))_{i,j} \\ &= 2 \sum_{i,j=1}^d (\Delta_t^s(x))_{i,j} \left( \sum_{k=1}^d -(Db(\varphi_{T+t}^x))_{i,k}(\Delta_t^s(x))_{k,j} + (Db(0) - Db(\varphi_{T+t}^x))_{i,k}(\Psi_t^{-1}\Psi_s)_{k,j} \right) \\ &= -2 \sum_{j=1}^d \sum_{i,k=1}^d (\Delta_t^s(x))_{i,j} (Db(\varphi_{T+t}^x))_{i,k} (\Delta_t^s(x))_{k,j} \\ &\quad + 2 \sum_{i,j,k=1}^d (\Delta_t^s(x))_{i,j} (Db(0) - Db(\varphi_{T+t}^x))_{i,k} (\Psi_t^{-1}\Psi_s)_{k,j}. \end{aligned}$$

By Hypothesis 1 we obtain

$$\frac{d}{dt}|\Delta_t^s(x)|^2 \leq -2\delta|\Delta_t^s(x)|^2 + 2 \sum_{i,j,k=1}^d (\Delta_t^s(x))_{i,j} (Db(0) - Db(\varphi_{T+t}^x))_{i,k} (\Psi_t^{-1}\Psi_s)_{k,j}. \quad (\text{A.6})$$

The Young inequality yields

$$\begin{aligned} 2 \sum_{i,j,k=1}^d |(\Delta_t^s(x))_{i,j} (Db(0) - Db(\varphi_{T+t}^x))_{i,k} (\Psi_t^{-1}\Psi_s)_{k,j}| \\ \leq \sum_{i,j,k=1}^d \left( \frac{\delta}{d} |(\Delta_t^s(x))_{i,j}|^2 + \frac{d}{\delta} |(Db(0) - Db(\varphi_{T+t}^x))_{i,k} (\Psi_t^{-1}\Psi_s)_{k,j}|^2 \right) \\ = \delta|\Delta_t^s(x)|^2 + \frac{d}{\delta} \sum_{i,j,k=1}^d |(Db(0) - Db(\varphi_{T+t}^x))_{i,k} (\Psi_t^{-1}\Psi_s)_{k,j}|^2. \end{aligned} \quad (\text{A.7})$$

Since  $b \in C^2$ , there exists a positive constant  $C := C(|x|)$  such that

$$|Db(y) - Db(0)| \leq C|y| \quad \text{for any } y \text{ with } |y| \leq |x|.$$

By Hypothesis 1 we observe that  $|\varphi_t^x| \leq |x|e^{-\delta t} \leq |x|$ . Lemma A.3.ii) yields the following inequality  $|(\Psi_t^{-1}\Psi_s)|^2 \leq de^{-2\delta(t-s)}$ . Then

$$\begin{aligned} \sum_{i,j,k=1}^d |(Db(0) - Db(\varphi_{T+t}^x))_{i,k}(\Psi_t^{-1}\Psi_s)_{k,j}|^2 &\leq C^2 \sum_{i,j,k=1}^d |\varphi_{T+t}^x|^2 |(\Psi_t^{-1}\Psi_s)_{k,j}|^2 \\ &\leq C^2 |\varphi_{T+t}^x|^2 d^2 e^{-4\delta(t-s)}. \end{aligned} \tag{A.8}$$

Combining (A.6), (A.7) and (A.8) we infer

$$\frac{d}{dt} |\Delta_t^s(x)|^2 \leq -\delta |\Delta_t^s(x)|^2 + \frac{C^2 d^3}{\delta} |\varphi_{T+t}^x|^2 e^{-4\delta(t-s)}.$$

Since  $\Delta_s^s(x) = 0$ , the preceding differential inequality with the help of Lemma A.2 imply

$$|\Delta_t^s(x)|^2 \leq \frac{C^2 d^3}{\delta} e^{-\delta(t-s)} \int_s^t |\varphi_{T+u}^x|^2 e^{-3\delta(u-s)} du.$$

Observe that  $|\varphi_{T+t}^x| = |\varphi_t^{\varphi_T^x}| \leq e^{-\delta t} |\varphi_T^x|$ . Then

$$|\Delta_t^s(x)|^2 \leq \frac{C^2 d^3}{\delta} |\varphi_T^x|^2 e^{-\delta(t-s)} \int_s^t e^{-\delta u} e^{-3\delta(u-s)} du.$$

The integral version of the Grönwall-Bellman lemma given in [57], Lemma 1, yields

$$|\Delta_t^s(x)|^2 \leq \frac{C^2 d^3}{4\delta^2} |\varphi_T^x|^2 e^{-\delta t} (1 - e^{-4\delta(t-s)}). \quad \square$$

## B Freidlin-Wentzell first order approximation

The result of this section yields a precise quantification of the inhomogeneous linearization error of  $X^{\varepsilon,x}$  by  $Y^\varepsilon(x)$  given in (1.26) under the Hypothesis 1 and 2 for any moment  $\beta > 0$ .

**Lemma B.1** (Quantitative first order expansion).

Assume Hypothesis 1 and 2 for some  $\beta > 0$ . For  $(t_\varepsilon)_{\varepsilon>0}$  with  $t_\varepsilon \rightarrow \infty$  as  $\varepsilon \rightarrow 0$  let the following limit hold true

$$\lim_{\varepsilon \rightarrow 0} t_\varepsilon \varepsilon^{\frac{1}{1+2(\beta \wedge 1)}} = 0.$$

Then for any  $\alpha \in (0, 2)$ ,  $K > 0$  and  $\Delta_\varepsilon = \varepsilon^{\frac{\alpha}{2}}$  there exist positive constants  $\varepsilon_0 = \varepsilon_0(K, \alpha, \beta, \delta)$  and  $C = C(K, \alpha, \beta, \delta)$  such that for all  $\varepsilon \in (0, \varepsilon_0]$

$$\sup_{|x| \leq K} \mathbb{P}(|X_{t_\varepsilon}^{\varepsilon,x} - Y_{t_\varepsilon}^\varepsilon(x)| \geq \Delta_\varepsilon^{1/\alpha} \varepsilon) \leq C(K) \varepsilon^{\frac{\beta \wedge 1}{1+2(\beta \wedge 1)}}.$$

*Proof.* Let  $t \geq 0$ . Recall that  $Y_t^\varepsilon(x) = \varphi_t^x + \varepsilon Y_t^x$ , by (1.26), where

$$\varepsilon Y_t^x = - \int_0^t Db(\varphi_s^x) \varepsilon Y_s^x ds + \varepsilon dL_t.$$

That is  $W_t^\varepsilon = \varepsilon Y_t^x$  satisfies

$$W_t^\varepsilon = - \int_0^t Db(\varphi_s^x) W_s^\varepsilon ds + \varepsilon dL_t.$$

Hence

$$\begin{aligned} X_t^{\varepsilon,x} - Y_t^\varepsilon(x) &= X_t^{\varepsilon,x} - \varphi_t^x - W_t^\varepsilon \\ &= \int_0^t \left( -b(X_s^{\varepsilon,x}) - (-b(Y_s^\varepsilon(x))) \right) ds \\ &\quad + \int_0^t \left( -b(Y_s^\varepsilon(x)) - (-b(\varphi_s^x) - Db(\varphi_s^x)W_s^\varepsilon) \right) ds. \end{aligned}$$

The chain rule yields

$$\begin{aligned} |X_t^{\varepsilon,x} - Y_t^\varepsilon(x)|^2 &= -2 \int_0^t \langle b(X_s^{\varepsilon,x}) - b(Y_s^\varepsilon(x)), X_s^{\varepsilon,x} - Y_s^\varepsilon(x) \rangle ds \\ &\quad + 2 \int_0^t \langle -b(Y_s^\varepsilon(x)) - (-b(\varphi_s^x) - Db(\varphi_s^x)\varepsilon Y_s^x), X_s^{\varepsilon,x} - Y_s^\varepsilon(x) \rangle ds. \end{aligned}$$

By the mean value theorem, the Cauchy-Schwarz and the Young inequality we have

$$\begin{aligned} | -b(Y_s^\varepsilon(x)) - (-b(\varphi_s^x) - Db(\varphi_s^x)\varepsilon Y_s^x) | &= | -b(\varphi_s^x + \varepsilon Y_s^x) - (-b(\varphi_s^x) - Db(\varphi_s^x)\varepsilon Y_s^x) | \\ &\leq \int_0^1 \int_0^1 \|D^2b(\varphi_s^x + \theta_1\theta_2\varepsilon Y_s^x)\| d\theta_1 d\theta_2 |W_s^\varepsilon|^2. \end{aligned}$$

Together with Hypothesis 1 we obtain

$$\begin{aligned} |X_t^{\varepsilon,x} - Y_t^\varepsilon(x)|^2 &\leq -\delta \int_0^t |X_s^{\varepsilon,x} - Y_s^\varepsilon(x)|^2 ds \\ &\quad + \frac{1}{\delta} \int_0^t \left( \int_0^1 \int_0^1 \|D^2b(\varphi_s^x + \theta_1\theta_2\varepsilon Y_s^x)\| d\theta_1 d\theta_2 |W_s^\varepsilon|^2 \right)^2 ds. \end{aligned}$$

Then the integral version of the Grönwall-Bellman lemma given in [57], Lemma 1, yields

$$|X_t^{\varepsilon,x} - Y_t^\varepsilon(x)|^2 \leq \frac{\varepsilon^{2-\theta}}{\delta} \int_0^t \left( \int_0^1 \int_0^1 \|D^2b(\varphi_s^x + \theta_1\theta_2\varepsilon Y_s^x)\| d\theta_1 d\theta_2 \right)^2 \varepsilon^\theta |Y_s^x|^2 ds.$$

Let  $M > 0$  and  $\theta \in (0, 1)$  and introduce

$$A_M^\varepsilon := \left\{ \varepsilon^\theta \sup_{0 \leq s \leq t} |Y_s^1(x)|^2 \leq M \right\}.$$

For  $\varepsilon \in (0, 1]$  we have

$$A_M^\varepsilon \subset \left\{ \sup_{0 \leq s \leq t} |Y_s^1(x)|^2 \leq M \right\}.$$

Then on the event  $A_M^\varepsilon$  it follows

$$|X_t^{\varepsilon,x} - Y_t^\varepsilon(x)|^2 \leq \varepsilon^{2-\theta} C_M t.$$

Observe that

$$\mathbb{P}(|X_{t_\varepsilon}^{\varepsilon,x} - Y_{t_\varepsilon}^\varepsilon(x)| \geq \Delta_\varepsilon^{1/\alpha} \varepsilon) \leq \mathbb{P}(|X_{t_\varepsilon}^{\varepsilon,x} - Y_{t_\varepsilon}^\varepsilon(x)| \geq \Delta_\varepsilon^{1/\alpha} \varepsilon, A_M^\varepsilon) + \mathbb{P}((A_M^\varepsilon)^c).$$

Then

$$\begin{aligned} \mathbb{P}\left(|X_{t_\varepsilon}^{\varepsilon,x} - Y_{t_\varepsilon}^\varepsilon(x)| \geq \Delta_\varepsilon^{1/\alpha} \varepsilon, A_M^\varepsilon\right) &\leq \mathbb{P}\left(C_M \varepsilon^{2-\theta} t_\varepsilon \geq \Delta_\varepsilon^{1/\alpha} \varepsilon, A_M^\varepsilon\right) \\ &= \mathbb{P}\left(C_M t_\varepsilon \geq \frac{\Delta_\varepsilon^{1/\alpha}}{\varepsilon^{1-\theta}}, A_M^\varepsilon\right). \end{aligned}$$

Choosing  $\Delta_\varepsilon = \varepsilon^{\alpha/2}$  we obtain

$$\mathbb{P}\left(|X_{t_\varepsilon}^{\varepsilon,x} - Y_{t_\varepsilon}^\varepsilon(x)| \geq \Delta_\varepsilon^{1/\alpha} \varepsilon, A_M^\varepsilon\right) \leq \mathbb{P}\left(C_M t_\varepsilon \geq \frac{1}{\sqrt{\varepsilon^{1-\theta}}}, A_M^\varepsilon\right) = 0 \quad \text{for } \varepsilon \ll 1.$$

In the sequel, we estimate the term

$$\mathbb{P}\left(\sup_{0 \leq s \leq t_\varepsilon} |Y_s^1(x)|^2 > \frac{M}{\varepsilon^\theta}\right).$$

By Theorem 1 in [72] we have

$$\sup_{0 \leq s \leq t_\varepsilon} |Y_s^1(x)| \leq 6\sqrt{[Y^1(x)]_{t_\varepsilon}} + 2 \int_0^{t_\varepsilon} H_{s-} dY_s^1(x) \quad \text{a.s.,}$$

where

$$H_{s-} = \frac{Y_s^1(x)}{\sqrt{\sup_{0 \leq s \leq t_\varepsilon} |Y_s^1(x)|^2 + [Y^1(x)]_{s-}}}$$

In particular, we have

$$\begin{aligned} [Y^1(x)]_t &= [L]_t = \int_0^t \int_{|z| \leq 1} |z|^2 N(dsdz) \quad \text{such that} \\ \int_0^t H_{s-} dY_s^1(x) &= \int_0^t \langle H_{s-}, -Db(\varphi_s^x) Y_s^1(x) \rangle ds + \int_0^t \int_{|z| \leq 1} \langle H_{s-}, z \rangle \tilde{N}(dsdz) \\ &\quad + \int_0^t \int_{|z| > 1} \langle H_{s-}, z \rangle N(dsdz). \end{aligned}$$

We apply Hypothesis 1 and obtain a.s.

$$\int_0^t \langle H_{s-}, -Db(\varphi_s^x) Y_s^1(x) \rangle ds \leq 0.$$

Hence

$$\begin{aligned} &\mathbb{P}\left(\sup_{0 \leq s \leq t_\varepsilon} |Y_s^1(x)| > \frac{M}{\varepsilon^\theta}\right) \\ &\leq \mathbb{P}\left(6\left(\int_0^{t_\varepsilon} \int_{|z| \leq 1} |z|^2 N(dsdz)\right)^{1/2} + \int_0^{t_\varepsilon} \int_{|z| \leq 1} \langle H_{s-}, z \rangle \tilde{N}(dsdz) \right. \\ &\quad \left. + \int_0^{t_\varepsilon} \int_{|z| > 1} \langle H_{s-}, z \rangle N(dsdz) > \frac{M}{\varepsilon^\theta}\right) \\ &\leq \mathbb{P}\left(\int_0^{t_\varepsilon} \int_{|z| \leq 1} |z|^2 N(dsdz) > \frac{1}{18^2} \frac{M^2}{\varepsilon^{2\theta}}\right) + \mathbb{P}\left(\int_0^{t_\varepsilon} \int_{|z| \leq 1} \langle H_{s-}, z \rangle \tilde{N}(dsdz) > \frac{1}{3} \frac{M}{\varepsilon^\theta}\right) \\ &\quad + \mathbb{P}\left(\int_0^{t_\varepsilon} \int_{|z| > 1} \langle H_{s-}, z \rangle N(dsdz) > \frac{1}{3} \frac{M}{\varepsilon^\theta}\right). \end{aligned}$$

We continue term by term. First we obtain

$$\mathbb{P}\left(\int_0^{t_\varepsilon} \int_{|z| \leq 1} |z|^2 N(dsdz) > \left(\frac{1}{18}\right)^2 \frac{M^2}{\varepsilon^{2\theta}}\right) \leq t_\varepsilon \varepsilon^{2\theta} \frac{(18)^2}{M^2} \int_{|z| \leq 1} |z|^2 \nu(dz) = C t_\varepsilon \varepsilon^{2\theta}. \quad (\text{B.1})$$



By the Markov inequality we bound the second term and obtain

$$\begin{aligned} \mathbb{P}\left(\int_0^{t_\varepsilon} \int_{|z| \leq 1} \langle H_{s-}, z \rangle \tilde{N}(dsdz) > \frac{1}{3} \frac{M}{\varepsilon^\theta}\right) &\leq \varepsilon^{2\theta} \left(\frac{3}{M}\right)^2 \mathbb{E}\left[\left(\int_0^{t_\varepsilon} \int_{|z| \leq 1} \langle H_{s-}, z \rangle \tilde{N}(dsdz)\right)^2\right] \\ &= \varepsilon^{2\theta} \left(\frac{3}{M}\right)^2 \mathbb{E}\left[\int_0^{t_\varepsilon} \int_{|z| \leq 1} \langle H_{s-}, z \rangle^2 \nu(dz) ds\right] \\ &= \varepsilon^{2\theta} t_\varepsilon \left(\frac{3}{M}\right)^2 \int_{|z| \leq 1} |z|^2 \nu(dz). \end{aligned} \tag{B.2}$$

Finally,

$$\begin{aligned} \mathbb{P}\left(\int_0^{t_\varepsilon} \int_{|z| > 1} \langle H_{s-}, z \rangle N(ds dz) > \frac{1}{3} \frac{M}{\varepsilon^\theta}\right) &\leq \mathbb{P}\left(\int_0^{t_\varepsilon} \int_{|z| > 1} |z| N(ds dz) > \frac{1}{3} \frac{M}{\varepsilon^\theta}\right) \\ &\leq \varepsilon^{(\beta \wedge 1)\theta} \left(\frac{3}{M}\right)^{(\beta \wedge 1)} \mathbb{E}\left[\left(\int_0^{t_\varepsilon} \int_{|z| > 1} |z| N(ds dz)\right)^{(\beta \wedge 1)}\right] \\ &\leq \varepsilon^{(\beta \wedge 1)\theta} \left(\frac{3}{M}\right)^{(\beta \wedge 1)} \mathbb{E}\left[\int_0^{t_\varepsilon} \int_{|z| > 1} |z|^{(\beta \wedge 1)} N(ds dz)\right] \end{aligned} \tag{B.3}$$

$$= \varepsilon^{(\beta \wedge 1)\theta} t_\varepsilon \left(\frac{3}{M}\right)^{(\beta \wedge 1)} \int_{|z| > 1} |z|^{(\beta \wedge 1)} \nu(dz), \tag{B.4}$$

where we have used the subadditivity of the power  $\beta \wedge 1$  in the sense of Subsection 1.1.2, see formula (1.6) in [63]. Optimizing over  $\theta$  we obtain  $\theta = \frac{1}{1+2(\beta \wedge 1)}$ .  $\square$

### C The linear inhomogeneous dynamics $Y^\varepsilon(x)$

This section gathers properties of the inhomogeneous first order expansion  $Y^\varepsilon(x)$  of  $X^{\varepsilon,x}$  mainly with the help of Fourier techniques.

#### C.1 $\beta$ -Hölder continuity of the characteristic exponent of a Lévy process

It is classical that  $\beta \geq 1$  in Hypothesis 2 implies that the characteristic function is continuously differentiable, and hence locally Lipschitz continuous. This remains valid for the characteristic exponent  $\psi$ . In the sequel, we provide an elementary proof for the respective fractional case  $\beta \in (0, 1)$ .

**Proposition 7** (Local Hölder continuity of the characteristic exponent).

Let  $L = (L_t)_{t \geq 0}$  be a Lévy process on  $\mathbb{R}^d$ . Denote by  $\psi$  its characteristic exponent and by  $\nu$  its Lévy measure. Assume that

$$\int_{|z| \geq 1} |z|^\beta \nu(dz) < \infty \quad \text{for some } \beta > 0.$$

Then we have the following.

1. If  $\beta \geq 1$ ,  $\psi$  is  $\mathcal{C}^1$ . In particular, it is Lipschitz continuous.
2. If  $\beta \in (0, 1)$ ,  $\psi$  is locally Hölder continuous with Hölder index  $\beta$ .

*Proof.* The proof of item (1) is given in Theorem 15.32 of [45]. We continue with the proof of (2). Assume that  $\beta \in (0, 1)$ . We prove that  $\psi$  is locally Hölder continuous. Recall that

$$\psi(z) = \int_{\mathbb{R}^d} (e^{i\langle u, z \rangle} - 1 - i\langle u, z \rangle \mathbf{1}_{\{|u| \leq 1\}}(u)) \nu(du), \quad z \in \mathbb{R}^d.$$

For any  $z \in \mathbb{R}^d$ , let

$$f_1(z) = \int_{|u| \leq 1} (e^{i\langle u, z \rangle} - 1 - i\langle u, z \rangle) \nu(du) \quad \text{and} \quad f_2(z) = \int_{|u| > 1} (e^{i\langle u, z \rangle} - 1) \nu(du).$$

First, we analyze  $f_2$ . Let  $z_1, z_2 \in \mathbb{R}^d$ . Notice that

$$\begin{aligned} |f_2(z_1) - f_2(z_2)| &\leq \int_{|u| > 1} |e^{i\langle u, z_1 \rangle} - e^{i\langle u, z_2 \rangle}| \nu(du) \\ &= \int_{|u| > 1} |e^{i\langle u, z_1 - z_2 \rangle} - 1| \nu(du) = \sqrt{2} \int_{|u| > 1} \sqrt{1 - \cos(\langle u, z_1 - z_2 \rangle)} \nu(du) \\ &= 2 \int_{|u| > 1} \left| \sin \left( \frac{\langle u, z_1 - z_2 \rangle}{2} \right) \right| \nu(du) \\ &= 2 \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| > 1}} \left| \sin \left( \frac{\langle u, z_1 - z_2 \rangle}{2} \right) \right| \nu(du) + 2 \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| \leq 1}} \left| \sin \left( \frac{\langle u, z_1 - z_2 \rangle}{2} \right) \right| \nu(du) \\ &\leq 2 \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| > 1}} |\langle u, z_1 - z_2 \rangle|^\beta \nu(du) + 2 \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| \leq 1}} \left| \sin \left( \frac{\langle u, z_1 - z_2 \rangle}{2} \right) \right| \nu(du) \\ &\leq 2|z_1 - z_2|^\beta \int_{|u| > 1} |u|^\beta \nu(du) + 2 \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| \leq 1}} \left| \sin \left( \frac{\langle u, z_1 - z_2 \rangle}{2} \right) \right| \nu(du) \\ &= C_\beta |z_1 - z_2|^\beta + 2 \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| \leq 1}} \left| \sin \left( \frac{\langle u, z_1 - z_2 \rangle}{2} \right) \right| \nu(du), \end{aligned}$$

where  $C_\beta = 2 \int_{|u| > 1} |u|^\beta \nu(du) < \infty$ . Let  $\tilde{C}_\beta = \sup \left\{ \frac{|\sin(\theta/2)|}{|\theta|^\beta} : |\theta| \leq 1 \right\}$ . Since  $\beta \in (0, 1)$  we have  $\tilde{C}_\beta < \infty$ . Indeed, notice that  $\lim_{\theta \rightarrow 0} \frac{|\sin(\theta/2)|}{|\theta|^\beta} = 0$  then  $C < \infty$ . Furthermore,

$$\begin{aligned} \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| \leq 1}} \left| \sin \left( \frac{\langle u, z_1 - z_2 \rangle}{2} \right) \right| \nu(du) &\leq \int_{\substack{|u| > 1, \\ |\langle u, z_1 - z_2 \rangle| \leq 1}} \tilde{C}_\beta |\langle u, z_1 - z_2 \rangle|^\beta \nu(du) \\ &\leq \tilde{C}_\beta |z_1 - z_2|^\beta \int_{|u| > 1} |u|^\beta \nu(du). \end{aligned}$$

Hence,  $|f_2(z_2) - f_2(z_1)| \leq C(\beta) |z_2 - z_1|^\beta$  for any  $z_1, z_2 \in \mathbb{R}^d$ . In the sequel, we analyze  $f_1$ . We calculate for  $z_1, z_2 \in \mathbb{R}^d$

$$\begin{aligned} |f_1(z_1) - f_1(z_2)| &\leq \int_{|u| \leq 1} |e^{i\langle u, z_1 \rangle} - e^{i\langle u, z_2 \rangle} - i\langle u, z_1 - z_2 \rangle| \nu(du) \\ &\leq C |z_1 - z_2| \int_{|u| \leq 1} |u|^2 \nu(du) = C_1 |z_1 - z_2|, \end{aligned}$$

where we have used the mean value theorem for the integrand

$$\begin{aligned} |e^{i\langle u, z_1 \rangle} - e^{i\langle u, z_2 \rangle} - i\langle u, z_1 - z_2 \rangle| &= \left| \int_0^1 e^{i\langle u, z_1 + \theta(z_2 - z_1) \rangle} \langle u, z_1 - z_2 \rangle d\theta - \langle u, z_1 - z_2 \rangle \right| \\ &= \int_0^1 |e^{i\langle u, z_1 + \theta(z_2 - z_1) \rangle} - 1| d\theta |\langle u, z_1 - z_2 \rangle| \end{aligned}$$

with  $C_1 = C \int_{|u| \leq 1} |u|^2 \nu(du) < \infty$ . If  $|z_1| \leq \frac{1}{2}$  and  $|z_2| \leq \frac{1}{2}$ , then  $|z_1 - z_2| \leq |z_1 - z_2|^\beta$ . This concludes the proof of (2).  $\square$

**Remark C.1.**

1. Note that the above calculations for  $f_1$  give an elementary proof of the fact that any pure jump Lévy process with uniformly bounded jumps-sizes has a globally Lipschitz continuous characteristic exponent  $\psi$ .
2. The calculations for  $f_2$  yield that any compound Poisson process with  $\beta$ -integrability

$$\int_{|z|>1} |z|^\beta \nu(dz) < \infty \quad \text{for some } \beta > 0$$

has a locally Hölder continuous characteristic exponent  $\psi$  with Hölder index  $\beta$ . This extends the well-known result that the existence of integer moments translates to the respective order of differentiability of the characteristic function to the case of fractional moments.

**C.2 Continuous dependence of the total variation in the nonlinearity**

**Lemma C.2** (Continuous dependence on the initial value).

Let  $t > 0$ ,  $x \in \mathbb{R}^d$  and denote by  $(Y_t^x)_{t \geq 0}$  the unique strong solution of (1.25) as well as by  $g_t^x$  the respective density of  $Y_t^x$ .

Then  $x \mapsto g_t^x(u)$  is continuous for any fixed  $t > 0$  and  $u \in \mathbb{R}^d$ . In addition, the map  $x \mapsto \|Y_t^x - U\|_{TV}$  is continuous for any fixed  $t > 0$  and  $U$  any random vector on  $\mathbb{R}^d$ .

*Proof.* Let  $x, x' \in \mathbb{R}^d$ . The Fourier inversion formula yields

$$\begin{aligned} g_t^x(u) - g_t^{x'}(u) &= C_\pi \int_{\mathbb{R}^d} e^{i\langle u, \theta \rangle} \left( \hat{f}_t^x(\theta) - \hat{f}_t^{x'}(\theta) \right) d\theta \\ &= C_\pi \int_{|\theta| \leq K} e^{i\langle u, \theta \rangle} \left( \hat{f}_t^x(\theta) - \hat{f}_t^{x'}(\theta) \right) d\theta + C_\pi \int_{|\theta| > K} e^{i\langle u, \theta \rangle} \left( \hat{f}_t^x(\theta) - \hat{f}_t^{x'}(\theta) \right) d\theta \end{aligned} \quad (C.1)$$

for any  $u \in \mathbb{R}^d$ . We start with the first term of the right-hand side of the preceding inequality. Recall

$$\hat{f}_t^x(\theta) = \exp \left( \int_0^t \psi(\Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta) ds \right), \quad \hat{f}_t^{x'}(\theta) = \exp \left( \int_0^t \psi(\Phi_s^*(x')(\Phi_t^*)^{-1}(x')\theta) ds \right).$$

For any  $|\theta| \leq K$  we have

$$\sup_{0 \leq s \leq t} |\psi(\Phi_s^*(x')(\Phi_t^*)^{-1}(x')\theta) - \psi(\Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta)| \rightarrow 0, \quad \text{as } x' \rightarrow x.$$

Indeed, by Proposition 7, Item (2), in Appendix C there exists a positive constant  $C_K$  such that for all  $|\theta| \leq K$ ,  $x, x' \in \mathbb{R}^d$  we have

$$\begin{aligned} &|\psi(\Phi_s^*(x')(\Phi_t^*)^{-1}(x')\theta) - \psi(\Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta)| \\ &\leq C_K |\theta|^{1 \wedge \beta} |\Phi_s^*(x')(\Phi_t^*)^{-1}(x') - \Phi_s^*(x)(\Phi_t^*)^{-1}(x)|^{1 \wedge \beta} \\ &\leq C_K |\theta|^{1 \wedge \beta} (|\Phi_s^*(x')(\Phi_t^*)^{-1}(x')|^{1 \wedge \beta} + |\Phi_s^*(x)(\Phi_t^*)^{-1}(x)|^{1 \wedge \beta}) \\ &\leq 2C_K |\theta|^{1 \wedge \beta} (\sqrt{d} e^{-\delta(t-s)})^{1 \wedge \beta}. \end{aligned}$$

Then the dominated convergence theorem in the exponent yields for any  $|\theta| \leq K$

$$\hat{f}_t^{x'}(\theta) \rightarrow \hat{f}_t^x(\theta), \quad \text{as } x' \rightarrow x.$$

Again, by dominated convergence we have

$$\int_{|\theta| \leq K} e^{i\langle u, \theta \rangle} \left( \hat{f}_t^x(\theta) - \hat{f}_t^{x'}(\theta) \right) d\theta \rightarrow 0, \quad \text{as } x' \rightarrow x.$$

We continue with the second term of the right-hand side. Let  $|\theta| > K$  and we assume that  $|x'| \leq r$  and  $|x| \leq r$  where  $r = 2|x|$ . We analyze

$$\begin{aligned} |\hat{f}_t^x(\theta)| &= \exp\left(\int_0^t \int_{\mathbb{R}^d} (\cos(\langle \Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta, w \rangle) - 1) \nu(dw) ds\right) \\ &\leq \exp\left(\int_0^t \int_{|\langle \Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta, w \rangle| \leq \pi} (\cos(\langle \Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta, w \rangle) - 1) \nu(dw) ds\right) \\ &\leq \exp\left(-\frac{2}{\pi^2} \int_0^t \int_{|\langle \Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta, w \rangle| \leq \pi} |\langle \Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta, w \rangle|^2 \nu(dw) ds\right). \end{aligned}$$

By Lemma A.3 we have

$$|\Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta| \geq c_1 e^{-c_2(r)(t-s)} |\theta| \quad \text{for any } s \in [0, t], |x| \leq r, \theta \in \mathbb{R}^d.$$

Note that

$$c_1 e^{-c_2(r)(t-s)} |\theta| \geq c_1 K e^{-c_2(r)t}.$$

Since  $t > 0$  is fixed. The choice  $K > \frac{e^{c_2(r)t} C_{\triangleleft}}{c_1}$  yields  $c_1 e^{-c_2(r)(t-s)} |\theta| > C_{\triangleleft}$  where  $C_{\triangleleft}$  is the constant that appears in Lemma 1.7. Then we have for  $|\theta| > K$

$$|\hat{f}_t^x(\theta)| \leq \exp\left(-\frac{2}{\pi^2} c_{\triangleleft} \int_0^t |\Phi_s^*(x)(\Phi_t^*)^{-1}(x)\theta|^\alpha ds\right) \leq \exp\left(-\frac{2}{\pi^2} c_{\triangleleft} |\theta|^\alpha \int_0^t c_1^\alpha e^{-c_2(r)s\alpha} ds\right).$$

Then

$$\begin{aligned} \int_{|\theta| > K} |\hat{f}_t^{x'}(\theta) - \hat{f}_t^x(\theta)| d\theta &\leq \int_{|\theta| > K} |\hat{f}_t^{x'}(\theta)| d\theta + \int_{|\theta| > K} |\hat{f}_t^x(\theta)| d\theta \\ &\leq 2 \int_{|\theta| > K} \exp\left(-\frac{2}{\pi^2} c_{\triangleleft} |\theta|^\alpha \int_0^t c_1^\alpha e^{-c_2(r)s\alpha} ds\right) d\theta < \infty. \end{aligned}$$

Sending  $x' \rightarrow x$  and subsequently  $K \rightarrow \infty$  we obtain

$$\lim_{K \rightarrow \infty} \limsup_{x' \rightarrow x} \int_{|\theta| > K} |\hat{f}_t^{x'}(\theta) - \hat{f}_t^x(\theta)| d\theta = 0.$$

By (C.1) we obtain

$$\lim_{x' \rightarrow x} \int_{\mathbb{R}^d} |\hat{f}_t^{x'}(\theta) - \hat{f}_t^x(\theta)| d\theta = 0.$$

The preceding limit yields that  $x \in \mathbb{R}^d \mapsto g_t^x(u) \in [0, \infty)$  is continuous for any  $t > 0$  and  $u \in \mathbb{R}^d$  fixed. This proves the first part of the statement.

We show the second part of the statement. The Scheffé lemma applied to the densities  $g_t^x, g_t^{x'}$  implies for any  $t > 0$

$$\|Y_t^{x'} - Y_t^x\|_{\text{TV}} \rightarrow 0, \quad \text{as } x' \rightarrow x.$$

The triangle inequality yields for any random vector  $U$  on  $\mathbb{R}^d$

$$\left| \|Y_t^{x'} - U\|_{\text{TV}} - \|Y_t^x - U\|_{\text{TV}} \right| \leq \|Y_t^{x'} - Y_t^x\|_{\text{TV}}$$

for any  $t > 0$  and  $x, x' \in \mathbb{R}^d$ . Combining both preceding expressions finishes the proof.  $\square$

**C.3 Ergodicity of the inhomogeneous O-U process  $Y_t^\varepsilon(x)$**

**C.3.1 Existence of the limiting distribution  $\mu_*^\varepsilon$  and its convergence in law**

**Lemma C.3.**

1. For any  $\varepsilon \in (0, 1)$  and  $x \in \mathbb{R}^d$  we have that  $Y_t^\varepsilon(x)$  converges in distribution to  $\mu_*^\varepsilon$  as  $t \rightarrow \infty$ , where  $\varepsilon Z_\infty$  has the law of  $\mu_*^\varepsilon$ .
2. For any  $\varepsilon \in (0, 1]$ ,  $K > 0$ ,  $t \geq 0$  and  $(x_{\varepsilon,t})_{\varepsilon,t}$  with  $|x_{\varepsilon,t}| \leq K$  we have that  $Y_t^\varepsilon(x_{\varepsilon,t})$  converges in distribution to  $\mu_*^\varepsilon$  as  $t \rightarrow \infty$ , where  $\varepsilon Z_\infty$  has the law of  $\mu_*^\varepsilon$ .

**Proof.** We start with the proof of (1). Let  $x \in \mathbb{R}^d$  and  $\varepsilon \in (0, 1]$  be fixed. Recall that  $Y_t^\varepsilon(x) = \varphi_t^x + \varepsilon Y_t^x$  for any  $t \geq 0$  by (1.26), where  $Y^x = (Y_t^x)_{t \geq 0}$  is the solution of the stochastic differential equation

$$dY_t^x = -Db(\varphi_t^x)Y_t^x dt + dL_t \quad \text{with } Y_0 = 0$$

and  $(\varphi_t^x)_{t \geq 0}$  is the solution of (1.6). By the variation of constants formula, it is not hard to see that

$$Y_t^x = \Phi_t^{-1}(x) \int_0^t \Phi_s(x) dL_s \quad \text{for } t \geq 0,$$

where  $\Phi(x) := (\Phi_t(x))_{t \geq 0}$  is the solution of the matrix differential equation

$$\frac{d}{dt} \Phi_t(x) = \Phi_t(x) Db(\varphi_t^x) \quad \text{with } \Phi_0 = I_d.$$

In addition, examining the Wronskian at 0 we have that  $\det \Phi_t(x) \neq 0$  for all  $t \geq 0$ . The inverse matrix  $\Phi_t^{-1}(x)$  exists for any  $t \geq 0$  and  $\Phi^{-1}(x) := (\Phi_t^{-1}(x))_{t \geq 0}$  is the solution of the matrix differential equation

$$\frac{d}{dt} \Phi_t^{-1}(x) = -Db(\varphi_t^x) \Phi_t^{-1}(x) \quad \text{with } \Phi_0 = I_d.$$

Recall that  $Z = (Z_t)_{t \geq 0}$  is the solution of

$$dZ_t = -Db(0)Z_t dt + dL_t \quad \text{with } Z_0 = 0.$$

Since this equation is also linear, its solution  $Z$  is also given explicitly by the variation-of-constants formula

$$Z_t = \Psi_t^{-1} \int_0^t \Psi_s dL_s \quad \text{for any } t \geq 0,$$

where  $\Psi := (\Psi_t)_{t \geq 0}$  is the solution of the matrix differential equation

$$\frac{d}{dt} \Psi_t = \Psi_t Db(0) \quad \text{with } \Psi_0 = I_d.$$

Note that the inverse matrix  $\Psi_t^{-1}$  exists for any  $t \geq 0$  and  $\Psi^{-1} := (\Psi_t^{-1})_{t \geq 0}$  satisfies the matrix differential equation

$$\frac{d}{dt} \Psi_t^{-1} = -Db(0) \Psi_t^{-1} \quad \text{with } \Psi_0 = I_d.$$

Since we are interested in convergence in distribution, we analyze the characteristic function of  $Y_t^x$  and  $Z_t$  for  $t > 0$ . By Theorem 3.1 in [66] we know that

$$\mathbb{E} \left[ e^{i\langle z, Z_t \rangle} \right] = \exp \left( \int_0^t \psi(\Psi_s^*(\Psi_t^{-1})^* z) ds \right) \quad \text{for } z \in \mathbb{R}^d, \tag{C.2}$$

where  $\psi : \mathbb{R}^d \rightarrow \mathbb{C}$  is the characteristic exponent of the Lévy process  $L$ . For the inhomogeneous process  $(Y_t^x)_{t \geq 0}$ , a standard discretization procedure combined with (C.2) yields

$$\mathbb{E} \left[ e^{i\langle z, Y_t^x \rangle} \right] = \exp \left( \int_0^t \psi(\Phi_s^*(x)(\Phi_t^{-1}(x))^* z) ds \right) \quad \text{for } z \in \mathbb{R}^d.$$

By Lemma A.3 part iii) we note that there exist uniform positive constants  $c_3$  and  $c_4$  such that

$$|\Phi_s^*(x)(\Phi_t^{-1}(x))^* z| \leq c_3 e^{-c_4(t-s)} |z| \quad \text{and} \quad |\Psi_s^*(\Psi_t^{-1})^* z| \leq c_3 e^{-c_4(t-s)} |z|$$

for any  $t \geq 0, s \in [0, t]$  and  $z \in \mathbb{R}^d$ . For  $t > 0$  and  $z \in \mathbb{R}^d$ , we define the error term by

$$\Theta_t^z(x) := \int_0^t \psi(\Phi_s^*(x)(\Phi_t^{-1}(x))^* z) ds - \int_0^t \psi(\Psi_s^*(\Psi_t^{-1})^* z) ds.$$

Since we are assuming that the Lévy process  $L = (L_t)_{t \geq 0}$  has  $\beta$ -moment for some  $\beta > 0$  (see Hypothesis 2), the characteristic exponent  $\psi$  is differentiable for  $\beta \geq 1$  and locally Hölder continuous with index  $\beta$  for  $\beta \in (0, 1)$ , a proof is given in Proposition 7 in Appendix C. Let  $z \in \mathbb{R}^d$  with  $|z| \leq \frac{1}{2c_3}$ . Then there exists a positive constant  $C_1 := C(c_3, \beta)$  such that

$$\begin{aligned} |\Theta_t^z(x)| &\leq \int_0^t |\psi(\Phi_s^*(x)(\Phi_t^{-1}(x))^* z) - \psi(\Psi_s^*(\Psi_t^{-1})^* z)| ds \\ &\leq C_1 \int_0^t |\Phi_s^*(x)(\Phi_t^{-1}(x))^* - \Psi_s^*(\Psi_t^{-1})^*|^{\beta \wedge 1} ds \quad \text{for } t \geq 0. \end{aligned} \tag{C.3}$$

By Lemma A.3, part v), there is a positive constant  $C = C(|x|)$  such that

$$|\Phi_s^*(x)(\Phi_t^{-1}(x))^* - (\Psi_s)^*(\Psi_t^{-1})^*|^2 \leq \frac{C^2(|x|)d^3}{4\delta^2} |\varphi_0^x|^2 e^{-\delta t} (1 - e^{-4\delta(t-s)}) \leq C_2(|x|) e^{-\delta t}$$

for any  $s \in [0, t]$ , where  $C_2(|x|)$  is a constant that depends continuously on  $|x|$ . Using then the preceding inequality in (C.3) we obtain

$$|\Theta_t^z(x)| \leq C_3(|x|) e^{-\frac{\delta}{2}(\beta \wedge 1)t} \quad \text{for } t \geq 0, \tag{C.4}$$

where  $C_3(|x|)$  is a constant that depends continuously on  $|x|$ . Sending  $t \rightarrow \infty$ , we obtain  $\Theta_t^z(x) \rightarrow 0$  for any  $|z| \leq \frac{1}{2c_3}$ . In the sequel, we prove  $Y_t^x \xrightarrow{d} Z_\infty$ . By Theorem 4.1 in [66] we know that  $Z_t \xrightarrow{d} Z_\infty$ , that is,

$$\lim_{t \rightarrow \infty} \mathbb{E} \left[ e^{i\langle z, Z_t \rangle} \right] = \exp \left( \int_0^\infty \psi(e^{-Db(0)s} z) ds \right) =: \chi(z) \quad \text{for } z \in \mathbb{R}^d. \tag{C.5}$$

Recall that for each  $t > 0, Z_t$  is infinitely divisible (see for instance Theorem 9.1 in [64]), then  $\mathbb{E}[e^{i\langle z, Z_t \rangle}] \neq 0$  for any  $z \in \mathbb{R}^d$  (see Lemma 7.5 in [64]). Hence, (C.4) implies

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E} \left[ e^{i\langle z, Y_t^x \rangle} \right]}{\mathbb{E} \left[ e^{i\langle z, Z_t \rangle} \right]} = \lim_{t \rightarrow \infty} \exp(\Theta_t^z(x)) = 1 \quad \text{for } |z| \leq \frac{1}{2c_3}. \tag{C.6}$$

By (C.5) we infer  $\lim_{t \rightarrow \infty} \mathbb{E} \left[ e^{i\langle z, Y_t^x \rangle} \right] = \chi(z)$  for  $|z| \leq \frac{1}{2c_3}$ . Since  $\chi$  is a characteristic function, it is uniquely determined by its values in an open neighborhood of the origin. As a result we obtain  $\lim_{t \rightarrow \infty} \mathbb{E} \left[ e^{i\langle z, Y_t^x \rangle} \right] = \chi(z)$  for any  $z \in \mathbb{R}^d$ . By the Lévy continuity theorem we obtain  $Y_t^x \xrightarrow{d} Z_\infty$ . Recall that  $Y_t^\varepsilon(x) = \varphi_t^x + \varepsilon Y_t^x, t \geq 0$ . Since  $\varphi_t^x \rightarrow 0$ , as  $t \rightarrow \infty$ , the Slutsky lemma yields  $Y_t^\varepsilon(x) \xrightarrow{d} \varepsilon Z_\infty$  as  $t \rightarrow \infty$ . This finishes the proof of (1).

We finish with the proof of (2). By (C.4) we have that the convergence (C.6) only depends of  $x$  via  $|x| \leq K$  and consequently is valid for all  $|x_{t,\varepsilon}| \leq K$  as stated in (2).

**C.3.2 Convergence of  $Y_t^\varepsilon(x)$  to  $\mu_*^\varepsilon$  in the total variation distance**

**Lemma C.4.** For any  $K > 0$  and  $\varepsilon > 0$  we have

$$\lim_{t \rightarrow \infty} \sup_{|x| \leq K} \|Y_t^\varepsilon(x) - \mu_*^\varepsilon\|_{TV} = 0, \tag{C.7}$$

where  $\varepsilon Z_\infty$  has the law of  $\mu_*^\varepsilon$ . In particular,

$$\lim_{t \rightarrow \infty} \sup_{|x| \leq K} \|Y_t^x - Z_\infty\|_{TV} = 0. \tag{C.8}$$

*Proof.* The idea is to show that convergence in distribution (Lemma C.3 (2)) combined with the Orey-Masuda cone condition (Lemma 1.7) implies  $\lim_{t \rightarrow \infty} \sup_{|x| \leq K} \|Y_t^\varepsilon(x) - \mu_*^\varepsilon\|_{TV} = 0$ . It is enough to prove that  $Z_\infty$  has a continuous density and  $\lim_{t \rightarrow \infty} \sup_{|x| \leq K} \|Y_t^x - Z_\infty\|_{TV} = 0$ . The latter implies  $\lim_{t \rightarrow \infty} \sup_{|x| \leq K} \|Y_t^\varepsilon(x) - \mu_*^\varepsilon\|_{TV} = 0$ . Let  $|x| \leq K$ . Indeed,

$$\begin{aligned} \|Y_t^\varepsilon(x) - \mu_*^\varepsilon\|_{TV} &\leq \|(\varphi_t^x + \varepsilon Y_t^x) - (\varphi_t^x + \varepsilon Z_\infty)\|_{TV} + \|(\varphi_t^x + \varepsilon Z_\infty) - (\varepsilon Z_\infty)\|_{TV} \\ &= \|Y_t^x - Z_\infty\|_{TV} + \|(\varphi_t^x/\varepsilon + Z_\infty) - Z_\infty\|_{TV} \\ &\leq \sup_{|x| \leq K} \|Y_t^x - Z_\infty\|_{TV} + \sup_{|x| \leq K} \|(\varphi_t^x/\varepsilon + Z_\infty) - Z_\infty\|_{TV}. \end{aligned} \tag{C.9}$$

We start with the second term in (C.9). Since  $Z_\infty$  has a continuous density (see for instance, Case 3 in Section 4 in [10]), the Scheffé lemma yields

$$\sup_{|x| \leq K} \|(\varphi_t^x/\varepsilon + Z_\infty) - Z_\infty\|_{TV} = \|(\varphi_t^{\tilde{x}}/\varepsilon + Z_\infty) - Z_\infty\|_{TV}$$

for some  $|\tilde{x}| \leq K$ . By Hypothesis 1 we have  $|\varphi_t^x| \leq e^{-\delta t}|x| \leq e^{-\delta t}K$  whenever  $|x| \leq K$ . Again by the Scheffé lemma we deduce

$$\|(\varphi_t^{\tilde{x}}/\varepsilon + Z_\infty) - Z_\infty\|_{TV} \rightarrow 0, \quad t \rightarrow \infty.$$

The latter, together with inequality (C.9) and  $\sup_{|x| \leq K} \|Y_t^x - Z_\infty\|_{TV} \rightarrow 0$  for  $t \rightarrow \infty$  implies

$$\sup_{|x| \leq K} \|Y_t^\varepsilon(x) - \varepsilon Z_\infty\|_{TV} \rightarrow 0 \quad \text{for } t \rightarrow \infty.$$

In the sequel, we dominate  $\sup_{|x| \leq K} \|Y_t^x - Z_\infty\|_{TV}$ . By Lemma C.2 we have

$$\sup_{|x| \leq K} \|Y_t^x - Z_\infty\|_{TV} = \|Y_t^{\tilde{x}} - Z_\infty\|_{TV}$$

for some  $|\tilde{x}| \leq K$ . For convenience of notation we drop the tilde and write  $x$ . The proof is divided in 2 steps.

**Step 1.** We start with the proof that for any  $0 < t \leq \infty$ ,  $Y_t^x$  has a continuous density. From Theorem 28.1 in [64], it is sufficient to show that

$$\int_{\mathbb{R}^d} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| dz < \infty \quad \text{for all } t > 0. \tag{C.10}$$

Fix  $t > 0$ . Since

$$\mathbb{E}[e^{i\langle z, Y_t^x \rangle}] = \exp\left(\int_0^t \psi(\Phi_s^*(x)(\Phi_t^{-1}(x))^* z) ds\right) \quad \text{for } z \in \mathbb{R}^d,$$

we have

$$\begin{aligned} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| &\leq \exp\left(\int_0^t \int_{\mathbb{R}^d} (\cos(\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z, \theta \rangle) - 1) \nu(d\theta) ds)\right) \\ &\leq \exp\left(\int_0^t \int_{|\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z, \theta \rangle)| \leq \pi} (\cos(\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z, \theta \rangle) - 1) \nu(d\theta) ds)\right) \\ &\leq \exp\left(-2 \int_0^t \int_{|\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z, \theta \rangle)| \leq 1} |\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z, \theta \rangle)|^2 \nu(d\theta) ds)\right), \end{aligned} \tag{C.11}$$

where the last inequality follows from the well-known inequality  $1 - \cos(x) \geq 2(\frac{x}{\pi})^2$  for  $|x| \leq \pi$ . By Lemma A.3 we know that there exist positive constants  $c_1, c_2(|x|)$  such that

$$c_1 e^{-c_2(|x|)(t-s)} |z| \leq |\Phi_s^*(x)(\Phi_t^{-1}(x))^* z| \tag{C.12}$$

for any  $t \geq 0, s \in [0, t]$  and  $z \in \mathbb{R}^d$ . Due to the boundedness of the characteristic function it is enough to prove that

$$\int_{|z| > R} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| dz < \infty \quad \text{for some } R > 0.$$

Let  $R > C_{\triangleleft} \frac{\pi e^{c_2(|x|)t}}{c_1}$  and  $|z| \geq R$ , where  $C_{\triangleleft}$  is given in Lemma 1.7. Then

$$\frac{|\Phi_s^*(x)(\Phi_t^{-1}(x))^* z|}{\pi} \geq \frac{c_1 e^{-c_2(|x|)(t-s)}}{\pi} |z| \geq C_{\triangleleft} e^{c_2(|x|)s} \geq C_{\triangleleft} \quad \text{for any } s \in [0, t].$$

By (C.11) and Lemma 1.7 we obtain

$$\begin{aligned} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| &\leq \exp\left(-\frac{2c_{\triangleleft} c_1^\alpha |z|^\alpha}{\pi^\alpha} \int_0^t e^{-c_2(|x|)\alpha(t-s)} ds\right) \\ &\leq \exp\left(-\frac{2c_{\triangleleft} c_1^\alpha |z|^\alpha}{\pi^\alpha c_2(|x|)\alpha} (1 - e^{-c_2(|x|)\alpha t})\right) \end{aligned}$$

for any  $|z| \geq R$ , which implies the existence of  $C_b^\infty$  density (see for instance Theorem 28.1 in [64]). For  $Y_\infty$ , we just notice that  $Y_\infty = Z_\infty$  in distribution and  $Z_\infty$  has a  $C_b^\infty$  density (see Case 3 Section 4 in [10]).

**Step 2.** Convergence in total variation. We prove  $\|Y_t^x - Z_\infty\|_{TV} \rightarrow 0$ , as  $t \rightarrow \infty$ . For any  $R > 0$  fixed we split

$$\int_{\mathbb{R}^d} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}] - \mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz = \left(\int_{|z| \leq R} + \int_{|z| > R}\right) |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}] - \mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz.$$

By Lemma C.3 and the uniform convergence of the characteristic functions on compact sets we have that

$$\lim_{t \rightarrow \infty} \int_{|z| \leq R} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}] - \mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz = 0.$$

Note that

$$\int_{|z| > R} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}] - \mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz \leq \int_{|z| > R} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| dz + \int_{|z| > R} |\mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz.$$

It is easy to see that the Orey-Masuda condition implies condition (H) in [10]. In the proof of Proposition 5.3 there it is shown that under condition (H) we have

$$\lim_{R \rightarrow \infty} \limsup_{t \rightarrow \infty} \int_{|z| > R} |\mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz = 0.$$



Therefore, the limit

$$\lim_{R \rightarrow \infty} \limsup_{t \rightarrow \infty} \int_{|z| > R} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| dz = 0 \tag{C.13}$$

yields the desired result. Indeed,

$$\begin{aligned} \limsup_{t \rightarrow \infty} \int_{\mathbb{R}^d} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}] - \mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz \\ \leq \limsup_{t \rightarrow \infty} \int_{|z| > R} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| dz + \limsup_{t \rightarrow \infty} \int_{|z| > R} |\mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz, \end{aligned}$$

where the left-hand side does not depend on  $R$ . Sending  $R \rightarrow \infty$  we obtain

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}^d} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}] - \mathbb{E}[e^{i\langle z, Z_\infty \rangle}]| dz = 0.$$

In the sequel, we prove inequality (C.13). By inequality (C.11) we have

$$|\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| \leq \exp \left( -2 \int_0^t \int_{|\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle| \leq 1} |\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle|^2 \nu(d\theta) ds \right).$$

By Lemma A.3 we know that there exist positive constants  $c_1, c_2(|x|)$  such that

$$c_1 e^{-c_2(|x|)(t-s)} |z| \leq |\Phi_s^*(x)(\Phi_t^{-1}(x))^* z| \quad \text{for } t \geq s, s \in [0, t], z \in \mathbb{R}^d.$$

Let  $R > \frac{\pi C_{\triangleleft}}{C_1}$ ,  $t > \frac{1}{c_2(|x|)} \ln \left( \frac{C_1 R}{\pi} \right) =: t_0(R)$ ,  $s \in [0, t]$  and  $|z| \geq R$ , where  $C_{\triangleleft}$  is given in Lemma 1.7. Then we obtain

$$\frac{|\Phi_s^*(x)(\Phi_t^{-1}(x))^* z|}{\pi} \geq \frac{c_1 e^{-c_2(|x|)(t-s)} R}{\pi} \geq C_{\triangleleft} \quad \text{whenever } s \in [t - t_0(R), t].$$

Observe that

$$\begin{aligned} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| &\leq \exp \left( -2 \int_0^{t-t_0(R)} \int_{|\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle| \leq 1} |\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle|^2 \nu(d\theta) ds \right) \\ &\quad \cdot \exp \left( -2 \int_{t-t_0(R)}^t \int_{|\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle| \leq 1} |\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle|^2 \nu(d\theta) ds \right) \\ &\leq \exp \left( -2 \int_{t-t_0(R)}^t \int_{|\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle| \leq 1} |\langle (\Phi_s^*(x)(\Phi_t^{-1}(x))^* z / \pi, \theta) \rangle|^2 \nu(d\theta) ds \right). \end{aligned}$$

The Orey-Masuda cone condition (Lemma 1.7) and equality (C.12) yield

$$\begin{aligned} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| &\leq \exp \left( -\frac{2C_{\triangleleft} c_1^\alpha |z|^\alpha}{\pi^\alpha} \int_0^t e^{-c_2(|x|)\alpha(t-s)} ds \right) \\ &\leq \exp \left( -\tilde{c}(|x|) |z|^\alpha (1 - e^{-c_2(|x|)\alpha t_0(R)}) \right) \end{aligned}$$

for any  $|z| \geq R$  and  $t > t_0(R)$ , where  $\tilde{c}(|x|) := \frac{2c_{\triangleleft} c_1^\alpha}{\pi^\alpha c_2(|x|)^\alpha} > 0$ . Therefore for  $|x| \leq K$  there are positive constants  $\tilde{c}(K)$  and  $c_2(K)$  such that

$$\limsup_{t \rightarrow \infty} \int_{|z| \geq R} |\mathbb{E}[e^{i\langle z, Y_t^x \rangle}]| dz \leq \int_{|z| \geq R} \exp \left( -\tilde{c}(K) |z|^\alpha (1 - e^{-c_2(K)\alpha t_0(R)}) \right) dz$$

for any  $R > \frac{\pi C_{\triangleleft}}{C_1}$ . Sending  $R \rightarrow \infty$ , the dominated convergence theorem implies (C.13).  $\square$

**C.4 Geometric profile characterization for rotationally invariant  $Z_\infty$**

**Lemma C.5.** *Let  $f \in C^1(\mathbb{R}^d, (0, \infty))$  be a smooth density such that  $f(z) = g(|z|)$  for some function  $g \in C^1((0, \infty), (0, \infty))$  with  $g'(s) < 0$  for all  $s > 0$  and  $g' \in L^1(\mathbb{R}^d)$ . Then the map*

$$(0, \infty) \ni r \mapsto \int_{\mathbb{R}^d} |f(z + re_1) - f(z)| dz \in (0, \infty)$$

is strictly increasing. In particular, it is injective.

*Proof.* First we rewrite

$$\int_{\mathbb{R}^d} |f(z + re_1) - f(z)| dz = 2 - 2 \int_{\mathbb{R}^d} (f(z + re_1) \wedge f(z)) dz.$$

By the definition of the minimum we have

$$\frac{d}{dr} (f(z + re_1) \wedge f(z)) = \begin{cases} 0 & \text{for all } z : f(z) < f(z + re_1), \\ \frac{d}{dr} f(z + re_1) & \text{for all } z : f(z) > f(z + re_1). \end{cases}$$

In the sequel we determine the shape of  $\{z \in \mathbb{R}^d \mid f(z) > f(z + re_1)\}$ . Since

$$f(z) < f(\tilde{z}) \quad \text{if and only if} \quad |z| > |\tilde{z}|,$$

the continuity of  $f$  yields

$$f(z + re_1) \wedge f(z) = \begin{cases} f(z) & \text{for all } z : |z + re_1| < |z|, \\ f(z + re_1) & \text{for all } z : |z + re_1| > |z|. \end{cases}$$

That is, we obtain geometrically the shifted half space

$$\{z \in \mathbb{R}^d \mid |z + re_1|^2 > |z|^2\} = \{z \in \mathbb{R}^d \mid z_1 > -r/2\}.$$

Consequently, it follows

$$\frac{d}{dr} (f(z + re_1) \wedge f(z)) = \begin{cases} 0 & \text{for all } z : z_1 < -\frac{r}{2}, \\ \frac{d}{dr} f(z + re_1) & \text{for all } z : z_1 > -\frac{r}{2}. \end{cases}$$

We continue with the computation of  $\frac{d}{dr} f(z + re_1)$ . For all  $z \neq 0$  we have

$$\frac{d}{dr} f(z + re_1) = \frac{d}{dr} g(|z + re_1|) = g'(|z + re_1|) \frac{\langle z + re_1, e_1 \rangle}{|z + re_1|} = g'(r|\frac{z}{r} + e_1|) \frac{(1 + \langle \frac{z}{r}, e_1 \rangle)}{|e_1 + \frac{z}{r}|},$$

such that the Leibniz integral rule and the implicit function theorem imply

$$\begin{aligned} \frac{d}{dr} \int_{\mathbb{R}^d} f(z + re_1) \wedge f(z) dz &= \int_{\mathbb{R}^d} \frac{d}{dr} f(z + re_1) \wedge f(z) dz \\ &= \int_{\frac{z_1}{r} > -\frac{1}{2}} g'(r|\frac{z}{r} + e_1|) \frac{(1 + \frac{z_1}{r})}{|e_1 + \frac{z}{r}|} dz \\ &= \int_{v_1 > -\frac{1}{2}} \underbrace{g'(r|v + e_1|)}_{<0} \underbrace{\frac{(1 + v_1)}{|e_1 + v|}}_{>0} r^d dv < 0. \end{aligned}$$

Consequently, we obtain the desired result

$$\begin{aligned} \frac{d}{dr} \int_{\mathbb{R}^d} |f(z + re_1) - f(z)| dz &= -2 \frac{d}{dr} \int_{\mathbb{R}^d} f(z + re_1) \wedge f(z) dz \\ &= -2 \int_{v_1 > -\frac{1}{2}} \underbrace{g'(r|v + e_1|)}_{<0} \underbrace{\frac{(1 + v_1)}{|e_1 + v|}}_{>0} r^d dv > 0. \quad \square \end{aligned}$$

### D Exponential ergodicity of coercive Lévy SDEs in $L^\beta, \beta > 0$

In this section we fix the following standing assumptions. Let  $b$  be a vector field satisfying Hypothesis 1 and  $A$  a  $d$ -squared matrix with real entries. Consider a Lévy process  $L = (L_t)_{t \geq 0}$  with values in  $\mathbb{R}^d$  with strongly locally layered stable Lévy measure  $\nu$  with given parameters  $(\nu_0, \nu_\infty, \Lambda, q, c_0, \alpha)$  satisfying Hypotheses 2, 3 and 4 and the strong solution  $X = (X_t)_{t \geq 0}$  of the SDE

$$\begin{cases} dX_t = -b(X_t)dt + AdL_t & \text{for } t \geq 0, \\ X_0 = x \in \mathbb{R}^d. \end{cases} \tag{D.1}$$

**Definition D.1** (Hörmander condition, nonlinear Kalman rank condition).

Under the standing assumptions we denote by  $B_0 = I_d$  be the identity matrix on  $\mathbb{R}^d$  and define for  $n \in \mathbb{N}$  the  $(d \times d)$ -matrix-valued function  $B_n(x)$  recursively by

$$B_n(x) := -b(x) \cdot DB_{n-1}(x) + Db(x)B_{n-1}(x), \quad x \in \mathbb{R}^d,$$

where  $b \cdot DF := \sum_{k=1}^d b_k \frac{\partial}{\partial x_k} F$ , and  $F$  is a  $(d \times d)$ -matrix-valued function. We say that the SDE (D.1) satisfies a Hörmander condition if its coefficients  $b = (b_1, \dots, b_d)^*$  and the matrix  $A$  satisfy the following: For each  $x \in \mathbb{R}^d$  there exists some  $n = n(x) \in \mathbb{N} \cup \{0\}$  such that

$$\text{Rank}[B_0A, B_1(x)A, \dots, B_n(x)A] = d. \tag{D.2}$$

**Lemma D.2** (Orey-Masuda type condition). Under the standing assumptions, the limit (1.8) implies

$$\lim_{h \rightarrow 0} h^{\alpha-2} \int_{|z| \leq h} |z|^2 \nu(dz) := \kappa_1 > 0.$$

The latter is Condition (1.2) in [73].

*Proof.* Observe that

$$\int_{|z| \leq h} |z|^2 \nu(dz) = \int_{\mathbb{S}^{d-1}} \int_0^h r^2 q(r, \theta) dr \Lambda(d\theta).$$

Let  $\eta > 0$  be fixed. By limit (1.8) we deduce that there exists  $r_0 := r_0(\eta) \in (0, 1)$  such that for any  $0 < r < r_0$  we have

$$r^{1-\alpha}(c_0(\theta) - \eta) < r^2 q(r, \theta) < r^{1-\alpha}(c_0(\theta) + \eta) \quad \text{for any } \theta \in \mathbb{S}^{d-1}.$$

Let  $h \in (0, r_0)$ . Then

$$\begin{aligned} \frac{1}{2-\alpha} \int_{\mathbb{S}^{d-1}} (c_0(\theta) - \eta) \Lambda(d\theta) &\leq h^{\alpha-2} \int_{\mathbb{S}^{d-1}} \int_0^h r^2 q(r, \theta) dr \Lambda(d\theta) \\ &\leq \frac{1}{2-\alpha} \int_{\mathbb{S}^{d-1}} (c_0(\theta) + \eta) \Lambda(d\theta). \end{aligned}$$

Sending  $h \rightarrow 0$  followed by sending  $\eta \rightarrow 0$ , we obtain

$$\lim_{h \rightarrow 0} h^{\alpha-2} \int_{\mathbb{S}^{d-1}} \int_0^h r^2 q(r, \theta) dr \Lambda(d\theta) = \frac{1}{2-\alpha} \int_{\mathbb{S}^{d-1}} c_0(\theta) \Lambda(d\theta) > 0,$$

where the last inequality follows from the fact that  $c_0 : \mathbb{S}^{d-1} \rightarrow (0, \infty)$  and  $c_0 \in L^1(\Lambda)$ .  $\square$

In the sequel we extend Theorem 4.1 in [60] to  $L^\beta$  for arbitrary  $\beta > 0$ .

**Theorem 4** (Exponential ergodicity).

Under the standing assumptions and the Hörmander condition (D.2) there exists a unique invariant distribution  $\mu$  for (D.1) satisfying exponential ergodicity in the total variation distance.

*Proof.* For  $\beta \geq 2$ , it is the statement of Theorem 4.1 in [60]. Let  $\beta \in (0, 2)$ . We apply Theorem 2.1 in [60]. Therefore, we verify Conditions LC, H1 and H2 in [60], p. 2-3.

**Condition LC in [60].** We stress that the fulfilment of Condition LC only requires Hypothesis 1 and Hypothesis 2. We define  $|\cdot|_c := \sqrt{|x|^2 + c^2}$  for  $c > 0$  satisfying for all  $x \in \mathbb{R}^d$

$$c \leq |x|_c \leq |x| + c, \quad \nabla|x|_c := \frac{x}{|x|_c} \quad \text{and} \quad 0 \leq \frac{|x|}{|x|_c} < 1.$$

In addition, we have

$$D^2|x|_c = \begin{pmatrix} \frac{|x|_c^2 - x_1^2}{|x|_c^3} & -\frac{x_1 x_2}{|x|_c^3} & -\frac{x_1 x_3}{|x|_c^3} & \cdots & -\frac{x_1 x_d}{|x|_c^3} \\ -\frac{x_1 x_2}{|x|_c^3} & \frac{|x|_c^2 - x_2^2}{|x|_c^3} & -\frac{x_2 x_3}{|x|_c^3} & & \\ -\frac{x_1 x_3}{|x|_c^3} & -\frac{x_2 x_3}{|x|_c^3} & \frac{|x|_c^2 - x_3^2}{|x|_c^3} & & \vdots \\ \vdots & & & \ddots & \\ -\frac{x_1 x_d}{|x|_c^3} & \cdots & & & \frac{|x|_c^2 - x_d^2}{|x|_c^3} \end{pmatrix}.$$

Consequently,

$$\begin{aligned} \|D^2|x|_c\|_1 &= \sum_{i,j} |(D^2|x|_c)_{ij}| = \sum_i |(D^2|x|_c)_{i,i}| + \sum_{i \neq j} |(D^2|x|_c)_{ij}| \\ &= \frac{d|x|_c^2 - \sum_i x_i^2 + \sum_{i \neq j} |x_i x_j|}{|x|_c^3} \leq \frac{1}{|x|_c} \left( d + \frac{|x|^2}{|x|_c^2} \right) \leq \frac{d+1}{c}. \end{aligned}$$

Let  $0 < \gamma \leq \beta \wedge 1$ . We calculate the gradient and the Hessian of  $|x|_c^\gamma$  as follows:

$$\nabla|x|_c^\gamma = \gamma|x|_c^{\gamma-1} \nabla|x|_c = \gamma|x|_c^{\gamma-1} \frac{x}{|x|_c} = \gamma|x|_c^{\gamma-2} x$$

and

$$\begin{aligned} \partial_{ii}|x|_c^\gamma &= \partial_i(\gamma|x|_c^{\gamma-2} x_i) = \gamma|x|_c^{\gamma-2} + \gamma(\gamma-2)|x|_c^{\gamma-4} x_i^2, \\ \partial_{ij}|x|_c^\gamma &= \partial_i(\gamma|x|_c^{\gamma-2} x_j) = \gamma x_j \partial_i(|x|_c^{\gamma-2}) = \gamma(\gamma-2)|x|_c^{\gamma-4} x_i x_j \quad \text{for } i \neq j. \end{aligned}$$

Hence,

$$\sum_{i,j} |\partial_{ij}|x|_c^\gamma| = \gamma(2-\gamma)|x|_c^{\gamma-4} \sum_{i,j} |x_i x_j| + d\gamma|x|_c^{\gamma-2} \gamma(2-\gamma)|x|_c^{\gamma-4} \|x\|^2 + d\gamma|x|_c^{\gamma-2},$$

where  $\|\cdot\|$  denotes the 1-norm. Since  $\|x\| \leq \sqrt{d}|x|$ , we obtain

$$\sum_{i,j} |\partial_{ij}|x|_c^\gamma| \leq (\gamma(2-\gamma)d + d\gamma)|x|_c^{\gamma-2} \leq (\gamma(2-\gamma)d + d\gamma)c^{\gamma-2}.$$

With the help of the preceding calculations Itô's formula yields

$$\begin{aligned} |X_t|_c^\gamma &= |x|_c^\gamma - \gamma \int_0^t \langle |X_s|_c^{\gamma-2} X_s, b(X_s) \rangle ds + \int_0^t \int_{|z|<1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) \tilde{N}(dsdz) \\ &\quad + \int_0^t \int_{|z|\geq 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) N(dsdz) \\ &\quad + \int_0^t \int_{|z|<1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma - \langle \gamma |X_s|_c^{\gamma-2} X_s, Az \rangle) \nu(dz) ds, \end{aligned}$$

where  $N$  is a Poisson random measure with compensator  $dt \otimes \nu(dz)$ . Moreover, we have the Lévy-Iô decomposition such that  $\mathbb{P}$ -a.s. for all  $t \geq 0$

$$L_t = \int_0^t \int_{|z| \leq 1} z \tilde{N}(dsdz) + \int_0^t \int_{|z| > 1} z N(dsdz),$$

where  $\tilde{N}$  is the compensated version of  $N$ . Taking expectations we obtain

$$\begin{aligned} \mathbb{E}[|X_t|_c^\gamma] &= |x|_c^\gamma - \gamma \int_0^t \mathbb{E}[\langle |X_s|_c^{\gamma-2} X_s, b(X_s) \rangle] ds \\ &\quad + \mathbb{E} \left[ \int_0^t \int_{|z| < 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) \tilde{N}(dsdz) \right] \\ &\quad + \mathbb{E} \left[ \int_0^t \int_{|z| \geq 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) N(dsdz) \right] \\ &\quad + \int_0^t \mathbb{E} \left[ \int_{|z| < 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma - \langle \gamma |X_s|_c^{\gamma-2} X_s, Az \rangle) \nu(dz) \right] ds. \end{aligned}$$

First, since the moment of order  $\beta$  is finite, a localization argument yields

$$\mathbb{E} \left[ \int_0^t \int_{|z| < 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) \tilde{N}(dsdz) \right] = 0 \quad \text{for } t \geq 0.$$

Secondly, by the Itô isometry for Poisson random measures (see [4]) we obtain

$$\begin{aligned} \mathbb{E} \left[ \int_0^t \int_{|z| \geq 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) N(dsdz) \right] \\ = \mathbb{E} \left[ \int_0^t \int_{|z| \geq 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) \nu(dz) ds \right] \\ = \int_0^t \mathbb{E} \left[ \int_{|z| \geq 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) \nu(dz) \right] ds. \end{aligned}$$

Hence, for almost all  $t$  we have

$$\begin{aligned} \mathbb{E}[|X_t|_c^\gamma] &= |x|_c^\gamma - \gamma \int_0^t \mathbb{E}[\langle |X_s|_c^{\gamma-2} X_s, b(X_s) \rangle] ds \\ &\quad + \int_0^t \mathbb{E} \left[ \int_{|z| \geq 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma) \nu(dz) \right] ds \\ &\quad + \int_0^t \mathbb{E} \left[ \int_{|z| < 1} (|X_{s-} + Az|_c^\gamma - |X_{s-}|_c^\gamma - \langle \gamma |X_s|_c^{\gamma-2} X_s, Az \rangle) \nu(dz) \right] ds. \end{aligned}$$

Taking derivatives we obtain

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[|X_t|_c^\gamma] &= -\gamma \mathbb{E}[\langle |X_s|_c^{\gamma-2} X_s, b(X_t) \rangle] + \mathbb{E} \left[ \int_{|z| \geq 1} (|X_{t-} + Az|_c^\gamma - |X_{t-}|_c^\gamma) \nu(dz) \right] \\ &\quad + \mathbb{E} \left[ \int_{|z| < 1} (|X_{t-} + Az|_c^\gamma - |X_{t-}|_c^\gamma - \langle \gamma |X_s|_c^{\gamma-2} X_s, Az \rangle) \nu(dz) \right]. \end{aligned}$$

By Hypothesis 1 it follows that

$$\begin{aligned} -\gamma |X_s|_c^{\gamma-2} \langle X_s, b(X_s) \rangle &\leq -\delta \gamma |X_s|_c^{\gamma-2} |X_s|^2 = -\delta \gamma |X_s|_c^{\gamma-2} (|X_s|_c^2 - c^2) \\ &= -\delta \gamma |X_s|_c^\gamma + \delta \gamma c^2 |X_s|_c^{\gamma-2} \leq -\delta \gamma |X_s|_c^\gamma + \delta \gamma c^\gamma. \end{aligned}$$

Hence

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[|X_t|_c^\gamma] &\leq -\delta\gamma \mathbb{E}[|X_t|_c^\gamma] + \delta\gamma c^\gamma + \mathbb{E}\left[\int_{|z|\geq 1} (|X_{t-} + Az|_c^\gamma - |X_{t-}|_c^\gamma) \nu(dz)\right] \\ &\quad + \mathbb{E}\left[\int_{|z|<1} (|X_{t-} + Az|_c^\gamma - |X_{t-}|_c^\gamma - \langle \gamma |X_s|_c^{\gamma-2} X_s, Az \rangle) \nu(dz)\right]. \end{aligned}$$

For  $\gamma \in (0, 1]$ , the subadditivity of the power of order  $\gamma$  yields

$$\begin{aligned} \mathbb{E}\left[\int_{|z|\geq 1} (|X_{t-} + Az|_c^\gamma - |X_{t-}|_c^\gamma) \nu(dz)\right] &\leq \int_{|z|\geq 1} |Az|_c^\gamma \nu(dz) \\ &\leq \|A\|^\gamma \int_{|z|\geq 1} |z|^\gamma \nu(dz) = \|A\|^\gamma C_1, \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}\left[\int_{|z|<1} (|X_{t-} + Az|_c^\gamma - |X_{t-}|_c^\gamma - \langle \gamma |X_s|_c^{\gamma-2} X_s, Az \rangle) \nu(dz)\right] \\ \leq (\gamma(2-\gamma)d + d\gamma)c^{\gamma-2} \int_{|z|<1} |Az|^2 \nu(dz) \\ \leq \|A\|^2 c^{\gamma-2} (\gamma(2-\gamma)d + d\gamma) \int_{|z|<1} |z|^2 \nu(dz) =: \|A\|^2 c^{\gamma-2} C_2. \end{aligned}$$

Therefore we have

$$\frac{d}{dt} \mathbb{E}[|X_t|_c^\gamma] \leq -\delta\gamma \mathbb{E}[|X_t|_c^\gamma] + \delta\gamma c^\gamma + \|A\|^\gamma C_1 + \|A\|^2 c^{\gamma-2} C_2$$

and the Grönwall lemma yields

$$\begin{aligned} \mathbb{E}[|X_t|^\gamma] &\leq \mathbb{E}[|X_t|_c^\gamma] \leq |x|_c^\gamma e^{-\delta\gamma t} + \frac{1 - e^{-\delta\gamma t}}{\delta\gamma} (\delta\gamma c^\gamma + \|A\|^\gamma C_1 + \|A\|^2 c^{\gamma-2} C_2) \\ &\leq |x|_c^\gamma e^{-\delta\gamma t} + \frac{1}{\delta\gamma} (\delta\gamma c^\gamma + \|A\|^\gamma C_1 + \|A\|^2 c^{\gamma-2} C_2) \\ &\leq |x|^\gamma e^{-\delta\gamma t} + C_3, \end{aligned} \tag{D.3}$$

where  $C_3 = c^\gamma + \frac{1}{\delta\gamma} (\delta\gamma c^\gamma + \|A\|^\gamma C_1 + \|A\|^2 c^{\gamma-2} C_2)$ .

**Condition  $H_1$  in [60].** We emphasize that Condition  $H_1$  also only requires Hypothesis 1 and Hypothesis 2. In the sequel we consider the solution  $(X_t(x))_{t \geq 0}$  of (D.1) with initial condition  $x$ . By Hypothesis 1 we have for all  $x, y \in \mathbb{R}^d$

$$\begin{aligned} \frac{d}{dt} |X_t(x) - X_t(y)|_c^\gamma &= -\gamma |X_t(x) - X_t(y)|_c^{\gamma-2} \langle X_t(x) - X_t(y), b(X_t(x)) - b(X_t(y)) \rangle \\ &\leq -\gamma \delta |X_t(x) - X_t(y)|_c^{\gamma-2} |X_t(x) - X_t(y)|^2 \\ &= -\gamma \delta |X_t(x) - X_t(y)|_c^{\gamma-2} (|X_t(x) - X_t(y)|_c^2 - c^2) \\ &= -\gamma \delta |X_t(x) - X_t(y)|_c^\gamma + \gamma \delta |X_t(x) - X_t(y)|_c^{\gamma-2} c^2 \\ &\leq -\gamma \delta |X_t(x) - X_t(y)|_c^\gamma + \gamma \delta c^\gamma, \end{aligned} \tag{D.4}$$

where in the last inequality we use that  $|x|_c \geq c$  and  $\gamma \in (0, 1]$ . Grönwall's lemma applied to (D.4) yields

$$\mathbb{E}[|X_t(x) - X_t(y)|^\gamma] \leq \mathbb{E}[|X_t(x) - X_t(y)|_c^\gamma] \leq |x - y|_c^\gamma e^{-\delta\gamma t} + c^\gamma \leq |x - y|^\gamma e^{-\delta\gamma t} + 2c^\gamma. \tag{D.5}$$

Let  $R > 0$  and  $\Delta > 0$ . Here, we analyze the quantity

$$\mathbb{P}\left(\underbrace{|X_t(x) - X_t(y)| \leq \Delta}_{=:D}, \underbrace{|X_t(x)| \leq R_0}_{=:B}, \underbrace{|X_t(y)| \leq R_0}_{=:C}\right),$$

where  $x, y \in \overline{B}_{R_0}(0)$  for a suitable  $R_0 > 0$ . Observe that

$$\mathbb{P}((D \cap B \cap C)^c) = \mathbb{P}(D^c \cup B^c \cup C^c) \leq \mathbb{P}(D^c) + \mathbb{P}(B^c) + \mathbb{P}(C^c). \quad (\text{D.6})$$

For any  $\Delta > 0$  and  $R > 0$  we set  $R_0 = \max\{R, (4C_3)^{1/\gamma}\}$ , where  $C_3$  is the positive constant in estimate (D.3) and take  $x, y \in \overline{B}_R(0)$ . By estimate (D.3) we obtain for any  $t \geq 0$

$$\begin{aligned} \mathbb{P}(B^c) &= \mathbb{P}(|X_t(x)| > R_0) \leq \frac{\mathbb{E}[|X_t(x)|^\gamma]}{R_0^\gamma} \\ &\leq \frac{|x|^\gamma e^{-\delta\gamma t} + C_3}{R_0^\gamma} \leq \frac{R^\gamma e^{-\delta\gamma t} + C_3}{R_0^\gamma} = \frac{R^\gamma}{R_0^\gamma} e^{-\delta\gamma t} + \frac{C_3}{R_0^\gamma} \leq e^{-\delta\gamma t} + \frac{1}{4}. \end{aligned} \quad (\text{D.7})$$

Switching the role of  $x$  and  $y$  we have

$$\mathbb{P}(C^c) = \mathbb{P}(|X_t(y)| > R_0) \leq e^{-\delta\gamma t} + \frac{C_3}{R_0^\gamma} = e^{-\delta\gamma t} + \frac{1}{4}. \quad (\text{D.8})$$

We continue with the analysis of  $\mathbb{P}(D^c)$ . Again, let  $x, y \in \overline{B}_R(0)$ . Define  $T'_0(\Delta, R)$  as the unique positive solution of

$$\frac{((2R)^\gamma + c^\gamma)e^{-\delta\gamma T'_0}}{\Delta^\gamma} = \frac{1}{4},$$

where  $c^\gamma = \frac{\Delta^\gamma}{8}$ . Hence for any  $x, y \in \overline{B}_R(0)$  and  $t \geq T_0$  we have by (D.5) the estimate

$$\begin{aligned} \mathbb{P}(D^c) &= \mathbb{P}(|X_t(x) - X_t(y)| > \Delta) \leq \frac{\mathbb{E}[|X_t(x) - X_t(y)|^\gamma]}{\Delta^\gamma} \leq \frac{|x - y|^\gamma e^{-\delta\gamma t} + c^\gamma}{\Delta^\gamma} \\ &\leq \frac{((2R)^\gamma + c^\gamma)e^{-\delta\gamma t} + c^\gamma}{\Delta^\gamma} \leq \frac{3}{8}. \end{aligned} \quad (\text{D.9})$$

By taking  $T_0 = \max\{T'_0, \frac{\ln(32)}{\delta\gamma}\}$  and combining (D.6), (D.7), (D.8) and (D.9) it follows uniformly for any  $x, y \in \overline{B}_R(0)$  and  $t \geq T_0$  that

$$\mathbb{P}((D \cap B \cap C)^c) \leq 2e^{-\delta\gamma t} + \frac{7}{8} \leq \frac{15}{16} < 1.$$

The preceding inequality yields the weak form of irreducibility condition  $\mathbf{H}_1$  in [60] for the canonical coupling.

**Condition  $\mathbf{H}_2$  in [60].** The proof is virtually identical to [60] p.15-16.  $\square$

The following corollaries are Taylor-made statements for the error estimates in Subsection 2.4.

**Corollary D.3.** For any  $0 < \gamma \leq \beta \wedge 1$  there exists a positive constant  $C$  such that for all  $\varepsilon > 0$ ,  $x \in \mathbb{R}^d$  and  $t \geq 0$  we have

$$\mathbb{E}[|X_t^{\varepsilon, x}|^\gamma] \leq |x|^\gamma e^{-\delta\gamma t} + C\varepsilon^\gamma. \quad (\text{D.10})$$

*Proof.* The statement follows taking  $A = \varepsilon I_d$  and  $c = \varepsilon$  in inequality (D.3).  $\square$

**Corollary D.4.** For any  $0 < \gamma \leq \beta \wedge 1$  there exists a positive constant  $C = C(\delta, d, \gamma)$  such that

$$\mathbb{E}[|X_t^{\varepsilon,x}|^\gamma] \leq C\varepsilon^\gamma + |\varphi_t^x|^\gamma \tag{D.11}$$

for all  $t \geq 0$  and  $\varepsilon \in (0, 1]$ .

*Proof.* Note that the difference  $\tilde{X}_t^{\varepsilon,x} := X_t^{\varepsilon,x} - \varphi_t^x$  satisfies

$$d\tilde{X}_t^{\varepsilon,x} = -\left(\int_0^1 Db(\varphi_t^x + \theta\tilde{X}_t^{\varepsilon,x})d\theta\right)\tilde{X}_t^{\varepsilon,x}dt + \varepsilon dL_t, \quad \tilde{X}_0^{\varepsilon,x} = 0.$$

Hypothesis 1 together with the analogous computations to the proof of Condition **LC** in Theorem 4 in Appendix D yields that for  $A = \varepsilon I_d$ ,  $c = \varepsilon$  and  $0 < \gamma \leq 1 \wedge \beta$  there is a constant  $C > 0$  such that  $\varepsilon \in (0, 1]$ ,  $x \in \mathbb{R}^d$  and  $t \geq 0$  imply

$$\mathbb{E}[|\tilde{X}_t^{\varepsilon,x}|^\gamma] \leq C\varepsilon^\gamma.$$

Using the subadditivity of the  $\gamma$ -power we obtain (D.11). □

**Corollary D.5.** For any  $x \in \mathbb{R}^d$  and  $0 < \gamma \leq \beta \wedge 1$  there exists a positive constant  $C = C(|x|, \gamma)$  such that for all  $\vartheta \in (0, 1)$  and  $\varepsilon \in (0, 1)$  we have

$$\mathbb{P}(|X_{T_\varepsilon^x}^{\varepsilon,x}| \geq r_\varepsilon) \leq C(|x|)\varepsilon^{\gamma\vartheta}.$$

*Proof.* By Corollary D.4 and Lemma A.1 we have

$$\mathbb{E}[|X_{T_\varepsilon^x}^{\varepsilon,x}|^\gamma] \leq C_1\varepsilon^\gamma + C_2(|x|)\varepsilon^\gamma$$

for some positive constants  $C_1$  and  $C_2(|x|)$ . The preceding inequality with the help of Markov's inequality yields

$$\mathbb{P}(|X_{T_\varepsilon^x}^{\varepsilon,x}| \geq r_\varepsilon) \leq \frac{\mathbb{E}[|X_{T_\varepsilon^x}^{\varepsilon,x}|^\gamma]}{\varepsilon^{\gamma(1-\vartheta)}} \leq (C_1 + C_2(|x|))\varepsilon^{\gamma\vartheta},$$

which concludes the statement. □

## References

- [1] Aldous, D.: Random walks on finite groups and rapidly mixing Markov chains. Seminar on probability, XVII, 243–297, Lecture Notes in Math. **986**, Springer, Berlin, 1983. MR-0770418
- [2] Aldous, D. and Diaconis, P.: Shuffling cards and stopping times. *Amer. Math. Monthly* **93**, no. 5, (1986), 333–348. MR-841111
- [3] Aldous, D. and Diaconis, P.: Strong uniform times and finite random walks. *Adv. Appl. Math.* **8**, no. 1, (1987), 69–97. MR-0876954
- [4] Applebaum, D.: Infinite dimensional Ornstein-Uhlenbeck processes driven by Lévy processes. *Probab. Surv.* **12**, (2015), 33–54. MR-3385977
- [5] Barrera, G.: Abrupt convergence for a family of Ornstein Uhlenbeck processes. *Braz. J. Probab. Stat.* **32**, no. 1, (2018), 188–199. MR-3770869
- [6] Barrera, G., Högele, M. A. and Pardo, J. C.: Cutoff thermalization for Ornstein-Uhlenbeck systems with small Lévy noise in the Wasserstein distance. *J. Stat. Phys.* **184**, no. 27, (2021).
- [7] Barrera, G., Högele, M. A. and Pardo, J. C.: The cutoff phenomenon in Wasserstein distance for nonlinear stable Langevin systems with small Lévy noise, arXiv:2108.08351.
- [8] Barrera, G. and Jara, M.: Abrupt convergence of stochastic small perturbations of one dimensional dynamical systems. *J. Stat. Phys.* **163**, no. 1, (2016), 113–138. MR-3472096



- [9] Barrera, G. and Jara, M.: Thermalisation for small random perturbation of hyperbolic dynamical systems. *Ann. Appl. Probab.* **30**, no. 3, (2020), 1164–1208. MR-4133371
- [10] Barrera, G. and Pardo, J. C.: Cut-off phenomenon for Ornstein-Uhlenbeck processes driven by Lévy processes. *Electron. J. Probab.* **25**, no. 15, (2020), 1–33. MR-4073676
- [11] Barrera, J., Lachaud, B. and Ycart, B.: Cut-off for  $n$ -tuples of exponentially converging processes. *Stoch. Process. Appl.* **116**, no. 10, (2006), 1433–1446. MR-2260742
- [12] Barrera, J. and Ycart, B.: Bounds for left and right window cutoffs. *ALEA Lat. Am. J. Probab. Math. Stat.* **11**, no. 2, (2014), 445–458. MR-2260742
- [13] Basu R., Hermon, J. and Peres, Y.: Characterization of cutoff for reversible Markov chains. *Ann. Probab.* **45**, no. 3, (2017), 1448–1487. MR-3650406
- [14] Bayer, D. and Diaconis, P.: Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2**, no. 2, (1992), 294–313. MR-1161056
- [15] Ben-Hamou, A. and Salez, J.: Cutoff for non-backtracking random walks on sparse random graphs. *Ann. Probab.* **45**, no. 3, (2017), 1752–1770. MR-3650414
- [16] Ben-Hamou, A., Lubetzky, E. and Peres, Y.: Comparing mixing times on sparse random graphs. *Ann. Inst. Henri Poincaré Probab. Stat.* **55**, no. 2, (2019), 1116–1130. MR-3949966
- [17] Berestycki, N. *Mixing times of Markov chains: techniques and examples. A crossroad between probability, analysis and geometry.* University of Cambridge. <http://www.statslab.cam.ac.uk/~beresty/teach/Mixing/mixing3.pdf> Retrieved: 02.11.2020, 12:05 h. (EET).
- [18] Bertoncini, O., Barrera, J. and Fernández, R.: Cut-off and exit from metastability: two sides of the same coin. *C. R. Math. Acad. Sci. Paris* **346**, no. 11–12, (2008), 691–696. MR-2423280
- [19] Bordenave, C., Qiu, Y. and Zhang, Y.: Spectral gap of sparse bistochastic matrices with exchangeable rows. *Ann. Inst. Henri Poincaré Probab. Stat.* **56**, no. 4, (2020), 2971–2995. MR-4164862.
- [20] Bordenave, C., Caputo, P. and Salez, J.: Random walk on sparse random digraphs. *Probab. Theory Related Fields* **170**, no. 3–4, (2018), 933–960. MR-3773804
- [21] Bordenave, C., Caputo, P. and Salez, J.: Cutoff at the “entropic time” for sparse Markov chains. *Probab. Theory Related Fields* **173**, no. 1–2, (2019), 261–292. MR-3916108
- [22] Caballero, M. E., Pardo, J. C. and Pérez, J. L.: On Lamperti stable processes. *Probab. Math. Statist.* **30**, no. 1, (2010), 1–28. MR-2792485
- [23] Chae, M. and Walker, S. G.: Wasserstein upper bounds of the total variation for smooth densities. *Statist. Probab. Lett.* **163**, (2020), 108771, 6pp. MR-4083852
- [24] Chleboun, P. and Smith, A.: Cutoff for the square plaquette model on a critical length scale. *Ann. Appl. Probab.* **31**, no. 2, (2021), 668–702. MR-4254492
- [25] Chen, G. and Saloff-Coste, L.: The cutoff phenomenon for ergodic Markov processes. *Electron. J. Probab.*, **13**, no. 3, (2008), 26–78. MR-2375599
- [26] Dauxois, T.: Fermi, Pasta, Ulam, and a mysterious lady. *Physics Today* **61**, no. 1, (2008), 55–57. doi:10.1063/1.2835154
- [27] Debussche, A. and Fournier N.: Existence of densities for stable-like driven SDE’s with Hölder continuous coefficients. *J. Funct. Anal.* **264**, no. 8, (2013), 1757–1778. MR-3022725
- [28] Devroye, L. and Lugosi, G.: Combinatorial methods in density estimation. Springer Series in Statistics. Springer-Verlag, New York, 2001. xii+208 pp. MR-1843146
- [29] Diaconis, P.: Group representations in probability and statistics. Institute of Mathematical Statistics Lecture Notes-Monograph Series **11**. Institute of Mathematical Statistics, Hayward, CA, 1988. vi+198 pp. MR-0964069
- [30] Diaconis, P.: The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.* **93**, no. 4, (1996), 1659–1664. MR-1374011
- [31] Diaconis, P., Graham, R. L. and Morrison, J. A.: Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures Algorithms* **1**, no. 1, (1990), 51–72. MR-1068491
- [32] Diaconis, P. and Shahshahani, M.: Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.* **18**, no. 1, (1987), 208–218. MR-0871832

- [33] Dong, Z. and Peng, X.: Malliavin matrix of degenerate SDE and gradient estimate. *Electron. J. Probab.* **19**, no. 73, (2014), 1–26. MR-3256873
- [34] Eberle, A. and Zimmer, R.: Sticky couplings of multidimensional diffusions with different drifts. *Ann. Inst. Henri Poincaré Probab. Stat.* **55**, no. 4, (2019), 2370–2394. MR-4029157
- [35] Fermi, E., Pasta, J., Ulam, S. and Tsingou, M.: Studies of nonlinear problems. United States: N. p., (1955). doi:10.2172/4376203.
- [36] Fournier N. and Printems J.: Absolute continuity for some one-dimensional processes. *Bernoulli* **16**, no. 2, (2010), 343–360. MR-2668905
- [37] Freidlin, M. I. and Wentzell, A. D.: Random perturbations of dynamical systems. Translated from the 1979 Russian original by Joseph Szücs. Third edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 260. Springer, Heidelberg, 2012. xxviii+458 pp. MR-2953753
- [38] Grobman, D.: Homeomorphism of systems of differential equations. (Russian) *Dokl. Akad. Nauk SSSR* **128**, (1959), 880–881. MR-0121545
- [39] Hartman, P.: On local homeomorphisms of Euclidean spaces. *Bol. Soc. Mat. Mexicana* **5**, (1960), 220–241. MR-141856
- [40] Hermon, J. and Salez, J.: Cutoff for the mean-field zero-range process with bounded monotone rates. *Ann. Probab.* **48**, no. 2, (2020), 742–759. MR-4089493
- [41] Hermon, J. and Hutchcroft, T.: Supercritical percolation on nonamenable graphs: Isoperimetry, analyticity, and exponential decay of the cluster size distribution. *Invent. Math.* **224**, no. 2, (2021), 445–486. MR-4243018
- [42] Houdré, C. and Kawai, R.: On layered stable processes. *Bernoulli* **13**, no. 1, (2007), 252–278. MR-4243018
- [43] Ishikawa, Y., Kunita, H. and Tsuchiya, M.: Smooth density and its short time estimate for jump process determined by SDE. *Stoch. Process. Appl.* **128**, no. 9, (2018), 3181–3219. MR-3834856
- [44] Kallianpur, G. and Sundar, P.: Stochastic analysis and diffusion processes. Oxford Graduate Texts in Mathematics 24. Oxford University Press, Oxford, 2014. xii+352 pp. MR-3156223
- [45] Klenke, A.: Probability theory: A comprehensive course. Translated from the 2006 German original. Universitext. Springer-Verlag London, Ltd., London, 2008. xii+616 pp. MR-2372119
- [46] Kohatsu-Higa, A. and Takeuchi, A.: Jump SDEs and the study of their densities: A self-study book. Universitext. Springer, Singapore, 2019. xix+355 pp. MR-3967712
- [47] Kulik, A.: Exponential ergodicity of the solutions to SDE's with a jump noise. *Stoch. Process. Appl.* **119**, no. 2, (2009), 602–632. MR-2494006
- [48] Kulik, A.: Ergodic behavior of Markov processes. With applications to limit theorems. De Gruyter Studies in Mathematics 67. De Gruyter, Berlin, 2018. x+256 pp. MR-3791835
- [49] Labbé, C. and Lacoïn, H.: Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling. *Ann. Probab.* **47**, no. 3, (2019), 1541–1586. MR-3945753
- [50] Lachaud, B.: Cut-off and hitting times of a sample of Ornstein-Uhlenbeck process and its average. *J. Appl. Probab.* **42**, no. 4, (2005), 1069–1080. MR-2203823
- [51] Lacoïn, H.: The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.* **44**, no. 5, (2016), 3399–3430. MR-3551201
- [52] Lancia, C., Nardi, F.R. and Scoppola, B.: Entropy-driven cutoff phenomena. *J. Stat. Phys.* **149**, no. 1, (2012), 108–141. MR-2981642
- [53] Levine, D., Peres Y. and Wilmer, E.: Markov chains and mixing times. With a chapter by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI, 2009. xviii+371 pp. MR-2466937
- [54] Martínez, S. and Ycart, B.: Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space. *Adv. in Appl. Probab.* **33**, no. 1, (2001), 188–205. MR-1825322
- [55] Méliot, P. L.: The cut-off phenomenon for Brownian motions on compact symmetric spaces. *Potential Anal.* **40**, no. 4, (2014), 427–509. MR-3201989

- [56] Meyn, S. P. and Tweedie, R. L.: Stability of markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25**, no. 3, (1993), 518–548. MR-1234295
- [57] Mikami, T.: Asymptotic expansions of the invariant density of a Markov process with a small parameter. *Ann. Inst. H. Poincaré Sect. B* **24**, no. 3, (1988), 403–424. MR-0971101
- [58] Nersesyan, V. and Raquépas, R.: Exponential mixing under controllability conditions for SDEs driven by a degenerate Poisson noise. *Stochastic. Process. Appl.* **138**, (2021), 26–55. MR-4252192
- [59] Pachpatte, B.: Inequalities for differential and integral equations. Mathematics in Science and Engineering 197. Academic Press, Inc., San Diego, CA, 1998. x+611 pp. MR-1487077
- [60] Peng, X. and Zhang, R.: Exponential ergodicity for SDEs under the total variation. *J. Evol. Equ.* **18**, no. 3, (2018), 1051–1067. MR-3859440
- [61] Protter, P.: Stochastic integration and differential equations. A new approach. Applications of Mathematics (New York) 21. Springer-Verlag, Berlin, 1990. x+302 pp. MR-1037262
- [62] Rosiński, J.: Tempering stable processes. *Stochastic Process. Appl.* **177**, no. 6, (2007), 677–707. MR-2327834
- [63] Saint Loubert Bié, E.: Étude d’une EDPS conduite par un bruit poissonnien. *Probab. Theory Related Fields* **111**, no. 2, (1998), 287–321. MR-1633586
- [64] Saloff-Coste, L.: Random walks on finite groups. (English summary) Probability on discrete structures, 263–346, Encyclopaedia Math. Sci. **110**, Probab. Theory 1, Springer, Berlin, 2004. MR-2023654
- [65] Sato, K.: Lévy processes and infinitely divisible distributions. Translated from the 1990 Japanese original. Revised edition of the 1999 English translation. Cambridge Studies in Advanced Mathematics 68. Cambridge University Press, Cambridge, 2013. xiv+521 pp. MR-3185174
- [66] Sato, K. and Yamazato, M.: Operator-self-decomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Process. Appl.* **17**, no. 1, (1984), 73–100. MR-0738769
- [67] Picard, J.: On the existence of smooth densities for jump processes. *Probab. Theory Related Fields* **105**, no. 4, (1996), 481–511. MR-1402654
- [68] Reiss, R.: Approximate distributions of order statistics. With applications to nonparametric statistics. Springer Series in Statistics. Springer-Verlag, New York, 1989. xii+355 pp. MR-0988164
- [69] Sheu, S. J.: Asymptotic behavior of the invariant density of a diffusion Markov process with small diffusion. *SIAM J. Math. Anal.*, **17**, no. 2, (1986), 451–460. MR-0826705
- [70] Siegert, W.: Local Lyapunov exponents. Sublimiting growth rates of linear random differential equations. Lecture Notes in Mathematics 1963. Springer-Verlag, Berlin, 2009. x+254 pp. MR-2465137
- [71] Simon, T.: On the absolute continuity of multidimensional Ornstein-Uhlenbeck processes. *Probab. Theory Related Fields* **151**, no. 1-2, (2011), 173–190. MR-2834716
- [72] Siorpaes, P.: Applications of pathwise Burkholder-Davis-Gundy inequalities. *Bernoulli* **24**, no. 4B, (2018), 3222–3245. MR-3788172
- [73] Song, Y. and Zhang, X.: Regularity of density for SDEs driven by degenerate Lévy noises. *Electron. J. Probab.*, **20**, no. 21, (2015), 1–27. MR-3325091
- [74] Trefethen, L. N. and Trefethen, L. M.: How many shuffles to randomize a deck of cards? *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **456**, no. 2002, (2000), 2561–2568. MR-1796496
- [75] Tudoran, R. M.: On the coercivity of continuously differentiable vector fields. *Qual. Theory Dyn. Syst.* **19**, no. 58, (2020), 1–7. MR-4083676
- [76] Wang J.: Exponential ergodicity and strong ergodicity for SDEs driven by symmetric  $\alpha$ -stable processes. *Appl. Math. Lett.* **26**, no. 6, (2013), 654–658. MR-3028071
- [77] Wang, J.: On the existence and explicit estimates for the coupling property of Lévy processes with drift. *J. Theoret. Probab.* **27**, no. 3, (2014), 1021–1044. MR-3245997
- [78] Whitt, W.: Some useful functions for functional limit theorems. *Math. Oper. Res.* **5**, no.1, (1980), 67–85. MR-0561155

- [79] Ycart, B.: Cutoff for samples of Markov chains. *ESAIM Probab. Stat.* **3**, (1999), 89–106. MR-1716128

**Acknowledgments.** The research of GBV has been supported by the Academy of Finland, via the Matter and Materials Prof4 university profiling action. GBV also would like to express his gratitude to University of Helsinki for all the facilities used along the realization of this work. The research of MAH has been supported by the Proyecto de la Convocatoria 2020–2021: “Stochastic dynamics of systems perturbed with small Markovian noise with applications in biophysics, climatology and statistics” of the School of Sciences (Facultad de Ciencias) at Universidad de los Andes. JCP acknowledges support from CONACyT-MEXICO CB-250590. The authors would like to thank professor M. Jara and professor R. Imbuzeiro Oliveira both at IMPA for ideas how to construct the example given in Subsubsection 1.3.5.