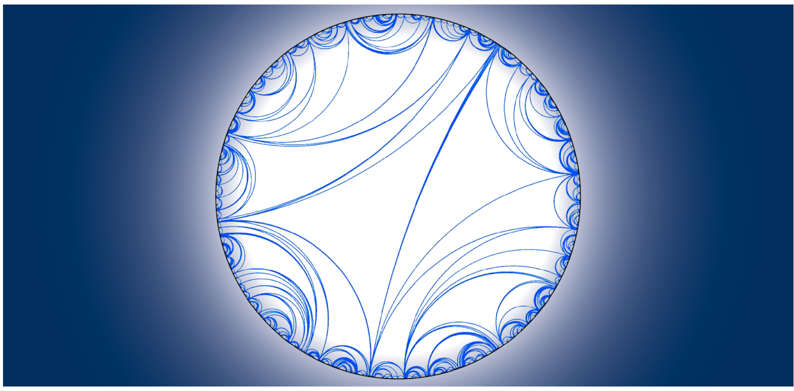
Universität Potsdam



Jan Gairing | Michael Högele | Tetiana Kosenkova | Alexei Kulik

Coupling Distances between Lévy Measures and Applications to Noise Sensitivity of SDE

Preprints des Instituts für Mathematik der Universität Potsdam
2 (2013) 16

Preprints des Instituts für Mathematik der Universität Potsdam

Preprints des Instituts für Mathematik der Universität Potsdam
2 (2013) 16

Jan Gairing | Michael Högele | Tetiana Kosenkova | Alexei Kulik

Coupling Distances between Lévy Measures and Applications to Noise Sensitivity of SDE

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.de abrufbar.

Universitätsverlag Potsdam 2013
http://verlag.ub.uni-potsdam.de/

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: 2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Preprints des Instituts für Mathematik der Universität Potsdam** wird herausgegeben vom Institut für Mathematik der Universität Potsdam.

ISSN (online) 2193-6943

Kontakt:
Institut für Mathematik
Am Neuen Palais 10
14469 Potsdam
Tel.: +49 (0)331 977 1028
WWW: http://www.math.uni-potsdam.de

Titelabbildungen:
1. Karla Fritze | Institutsgebäude auf dem Campus Neues Palais
2. Nicolas Curien, Wendelin Werner | Random hyperbolic triangulation
Published at: http://arxiv.org/abs/1105.5089

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2013/6888/
URN urn:nbn:de:kobv:517-opus-68886
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-68886

Coupling distances between Lévy measures and applications to noise sensitivity of SDE

Jan Gairing * Michael Högele† Tetiana Kosenkova ‡ Alexei Kulik §

November 15, 2013

Abstract

We introduce the notion of coupling distances on the space of Lévy measures in order to quantify rates of convergence towards a limiting Lévy jump diffusion in terms of its characteristic triplet, in particular in terms of the tail of the Lévy measure. The main result yields an estimate of the Wasserstein-Kantorovich-Rubinstein distance on path space between two Lévy diffusions in terms of the couping distances. We want to apply this to obtain precise rates of convergence for Markov chain approximations and a statistical goodness-of-fit test for low-dimensional conceptual climate models with paleoclimatic data.

Keywords: Lévy diffusion approximation; coupling methods; Wasserstein-Kantorovich-Rubinstein metric; Skorohod's invariance principle; statistical model selection.

2010 Mathematical Subject Classification:
60J60; 60J75; 60F17; 60G51; 60H10; 62G32; 62P12.

1 Introduction

This article introduces a family of distances on the set of Lévy measures on $\mathbb{R}^d$, which we shall call *coupling distances* since they are based on the coupling-type Wasserstein-Kantorovich-Rubinstein distance between probability laws. They measure the distance between the appropriately truncated and normalized tails of Lévy measures. Their construction aims at quantifying the rate of convergence in limit theorems and approximation schemes of Lévy driven jump processes in terms of the underlying Lévy measures. In particular we are interested in their distribution on path space.

Let us outline briefly two areas where this notion turns out to be useful.

Recall Gnedenko's theorem about the weak convergence of the row-wise sums $S_n = \sum_{k=1}^n \xi_{kn}$ of triangular arrays $(\xi_{kn})_{k=1,\dots,n}$ of independent and row-wise identically distributed random variables. See for instance Chapter 19, Theorem 2 in [19]. The main condition is the convergence in a proper sense of the family of Lévy measures $\Pi_n(du) = n\mathbf{P}(\xi_{1n} \in du)$ to the Lévy measure Π of a limiting infinitely divisible distribution. The very same condition appears in theorems of Skorokhod invariance principle-type on the convergence of step-wise processes to a Lévy process in $\mathbb{D}$, which are also constructed via partial row-wise sums [8]. These classical results allow for generalizations

*Institut für Mathematik, Humboldt-Universität zu Berlin, Germany; gairing@math.hu-berlin.de
†Institut für Mathematik, Universität Potsdam, Germany; hoegele@uni-potsdam.de
‡Kyiv National Taras Shevchenko University, Ukraine; tanya.kosenkova@gmail.com
§Institute of Mathematics, National Academy of Science, Kiev, Ukraine; kulik@imath.kiev.ua

to the case of step-wise processes constructed via Markov chains, as treated for instance in [22] and [23].

In all these results weak convergence is proved, but the rate of convergence is not addressed. A first step in this direction is to quantify the convergence of Lévy measures Π_n to Π in terms of an appropriate metric. This step is made in the current paper. A functional limit theorem with an explicit bound for the rate of convergence of step-wise processes for the Markov chain approximation is subject to a separate paper [24].

A second question we would like to discuss concerns the sensitivity of the solution to a one-dimensional Lévy driven stochastic differential equation (SDE) with respect to perturbations of the noise. Consider a sufficiently regular function $V : \mathbb{R} \to \mathbb{R}$, two Lévy processes $(Z_j(t))_{t\geqslant 0}$, $j = 1, 2$ and the SDEs

$$X_j(t) = x + \int_0^t V\left(X_j(s)\right) ds + Z_j(t) \qquad t \geqslant 0,\ x \in \mathbb{R}. \tag{1}$$

If the characteristic triplets (see Section 2 below) of the two Lévy processes $Z_j, j = 1, 2$ are close in a sense, it is natural to expect that the laws of the respective solutions $X_j, j = 1, 2$ should not deviate too much. To make such a statement precise, the proposed metric turns out to be particularly valuable. Such problems have a strong motivation coming from statistical inference.

Let us sketch the problem of model selection for low-dimensional climate models. Such systems can be derived as zero dimensional energy balance models perturbed by random fluctuations and have been studied extensively, see for instance [1], [2], [4], [12], [14] and [15]. These studies focus on the transition behavior, metastability and stochastic resonance phenomena for Gaussian perturbations. The study of climate dynamics suffers from a poor quantity of available (proxy) data covering climate intrinsic time scales. One of the richest (and best-studied) time series stems from annual temperature proxies taken from ice cores of the Greenland ice sheet dating back to 80.000 years before present [21]. In [5] Ditlevsen linked the fast non-Gaussian transition behavior exhibited in the data to the presence of jumps. The analysis of the relation between the frequency and the size of the large jumps justifies the assumption of a bistable model given by equations of type (1). This gave rise to further studies of a great variety of such models [10], [13], [16], [17] and [18]. In [8] and [9] the authors solve the corresponding model selection problem within the class of α-stable diffusions based on a fine analysis of sample path properties.

The deviation bounds obtained in Section 2 of the current article provide an quantitative instrument to asses model selection in the class of general Lévy diffusions with additive noise. The coupling (semi-) distances proposed here are weak enough in order to be statistically tractable and allow for an empirical evaluation of a large class of standard diffusion models. In a separate paper [7], we will carry out this procedure in an empirical analysis of the aforementioned time series. On the other hand they are sufficiently strong in a toplogical sense in order to measure convergence rates in functional limit theorems [24].

The present article is organized as follows. We first develop the concept of coupling distances of Lévy measures and address basic properties. Then we introduce set-up under consideration and state the main results in Theorem 1 and 2. Section 3 is devoted to the example of α-stable Lévy measures. The article closes with the proof of the theorems in Section 4.

2 Main Results

Recall that a *Lévy measure* on $\mathbb{R}^d$ is a σ-finite measure Π on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} (|u|^2 \wedge 1)\Pi(du) < \infty$$

and denote by $\mathfrak{L}$ the space of such Lévy measures. We do not exclude point masses in $\{0\}$ here. For a given $\Pi \in \mathfrak{L}$ and given $r > 0$ we want a decomposition of Π into two σ-finite measures $\Pi^{H,r}$ and $\Pi^{T,r}$ of the form

$$\Pi = \Pi^{H,r} + \Pi^{T,r}, \tag{2}$$

such that

- the total mass of $\Pi^{T,r}$ is r and
- there exists $\varepsilon(r) \geqslant 0$ for which

$$\operatorname{supp}(\Pi^{H,r}) = \{u : |u| \leqslant \varepsilon(r)\} \qquad \text{and} \qquad \operatorname{supp}(\Pi^{T,r}) = \{u : |u| \geqslant \varepsilon(r)\} \cup \{0\}.$$

Here "H" and "T" stand for "head" and "tail" of the measure, respectively. Such a decomposition always exists if Π has infinite intensity ($\Pi(\mathbb{R}^d) = \infty$) and is unique if Π is continuous. Let us assume for a while that we have such a unique decomposition for all $r > 0$; a work-around for the cases omitted here will be given below. For $r > 0$ we define a probability measure

$$\pi^r = \frac{1}{r}\Pi^{T,r}. \tag{3}$$

Recall that on a metric space (S, d) the *Wasserstein-Kantorovich-Rubinstein metric* of order 2, between two probability measures μ, ν on (S, d) is defined by

$$W_{2,d}(\mu, \nu) := \inf_{(\xi,\eta)\in\mathcal{C}(\mu,\nu)} \left(\mathbf{E}\, d^2(\xi,\ \eta)\right)^{1/2},$$

where $\mathcal{C}(\mu, \nu)$ denotes the set of all (μ, ν)-couplings. The space $\mathcal{C}(\mu, \nu)$ consists of all pairs (ξ, η) of S-valued random elements, defined on the same probability space, such that $\mathrm{Law}(\xi) = \mu$ and $\mathrm{Law}(\eta) = \nu$. For further details we refer for instance to [6], Chapter 11, or [26]. In what follows, we will consider the metric ρ on $\mathbb{R}^d$ defined by

$$\rho(x, y) = |x - y| \wedge 1, \qquad x, y \in \mathbb{R}^d.$$

The following notion will allow to transfer the concept of optimal couplings from probability distributions to the space of Lévy measures $\mathfrak{L}$.

Definition 1. For Lévy measures $\Pi_1, \Pi_2 \in \mathfrak{L}$ define

$$T_r(\Pi_1, \Pi_2) := r^{1/2} W_{2,\rho}(\pi_1^r, \pi_2^r), \quad r > 0, \tag{4}$$

$$T(\Pi_1, \Pi_2) := \sup_{r>0} T_r(\Pi_1, \Pi_2). \tag{5}$$

We shall call T_r and T *coupling distances* on the space $\mathfrak{L}$.

The (semi-) distances T_r capture a compound Poisson approximation by coupling the jumps in an optimal way, which we will make precise in Section 4.

Let us address the decomposition (2) for general Lévy measures.

- If $\Pi(\mathbb{R}^d) \leqslant r$, then clearly $\Pi^{H,r} := \mathbf{0}$ (the zero measure). In order to produce the mass required in (5) we formally define

$$\Pi^{T,r} := \Pi + (r - \Pi(\mathbb{R}^d))\delta_0 \tag{6}$$

artifically introducing a point mass in 0.

- In general, $\varepsilon(r)$ is defined as the unique $\varepsilon > 0$ such that for some (unique) $p \in [0, 1)$

$$\Pi(|u| > \varepsilon) + p\Pi(|u| = \varepsilon) = r.$$

In this case we set

$$\Pi^{T,r}(du) := \Big(1_{\{|u|>\varepsilon\}} + p1_{\{|u|=\varepsilon\}}\Big)\Pi(du), \qquad \Pi^{H,r} := \Pi - \Pi^{T,r}. \tag{7}$$

Remark 1. Note that the choice (7) is not a unique, because instead of the "symmetric" additional term $p1_{\{|u|=\varepsilon\}}\Pi(du)$ therein one could take an "asymmetric" one of a form $g(u)1_{\{|u|=\varepsilon\}}\Pi(du)$ with any function g such that

$$\int_{\{|u|=\varepsilon\}} g(u)\Pi(du) = p.$$

To specify uniquely the construction, we define the coupling distances T_r and T by (4) and (5) with π^r defined by (3) and the convention (7).

The following statement, proved in Section 3 below, gives the basic properties of the coupling distances T_r and T.

Proposition 1. *1. For every $r > 0$ the function T_r defines a semimetric on $\mathfrak{L}$, that is, it is nonnegative, symmetric, and satisfies the triangle inequality.*

2. For any $\Pi_1, \Pi_2 \in \mathfrak{L}$

$$T(\Pi_1, \Pi_2) \leqslant \Big(\int_{\mathbb{R}^d} (|u|^2 \wedge 1)\Pi_1(du)\Big)^{1/2} + \Big(\int_{\mathbb{R}^d} (|u|^2 \wedge 1)\Pi_2(du)\Big)^{1/2} < \infty. \tag{8}$$

The function T is a metric on $\mathfrak{L}$.

In Section 3 below we give two examples for the calculation of the coupling distance between two α-stable Lévy measures $\Pi_j, j = 1, 2$ with the respective "shape" parameter α_j and the "scaling" parameter c_j. These calculations illustrate the topology of these measures in terms of their parameters.

As a first application let us return to the analysis of one-dimensional SDE (1). We shall quantify the sensitivity of solutions with respect to perturbations with respect to the driving Lévy noise in terms of T. In what follows, we assume that the drift V satisfies the following *weak monotonicity* or *one-sided Lipschitz* condition:

$$(V(x) - V(y))(x - y) \leqslant L(x - y)^2, \quad x, y \in \mathbb{R}, \tag{9}$$

then it is well known that unique (strong) solution to (1) can be obtained via usual Picard - type successive approximation procedure.

Example 1. Generic examples satisfying (9) are polynomials of odd order and negative leading coefficient

$$V(x) = \beta_n x^n + \sum_{i=0}^{n-1} \beta_i x^i, \qquad \beta_n < 0, \beta_i \in \mathbb{R}, n \text{ odd}.$$

In applications such a polynomial is often considered as the gradient of an "energy potential" $U : \mathbb{R} \to \mathbb{R}$ with several local minima such that $V = -U'$.

- The FitzHugh-Nagumo model in neuroscience is an example of such a fourth order potential. Equation (1) models in this case the membrane voltage under random excitation, recently studied in the literature e.g. [3], [11], [27].
- Another class of examples can be derived from energy balance models in climatology, see [2], [5]. Here V describes the interaction between an idealized black body radiation of the Earth and albedo feedback. Under the proper choice of the parameter the potential U admits two local minima that correspond to two climate equilibrium states.

Now, let us consider two Lévy processes Z_j, $j = 1, 2$ with the characteristic triplets (a_j, b_j, Π_j). Recall that the law of Z_j is uniquely determined by its *cumulant function* or *Lévy exponent* ψ_j given as

$$\mathbf{E}e^{izZ_j(t)} = e^{t\psi_j(z)}, \quad z \in \mathbb{R}, \quad t \geqslant 0,$$

linked to the characteristic triplet via the Lévy-Khinchin formula

$$\psi_j(z) = ia_j z - \frac{1}{2}b_j z^2 + \int_{\mathbb{R}} \left[e^{izu} - 1 - iz\tau(u)\right] \Pi_j(du), \quad z \in \mathbb{R}, \tag{10}$$

where $a_j \in \mathbb{R}, b_j \geqslant 0, \Pi_j$ is a Lévy measure and

$$\tau(u) = (|u| \wedge 1)\,\mathrm{sign}\,(u), \qquad u \in \mathbb{R}.$$

This choice of the cutoff function τ has some technical advantages. We denote by X_j, $j = 1, 2$, the solutions to equation (1), driven by Z_j and initial values x_j. $\mathbb{D}(0, 1)$ denotes the space of càdlàg paths and is the state space of our solutions. Introduce the following metric ζ on $\mathbb{D}(0, 1)$ by

$$\zeta(x, y) = \sup_{t\in[0,1]} \rho(x(t), y(t)) = \|x - y\|_\infty \wedge 1, \quad \|x - y\|_\infty = \sup_{t\in[0,1]} |x(t) - y(t)|.$$

The main results of this article estimate the deviation between the laws of the solutions X_j on $(\mathbb{D}(0, 1), \zeta)$ in terms of the metric induced by T and T_r on the set of quadruplets of parameters (x_j, a_j, b_j, Π_j).

Theorem 1. *Let $V : \mathbb{R} \to \mathbb{R}$ be $\mathcal{C}^2$ and satisfy condition (9) for some constant $L > 0$. Let (a_j, b_j, Π_j) be two Lévy characteristics and $x_j \in \mathbb{R}$ given initial values, $j = 1, 2$. Then for any two solutions X_j of equation (1) driven by Lévy processes Z_j with the respective characteristics and for any $r > 0$ the following estimate holds true*

$$W^2_{2,\zeta}\Big(\mathrm{Law}(X_1), \mathrm{Law}(X_2)\Big) \leqslant Q^1_r e^{L/\arctan(1/2)} + Q^2_r$$

where

$$Q^1_r = 2\rho^2(x_1, x_2) + \frac{4}{\pi}\Big(\frac{3^{3/4}}{2}|a_1 - a_2| + (\sqrt{b_1} - \sqrt{b_2})^2 + U_r\,(\Pi_1) + U_r\,(\Pi_2) +$$

$$+ (\pi + 3^{3/4})T^2_r(\Pi_1, \Pi_2) + 3^{3/4}\min(\Pi_1(|u| > 1) + \Pi_2(|u| > 1), r)^{1/2}T_r(\Pi_1, \Pi_2)\Big),$$

$$Q^2_r = \frac{4}{\pi}\sqrt{3^{3/2}(\sqrt{b_1} - \sqrt{b_2})^2 + (2\pi)^2(U_r\,(\Pi_1) + U_r\,(\Pi_2) + T^2_r(\Pi_1, \Pi_2))},$$

and

$$U_r\,(\Pi_j) = \int_{|u|\leqslant\varepsilon_j(r)} u^2\Pi_j(du), \qquad j = 1, 2.$$

The proof of Theorem 1 will be postponed to Section 4. When $r \to \infty$, we have $\varepsilon_j(r) \to 0$ and $U_r(\Pi_j) \to 0$, $j = 1, 2$. By construction of the metric T we obtain the following more theoretical result after polishing the constants.

Theorem 2. *Under the assumptions of Theorem 1 there are constants $c_1, c_2 > 0$ such that*

$$W_{2,\zeta}^2\Big(\mathrm{Law}(X_1), \mathrm{Law}(X_2)\Big) \leqslant c_1\, Q^1 e^{L/\arctan(1/2)} + c_2\, Q^2, \tag{11}$$

where

$$\begin{aligned} Q^1 &= \rho^2(x_1, x_2) + |a_1 - a_2| + (\sqrt{b_1} - \sqrt{b_2})^2 + T^2(\Pi_1, \Pi_2) \\ &\quad + \big(\Pi_1(|u| > 1) + \Pi_2(|u| > 1)\big)^{1/2} T(\Pi_1, \Pi_2), \\ Q^2 &= \sqrt{(\sqrt{b_1} - \sqrt{b_2})^2 + T^2(\Pi_1, \Pi_2)}. \end{aligned}$$

Each of the theorems exists in its own right, since they target different ranges of applications. Clearly, Theorem 2 gives a shorter and a more elegant deviation bound in terms of the metric induced by the coupling distance T on the set of Lévy characteristics. The bound of Theorem 1 however is sharper but stated only in terms of the semi-distance T_r, $r > 0$. From the point of view of practical applications, for example the statistical inference scheme mentioned in the introduction, it has the advantage that its quantities are easier to handle. Obviously the right-hand side does not contain a supremum in r, which in particular can be seen as a free parameter defining the threshold $\varepsilon_j(r)$ between large and small jumps. We refer to [7] for further details.

To illustrate the notion of coupling distances for Lévy measures of infinite intensity we calculate upper bounds for this quantity for α-stable Lévy measures.

3 Two examples: the value of the coupling distance between one-sided α-stable measures

Recall that for a *one-sided α-stable process*, its Lévy measure has the form

$$\Pi(du) = \alpha c u^{-\alpha-1} \mathbf{1}(u)_{[0,\infty)}\, du,$$

where $\alpha \in (0, 2)$ and $c \geqslant 0$. Here we introduce the factor α for further convenience in the calculation. Parameters α and c are naturally interpreted as the "shape" and the "scaling" parameters, and one can expect heuristically that two one-sided α-stable measures are "close", if their respective parameters are close. As we will see in two examples below, the coupling distance T quantifies this convergence.

Example 2. Let $\Pi_j, j = 1, 2$ be two α-stable Lévy measures with the same shape parameter α and different scale parameters $c_1 \neq c_2$. We will show that in this case

$$T^2(\Pi_1, \Pi_2) \leqslant \left(\frac{2}{2-\alpha}\right) \left| c_1^{1/\alpha} - c_2^{1/\alpha} \right|^{\alpha}, \tag{12}$$

which tends to 0 for converging scale parameters. By the explicit formula

$$\int_\varepsilon^\infty \alpha c u^{-\alpha-1}\, du = c\varepsilon^{-\alpha},$$

valid for a one-sided α-stable Lévy measure, we have that

$$\varepsilon_j(r) = \left(\frac{r}{c_j}\right)^{-1/\alpha}, \quad j = 1, 2. \tag{13}$$

Using this, we obtain

$$\pi_j^r((-\infty, x]) = 1 - \frac{c_j}{r x^\alpha}, \quad x \geqslant \left(\frac{r}{c_j}\right)^{-1/\alpha} \quad j = 1, 2. \tag{14}$$

Recall that a *quantile function* of the distribution function F is defined as $F^{[-1]}(y) = \inf\{x : F(x) > y\}$, and that for a uniformly distributed random variable U the random variable $F^{[-1]}(U)$ has distribution function F. Denote by $F_{r,j}^{[-1]},\ j = 1, 2$ the quantile functions of $\pi_j^r,\ j = 1, 2$. Then

$$\left(F_{r,1}^{[-1]}(U), F_{r,2}^{[-1]}(U)\right) \in \mathcal{C}(\pi_1^r, \pi_2^r).$$

Due to the optimal coupling property of the Wasserstein-Kantorovich-Rubinstein distance we obtain

$$W_{2,\rho}^2(\pi_1^r, \pi_2^r) \leqslant \mathbf{E}\rho^2\left(F_{r,1}^{[-1]}(U), F_{r,2}^{[-1]}(U)\right) = \int_0^1 \left(\left|F_{r,1}^{[-1]}(y) - F_{r,2}^{[-1]}(y)\right|^2 \wedge 1\right) dy. \tag{15}$$

Using (14) we get the quantile function for $\pi_j^r,\ j = 1, 2$:

$$F_{r,j}^{[-1]}(y) = \left(\frac{r(1-y)}{c_j}\right)^{-1/\alpha}, \quad 0 \leqslant y \leqslant 1, \quad j = 1, 2. \tag{16}$$

Consequently,

$$\int_0^1 \left(\left|F_{r,1}^{[-1]}(y) - F_{r,2}^{[-1]}(y)\right|^2 \wedge 1\right) dy = \int_0^1 \left(\left[|r(1-y)|^{-2/\alpha}\left|c_1^{1/\alpha} - c_2^{1/\alpha}\right|^2\right] \wedge 1\right) dy.$$

To shorten the notation let us denote $\Delta_c = \left|c_1^{1/\alpha} - c_2^{1/\alpha}\right|$. By a change of variables $z = r(1-y)$, we get eventually

$$T^2(\Pi_1, \Pi_2) \leqslant \int_0^\infty \left[|z|^{-2/\alpha}\Delta_c^2\right] \wedge 1\ dz = \Delta_c^\alpha + \int_0^\infty |z|^{-2/\alpha}\Delta_c^2\ dz = \Delta_c^\alpha + \frac{\alpha}{2-\alpha}\Delta_c^\alpha,$$

as claimed.

Example 3. Let $\Pi_j, j = 1, 2$ be two one-sided Lévy measures with the same scale parameter c, but different shape parameters $0 < \alpha_1 < \alpha_2 < 2$, say. We will show that in this case for $c^* = \frac{1}{2}(1 + \sqrt{5})$

$$\begin{aligned} T^2(\Pi_1, \Pi_2) \leqslant &\left(2\frac{(\alpha_2-\alpha_1)^2\left(c^{*2} - 2\ln(2)\ln\left(\frac{\alpha_2-\alpha_1}{\alpha_1}\right) \wedge 0\right)}{(2-\alpha_1)(2-\alpha_2)(\alpha_1+\alpha_2-\alpha_1\alpha_2)} + \frac{2}{2-\alpha_1}\right) \\ &\times \left(c^* + \ln(\alpha_1) - \ln(\alpha_2 - \alpha_1)\right)^{-\alpha_2}, \end{aligned} \tag{17}$$

which tends to 0 if $\alpha_1 \nearrow \alpha_2$, strictly away from 0 and 2. Like in the previous example, we reduce the problem to the estimation of

$$I_r = \int_0^1 \left(\left|F_{r,1}^{[-1]}(y) - F_{r,2}^{[-1]}(y)\right|^2 \wedge 1\right) dy, \quad r > 0,$$

where

$$F_{r,j}^{[-1]}(y) = \left(\frac{r(1-y)}{c}\right)^{-1/\alpha_j}, \quad 0 \leqslant y \leqslant 1, \quad j = 1, 2. \tag{18}$$

Changing the variables $t = r(1-y)/c$, we get

$$I_r = \frac{c}{r} \int_0^{r/c} \left(\left(t^{-1/\alpha_1} - t^{-1/\alpha_2}\right)^2 \wedge 1\right) dt, \tag{19}$$

hence

$$T^2(\Pi_1, \Pi_2) \leqslant c \int_0^{\infty} \left(\left(t^{-1/\alpha_1} - t^{-1/\alpha_2}\right)^2 \wedge 1\right) dt. \tag{20}$$

On $(0,1)$, the function $\left(t^{-1/\alpha_1} - t^{-1/\alpha_2}\right)^2$ is decreasing from $+\infty$ to 0. On $(1,\infty)$, this function is bounded by 1. Consequently there exists unique $t_* \in (0,1)$ such that

$$t_*^{-1/\alpha_1} - t_*^{-1/\alpha_2} = 1, \tag{21}$$

$$\left(t^{-1/\alpha_1} - t^{-1/\alpha_2}\right)^2 \wedge 1 = \begin{cases} 1, & t \leqslant t_*, \\ \left(t^{-1/\alpha_1} - t^{-1/\alpha_2}\right)^2, & t > t_*. \end{cases}$$

The explicit calculation gives

$$\begin{aligned} \int_0^{\infty} \left(\left(t^{-1/\alpha_1} - t^{-1/\alpha_2}\right)^2 \wedge 1\right) dt &= t_* + \int_{t_*}^{\infty} \left(t^{-1/\alpha_1} - t^{-1/\alpha_2}\right)^2 dt \\ &= t_* \left[\frac{2(\alpha_2-\alpha_1)^2}{(2-\alpha_1)(2-\alpha_2)(\alpha_1+\alpha_2-\alpha_1\alpha_2)} t_*^{-2/\alpha_2} - \frac{2\alpha_1(\alpha_2-\alpha_1)}{(2-\alpha_1)(\alpha_1+\alpha_2-\alpha_1\alpha_2)} t_*^{-1/\alpha_2} + \frac{2}{2-\alpha_1}\right] \\ &\leqslant \frac{2(\alpha_2-\alpha_1)^2}{(2-\alpha_1)(2-\alpha_2)(\alpha_1+\alpha_2-\alpha_1\alpha_2)} t_*^{(\alpha_2-2)/\alpha_2} + \frac{2}{2-\alpha_1} t_*. \end{aligned} \tag{22}$$

Denote

$$y = t_*^{-1/\alpha_2}. \tag{23}$$

Then (21) can be written in the following form:

$$y^{\alpha_2/\alpha_1} - y = 1. \tag{24}$$

To simplify the notation we denote

$$\beta = \frac{\alpha_2 - \alpha_1}{\alpha_1}. \tag{25}$$

Then for (24) we get

$$y = \left(1 + \frac{1}{y}\right)^{1/\beta}. \tag{26}$$

If we differentiate equation (26) with respect to β and obtain

$$\frac{dy}{d\beta} = \frac{d}{d\beta}\left(1 + \frac{1}{y}\right)^{1/\beta} = -\left(1 + \frac{1}{y}\right)^{1/\beta} \left(\frac{1}{\beta^2} \ln(1 + \frac{1}{y}) + \frac{1}{\beta} \frac{y}{1+y} \frac{1}{y^2} \frac{dy}{d\beta}\right) \tag{27}$$

We resolve the equation for $\frac{dy}{d\beta}$ to obtain

$$\frac{dy}{d\beta} = -\left(1 + \frac{(1+1/y)^{1/\beta-1}}{y^2\beta}\right)^{-1} \frac{\ln(1+1/y)}{\beta^2} (1+1/y)^{1/\beta}$$
$$= -\left((1+1/y)^{-1/\beta} + \frac{(1+1/y)^{-1}}{y^2\beta}\right)^{-1} \frac{\ln(1+1/y)}{\beta^2} \ .$$

Using $\ln(2)x \leqslant \ln(1+x)$ for $0 \leqslant x \leqslant 1$ we get

$$\frac{dy}{d\beta} \leqslant -\frac{\ln(2)}{\beta y} \frac{\frac{1}{\beta}(1+1/y)}{1+\frac{1}{y^2\beta}}^{(1/\beta\wedge 1)} \ .$$

The differential of y is strictly negative. Hence y is decreasing in β, in particular $y(\beta) \leqslant y(1) = \frac{1}{2}(1+\sqrt{5})$, for $\beta \geqslant 1$. On the set $0 < \beta \leqslant 1$

$$\frac{dy}{d\beta} \leqslant -\frac{\ln(2)}{y\beta} \frac{\frac{1}{\beta}(1+1/y)}{1+\frac{1}{y^2\beta}} \leqslant -\frac{\ln(2)}{y\beta} \frac{\frac{1}{\beta}+\frac{1}{\beta y^2}}{1+\frac{1}{y^2\beta}} \leqslant -\frac{\ln(2)}{y\beta} \ .$$

With $c^* = y(1)$ we obtain the estimate

$$y(\beta) \leqslant \sqrt{c^{*2} - 2\ln(2)\ln(\beta)} \ , \qquad 0 < \beta \leqslant 1 \ .$$

In the light of convention (25) we remark that although the right hand side explodes as $\beta \searrow 0$ the terms of order $\beta y(\beta)$ that appear in equation (17) are bounded and converge to zero. In fact we have

$$(\alpha_2 - \alpha_1)^2 t_*^{-2/\alpha_2} \leqslant (\alpha_2-\alpha_1)^2 \left(c^{*2} - 2\ln(2)\ln\big(\frac{\alpha_2-\alpha_1}{\alpha_1}\big) \wedge 0\right) \ . \tag{28}$$

Now bearing in mind that $1/y \in (0,1)$ we estimate equation (27) from below

$$\frac{dy}{d\beta} \geqslant -\left(1+\frac{1}{y}\right)^{1/\beta} \frac{1}{y}\Big(\frac{1}{\beta^2} + \frac{1}{\beta}\frac{dy}{d\beta}\Big) \geqslant -\frac{1}{\beta} \ ,$$

we obtain $y \geqslant c^* - \ln(\beta)$ such that

$$t_* \leqslant \Big(c^* - \ln\big(\frac{\alpha_2-\alpha_1}{\alpha_1}\big)\Big)^{-\alpha_2}, \tag{29}$$

which tends to 0 as $\alpha_1 \nearrow \alpha_2$. Combining (28) and (29) we obtain the bound in (17) and this completes the proof.

4 Proofs

4.1 Proof of Proposition 1

Statement I follows immediately from the fact that $W_{2,\rho}$ is a metric and the following estimate. Let $\Pi \in \mathfrak{L}$ and the zero measure $\mathbf{0} \in \mathfrak{L}$. For every $r > 0$ the element $\mathbf{0}^r$ is the delta-measure δ_0 and for every $(\xi, \eta) \in \mathcal{C}(\pi^r, \mathbf{0}^r)$ one has $\eta = 0$ a.s. Therefore

$$\mathbf{E}\rho^2(\xi, 0) = \int_{\mathbb{R}^d} \rho^2(u,0)\pi^r(du) = \frac{1}{r}\int_{\mathbb{R}^d} (u^2 \wedge 1)\Pi^{T,r}(du).$$

Hence

$$T_r^2(\Pi,\mathbf{0}) = \int_{\mathbb{R}^d} (u^2\wedge 1)\Pi^{T,r}(du) \leqslant \int_{\mathbb{R}^d} (u^2\wedge 1)\Pi(du) < \infty. \tag{30}$$

Then by the triangle inequality

$$T_r(\Pi_1,\Pi_2) \leqslant T_r(\Pi_1,\mathbf{0}) + T_r(\Pi_2,\mathbf{0}) \leqslant \Big(\int_{\mathbb{R}^d} (u^2\wedge 1)\Pi_1(du)\Big)^{1/2} + \Big(\int_{\mathbb{R}^d} (u^2\wedge 1)\Pi_2(du)\Big)^{1/2}.$$

Since this bound does not depend on r, also T is finite and (8) holds true. Now, since $W_{2,\rho}$ is a true metric, $T(\Pi_1,\Pi_2)=0$ implies that

$$\pi_1^r = \pi_2^r, \quad \text{for all } r>0\ ,$$

which implies that $\Pi_1=\Pi_2$. This means that T is a metric. □

4.2 Proof of Theorem 1

The proof consists of two parts. In the first step ("construction") we construct a coupling of the Lévy processes (Z_1,Z_2) with given characteristic triplets and a certain optimality property. The second step ("estimation") uses Itô's formula to obtain a bound for the difference of the respective solutions of the SDE driven by the components of our coupling.

Let (a_j,b_j,Π_j), $j=1,2$ be two given Lévy triplets. It is well known that a Lévy process Z_j with this characteristic triplet has the *Itô-Lévy representation*

$$Z_j(t) = \Big(a_j - \Pi_j(|u|>1)\Big)t + \sqrt{b_j}W_j(t) + \int_0^t\int_{|u|\leqslant 1} u\,\tilde\nu_j(ds,du) + \int_0^t\int_{|u|>1} u\,\nu_j(ds,du),\quad t\geqslant 0,$$

where W_j is a Wiener process, ν_j is a Poisson point measure on $\mathbb{R}^+\times\mathbb{R}$ with the intensity measure $ds\Pi_j(du)$ and $\tilde\nu_j(ds,du)=\nu_j(ds,du)-ds\Pi_j(du)$ is the respective compensated Poisson point measure. Furthermore, W_j, $\tilde\nu_j\big|_{|u|\leqslant 1}$ and $\nu_j\big|_{|u|>1}$ are independent. Note that the additional term $\Pi_j(u>1)$ which appears in the drift corresponds to our choice of compensating term in (10) equal to $iz\tau(u)$, instead of usual $izu1_{\{|u|\leqslant 1\}}$.

The coupling (Z_1,Z_2) of our choice satisfies the following. The diffusive parts of the components Z_j, $j=1,2$ will be generated by a single Wiener processes W. For given $r>0$ we split the Lévy measures Π_j according to relation (2). The "large jump" part $Z^{T,r}=(Z_1^{T,r},Z_2^{T,r})$ of (Z_1,Z_2) is given as a compound Poisson process

$$Z^{T,r}(t) = \sum_{k=1}^{N^r(t)} \xi_k, \tag{31}$$

where N^r is a Poisson process with the intensity r, and the random vectors $\xi_k=(\xi_{k,1},\xi_{k,2}), k\geqslant 1$ are i.i.d. and independent of N^r. The law of $\xi_{k,j}$ equals $\pi_j^r=(1/r)\Pi_j^{T,r}$, $j=1,2$, and the coupling is chosen in such a way that for each k the joint law κ of $(\xi_{k,1},\xi_{k,2})$ is the optimal coupling in the Wasserstein-Kantorovich-Rubinstein sense [26], i.e.

$$\mathbf{E}[\rho^2(\xi_{k,1},\xi_{k,2})] = \iint \rho^2(u_1,u_2)\kappa(du_1,du_2) = W_{2,\rho}^2(\pi_1^r,\pi_2^r).$$

The jump part $\Delta\hat Z^r$ of the remaining processs $\hat Z^r=(Z_1,Z_2)-Z^{T,r}$ stems from a Lévy measure

$$\hat\Pi(du_1,du_2) = \delta_0(du_1)\Pi_2^{H,r}(du_2) + \Pi_1^{H,r}(du_1)\delta_0(du_2) \tag{32}$$

concentrated on the axes, and is realized by a random Poisson measure

$$\hat{\nu}^r = \hat{\nu}^r_1 \otimes \hat{\nu}^r_2 \tag{33}$$

with two independent components and which are jointly independent of W and $Z^{T,r}$. We introduce the compensated random Poisson measure

$$\tilde{\nu}^r(ds,du) := \hat{\nu}^r(ds,du) - ds\tau(u_1+u_2)\hat{\Pi}^r(du)$$

which is also concentrated on the axes almost surely. Therefore marginal processes $\hat{Z}^r_j, j = 1,2$ have the following shape

$$\hat{Z}^r_j(t) = a_j t + \sqrt{b_j}W(t) + \int_0^t \int_{\mathbb{R}} u_j \tilde{\nu}^r_j(ds,du) - t\int_{\mathbb{R}} \tau(u_j)\Pi^{T,r}_j(du). \tag{34}$$

In the second step ("estimation") of the proof we derive upper bounds for ζ-difference between the solutions $X_j, j = 1,2$ for equations (1) with initial conditions $x_j, j = 1,2$ driven by noise coupling (Z_1, Z_2) constructed above. We consider the difference process $Y(t) = X_1(t) - X_2(t)$ and the smooth and bounded auxiliary function

$$F(y) = \arctan y^2.$$

Due to

$$y \wedge 1 \leqslant \frac{4}{\pi}\arctan y, \qquad \text{for } y \geqslant 0 \tag{35}$$

we obtain

$$\mathbf{E}\sup_{t\in[0,1]} \rho^2(X_1(t), X_2(t)) \leqslant \frac{4}{\pi}\mathbf{E}\sup_{t\in[0,1]} F(Y(t)). \tag{36}$$

The process Y has the Itô differential representation

$$dY(t) = \Big(V(X_1(t)) - V(X_2(t))\Big)\,dt + d\Big(\hat{Z}^r_1(t) - \hat{Z}^r_2(t)\Big) + d\Big(Z^{T,r}_1(t) - \hat{Z}^{T,r}_2(t)\Big), \tag{37}$$

which can be extended further, using the identities for $\hat{Z}^r_j$ and $Z^{T,r}_j$ given above. Recall the definition of $\hat{\nu}^r$ in (33) and denote by $\nu^{T,r}$ the Poisson point measure on $\mathbb{R}^+ \times \mathbb{R}^2$, which is realized by the compound Poisson process $Z^{T,r} = (Z^{T,r}_1, Z^{T,r}_2)$ of equation (31). By construction its Lévy measure equals

$$\Pi^{T,r}(du_1, du_2) = r\kappa(du_1, du_2).$$

For convenience we abbreviate $\sigma_j = \sqrt{b_j}, j = 1,2$. Then

$$\begin{aligned} dY(t) = \Big(V(X_1(t)) - V(X_2(t))\Big)\,dt + (a_1 - a_2)\,dt - \int_{\mathbb{R}^2}\Big(\tau(u_1) - \tau(u_2)\Big)\Pi^{T,r}(du)\,dt \\ + (\sigma_1 - \sigma_2)\,dW(t) + \int_{\mathbb{R}^2}(u_1 - u_2)\Big[\hat{\nu}(dt,du) - dt\hat{\Pi}(du)\Big] + \int_{\mathbb{R}^2}(u_1 - u_2)\nu^{T,r}(dt,du). \end{aligned} \tag{38}$$

Now, we apply Itô's formula for the function $F(y)$ (see [20] Chapter 2)

$$\begin{aligned}
F(Y(t)) - F(Y(0)) = &\int_0^t F'(Y(s)) \Big(V(X_1(s)) - V(X_2(s))\Big) ds \\
&+ (a_1 - a_2) \int_0^t F'(Y(s))\, ds \\
&- \int_0^t \int_{\mathbb{R}^2} F'(Y(s)) \Big(\tau(u_1) - \tau(u_2)\Big) \Pi^{T,r}(du) ds \\
&+ (\sigma_1 - \sigma_2) \int_0^t F'(Y(s))\, dW(s) \\
&+ \frac{1}{2}(\sigma_1 - \sigma_2)^2 \int_0^t F''(Y(s))\, ds \\
&+ \int_0^t \int_{\mathbb{R}^2} \Big[F(Y(s-) + (u_1 - u_2)) - F(Y(s-)\Big] \Big[\hat{\nu}(ds, du) - \hat{\Pi}(du) ds\Big] \\
&+ \int_0^t \int_{\mathbb{R}^2} \Big[F(Y(s-) + (u_1 - u_2)) - F(Y(s-)) - F'(Y(s-))(u_1 - u_2)\Big] \hat{\Pi}(du) ds \\
&+ \int_0^t \int_{\mathbb{R}^2} \Big[F(Y(s-) + (u_1 - u_2)) - F(Y(s-)\Big] \nu^{T,r}(ds, du).
\end{aligned} \tag{39}$$

We separate the martingale parts, which come from line 4 and 6 in (39). Since F is bounded we can also compensate the compound Poisson part in line 8

$$\begin{aligned}
M_t = &(\sigma_1 - \sigma_2) \int_0^t F'(Y(s))\, dW(s) \\
&+ \int_0^t \int_{\mathbb{R}^2} \Big[F(Y(s-) + (u_1 - u_2)) - F(Y(s-)\Big] \Big[\hat{\nu}(ds, du) - ds\hat{\Pi}(du)\Big] \\
&+ \int_0^t \int_{\mathbb{R}^2} \Big[F(Y(s-) + (u_1 - u_2)) - F(Y(s-)\Big] \Big[\nu^{T,r}(ds, du) - ds\Pi^{T,r}(du)\Big].
\end{aligned} \tag{40}$$

Hence we can rewrite (39) as

$$F(Y(t)) = F(Y(0)) + \int_0^t g(X_1(s), X_2(s))\, ds + M_t, \tag{41}$$

where

$$\begin{aligned}
g(z_1, z_2) = &(V(z_1) - V(z_2)) F'(z_1 - z_2) + (a_1 - a_2) F'(z_1 - z_2) + \frac{1}{2}(\sigma_1 - \sigma_2)^2 F''(z_1 - z_2) \\
&+ \int_{\mathbb{R}^2} \Big[F(z_1 - z_2) + (u_1 - u_2)) - F(z_1 - z_2) - F'(z_1 - z_2)(u_1 - u_2)\Big] \hat{\Pi}(du) \\
&+ \int_{\mathbb{R}^2} \Big[F((z_1 - z_2) + (u_1 - u_2)) - F(z_1 - z_2) - F'(z_1 - z_2)\Big(\tau(u_1) - \tau(u_2)\Big)\Big] \Pi^{T,r}(du) \\
&=: g_1(z_1, z_2) + g_2(z_1, z_2) + g_3(z_1, z_2) + g_4(z_1, z_2) + g_5(z_1, z_2).
\end{aligned}$$

For the function F and its derivatives, we have the following explicit expressions and bounds:

$$F'(y) = \frac{2y}{1 + y^4}, \qquad |F'(y)| \leqslant \frac{3^{3/4}}{2} \tag{42}$$

$$F'(y)y = \frac{2y^2}{1+y^4} \leqslant (2y^2) \wedge 1 = 2(y^2 \wedge \frac{1}{2}) \leqslant \frac{F(y)}{\arctan(1/2)}, \tag{43}$$

$$F''(y) = 2\frac{1-3y^4}{(1+y^4)^2}, \qquad |F''(y)| \leqslant 2 \tag{44}$$

$$|F(y+\delta) - F(y) - F'(y)\delta| \leqslant \frac{\delta^2}{2} \sup_v |F''(v)| \leqslant \delta^2. \tag{45}$$

Hence, we can bound every summand on the r.h.s. of (42). By the Lipschitz condition (9) and (43)

$$g_1(z_1,z_2) \leqslant \frac{L}{\arctan(1/2)} F(z_1 - z_2).$$

Estimates (42) and (44) yield

$$g_2(z_1,z_2) \leqslant \frac{3^{3/4}}{2}|a_1 - a_2|, \qquad g_3(z_1,z_2) \leqslant (\sigma_1 - \sigma_2)^2.$$

and (45) and (32),

$$g_4(z_1,z_2) \leqslant \int_{\mathbb{R}^2} (u_1-u_2)^2 \hat{\Pi}(du_1,du_2) = \int_{\mathbb{R}} u^2 \Pi_1^{H,r}(du) + \int_{\mathbb{R}} u^2 \Pi_2^{H,r}(du) = U_r(\Pi_1) + U_r(\Pi_2). \tag{46}$$

To estimate $g_5(z_1,z_2)$, we rewrite it in the following way

$$g_5(z_1,z_2) = \Big(\int_{|u_1-u_2|\leqslant 1} + \int_{|u_1-u_2|>1} \Big) \Big[\cdots \Big] \Pi^{T,r}(du),$$

and note that in any case the absolute value of the term under the $\Big[\cdots \Big]$ does not exceed $\pi + 3^{3/4}$. In the case when $|u_1 - u_2| \leqslant 1$, we have the inequality

$$F((z_1-z_2)+(u_1-u_2)) - F(z_1-z_2) \leqslant F'(z_1-z_2)(u_1-u_2) + (u_1-u_2)^2,$$

which comes from the Taylor expansion because $|F''(z)| \leqslant 2$. Hence, after simple rearrangements, we get

$$\begin{aligned} g_5(z_1,z_2) &\leqslant (\pi + 3^{3/4}) \int_{\mathbb{R}^2} \Big((u_1-u_2)^2 \wedge 1 \Big) \Pi^{T,r}(du) \\ &\quad + F'(z_1-z_2) \int_{|u_1-u_2|\leqslant 1} \Big[(u_1-u_2) - (\tau(u_1) - \tau(u_2)) \Big] \Pi^{T,r}(du) \\ &= (\pi + 3^{3/4}) T_r^2(\Pi_1,\Pi_2) + \frac{3^{3/4}}{2} \int_{|u_1-u_2|\leqslant 1} \Big[(u_1-u_2) - (\tau(u_1)-\tau(u_2)) \Big] \Pi^{T,r}(du). \end{aligned}$$

Here we have used that $\Pi^{T,r} = r\kappa$, where κ is the law on $\mathbb{R}^2$ which minimizes the expectation in the definition of $W_{2,\rho}(\pi_1^r, \pi_2^r)$. Consequently,

$$\int_{\mathbb{R}^2} \Big(|u_1-u_2|^2 \wedge 1 \Big) \Pi^{T,r}(du) = r \int_{\mathbb{R}^2} \rho^2(u_1,u_2)\kappa(du) = r W_\rho^2(\pi_1^r, \pi_2^r) = T_r^2(\Pi_1,\Pi_2). \tag{47}$$

The absolute value of the integrand in the remaining integral can be estimated by

$$|u_1-u_2| + |\tau(u_1) - \tau(u_2)| \leqslant 2(|u_1-u_2| \wedge 1),$$

on $\{(u_1,u_2) : |u_1-u_2| \leqslant 1\}$. If, in addition, $|u_1| \leqslant 1, |u_2| \leqslant 1$, then $\tau(u_j) = u_j, j = 1,2$ and the integrand vanishes. Hence by the Cauchy-Schwarz-Bunjakovski inequality and equation (47)

$$\int_{|u_1-u_2|\leqslant 1}\Big[\tau(u_1-u_2)-(\tau(u_1)-\tau(u_2))\Big]\Pi^{T,r}(du) \leqslant 2\int_{|u_1|>1 \text{ or } |u_2|>1}\Big(|u_1-u_2|\wedge 1\Big)\Pi^{T,r}(du)$$

$$\leqslant 2\left(\int_{\mathbb{R}^2}\Big((u_1-u_2)^2\wedge 1\Big)\Pi^{T,r}(du)\right)^{1/2}\Big(\Pi^{T,r}(\{u : |u_1|>1 \text{ or } |u_2|>1\})\Big)^{1/2}$$

$$\leqslant 2\Big(\min(\Pi_1(|u|>1)+\Pi_2(|u|>1), r)\ \ rW^2_{2,\rho}(\pi^r_1,\pi^r_2)\Big)^{1/2}. \qquad (48)$$

Therefore we obtain

$$g_5(z_1,z_2) \leqslant (\pi+3^{3/4})T^2_r(\Pi_1,\Pi_2) + 3^{3/4}\min(\Pi_1(|u|>1)+\Pi_2(|u|>1), r)^{1/2}\, T_r(\Pi_1,\Pi_2).$$

Then, summarizing all the above, we get

$$g(z_1,z_2) \leqslant \frac{L}{\arctan(1/2)}F(z_1-z_2) + \frac{3^{3/4}}{2}|a_1-a_2| + (\sigma_1-\sigma_2)^2 + U_r(\Pi_1) + U_r(\Pi_2) +$$

$$+ (\pi+3^{3/4})T^2_r(\Pi_1,\Pi_2) + 3^{3/4}\min(\Pi_1(|u|>1)+\Pi_2(|u|>1), r)^{1/2}\, T_r(\Pi_1,\Pi_2). \qquad (49)$$

Denote by

$$\hat{Q}_r = \frac{3^{3/4}}{2}|a_1-a_2| + (\sigma_1-\sigma_2)^2 + U_r(\Pi_1) + U_r(\Pi_2) +$$

$$+ (\pi+3^{3/4})T^2_r(\Pi_1,\Pi_2) + 3^{3/4}\min(\Pi_1(|u|>1)+\Pi_2(|u|>1), r)^{1/2}\, T_r(\Pi_1,\Pi_2),$$

and also we denote

$$Q_r = \hat{Q}_r + (\pi/2)\rho^2(x_1,x_2).$$

Recall that

$$F(y) \leqslant y \wedge (\pi/2) \leqslant (\pi/2)(y\wedge 1),$$

and therefore $F(Y(0)) \leqslant (\pi/2)\rho^2(x_1,x_2)$. Then, due to $F \geqslant 0$ and $\hat{Q}_r \geqslant 0$, representation (41) and bound (49) yield that for $t \in [0,1]$

$$F(Y(t)) \leqslant F(Y(0)) + \frac{L}{\arctan(1/2)}\int_0^t F(Y(s))\,ds + \hat{Q}_r + M_t$$

$$\leqslant Q_r + \frac{L}{\arctan(1/2)}\int_0^1 F(Y(s))\,ds + M_t \qquad a.s.$$

Hence we get that

$$\mathbf{E}F(Y(t)) \leqslant Q_r + \frac{L}{\arctan(1/2)}\int_0^1 \mathbf{E}F(Y(s))\,ds \qquad (50)$$

and

$$\mathbf{E}\sup_{t\in[0,1]} F(Y(t)) \leqslant Q_r + \frac{L}{\arctan(1/2)}\int_0^1 \mathbf{E}F(Y(s))\,ds + \mathbf{E}\sup_{t\in[0,1]}|M_t|. \qquad (51)$$

Applying the Gronwall lemma, we obtain from (50)

$$\mathbf{E}F(Y(t)) \leqslant Q_r \exp(\frac{Lt}{\arctan(1/2)}) \qquad (52)$$

which makes it possible to re-write (51) to the form

$$\mathbf{E}\sup_{t\in[0,1]}F(t)\leqslant Q_r\exp(\frac{L}{\arctan(1/2)})+\mathbf{E}\sup_{t\in[0,1]}|M_t|. \tag{53}$$

To estimate martingale term in (53) we apply Doob's maximal moment inequality

$$\mathbf{E}\sup_{t\in[0,T]}|M_t|^p\leqslant\left(\frac{p}{p-1}\right)^p\mathbf{E}|M_T|^p.$$

with $p=2$:

$$\begin{aligned}\mathbf{E}\sup_{t\in[0,1]}M_t^2\leqslant 4\mathbf{E}M_1^2&=4(\sigma_1-\sigma_2)^2\int_0^1\mathbf{E}(F'(Y(s)))^2\,ds\\&+4\int_0^1\int_{\mathbb{R}^2}\mathbf{E}\Big[F(Y(s)+(u_1-u_2))-F(Y(s))\Big]^2ds\hat{\Pi}(du)\\&+4\int_0^1\int_{\mathbb{R}^2}\mathbf{E}\Big[F(Y(s)+(u_1-u_2))-F(Y(s))\Big]^2ds\Pi^{T,r}(du);\end{aligned}\tag{54}$$

the last identity comes from (40), note that the noises involved in the three summands in the right hand side of (40) are independent.

Now, let us estimate three terms in the r.h.s. of (54).

1. Using the inequality (42) and the bound (52) we have

$$4(\sigma_1-\sigma_2)^2\int_0^1\mathbf{E}\,(F'(Y(s)))^2\,ds\leqslant 4(\sigma_1-\sigma_2)^2\frac{3^{3/2}}{4}=3^{3/2}(\sigma_1-\sigma_2)^2.$$

2. Using Taylor estimate and inequality (42) and the analogous reasoning to (46) yield

$$[F(Y(s)+(u_1-u_2))-F(Y(s))\Big]^2\leqslant\Big[\pi\wedge\Big(\frac{3^{3/4}}{2}|u_1-u_2|\Big)\Big]^2\leqslant\pi^2\Big(|u_1-u_2|^2\wedge 1\Big).$$

Then

$$\begin{aligned}4\int_0^1\int_{\mathbb{R}^2}\mathbf{E}\Big[F(Y(s)+(u_1-u_2))-F(Y(s))\Big]^2ds\hat{\Pi}(du)&\leqslant 4\pi^2\int_{\mathbb{R}^2}\Big(|u_1-u_2|^2\wedge 1\Big)\hat{\Pi}(du)\\&=4\pi^2\left[U_r\left(\Pi_1\right)+U_r\left(\Pi_2\right)\right].\end{aligned}$$

3. Similarly,

$$\begin{aligned}4\int_0^1\int_{\mathbb{R}^2}\mathbf{E}\Big[F(Y(s)+(u_1-u_2))-F(Y(s))\Big]^2ds\Pi^{T,r}(du)&\leqslant 4\pi^2\int_{\mathbb{R}^2}\Big(|u_1-u_2|^2\wedge 1\Big)\Pi^{T,r}(du)\\&=4\pi^2T_r^2(\Pi_1,\Pi_2).\end{aligned}$$

Eventually the martingale estimates amount to

$$\mathbf{E}\sup_{t\in[0,1]}|M_t|\leqslant\Big(\mathbf{E}\sup_{t\in[0,1]}M_t^2\Big)^{1/2}\leqslant\sqrt{3^{3/2}|\sigma_1-\sigma_2|^2+(2\pi)^2(U_r\left(\Pi_1\right)+U_r\left(\Pi_2\right)+T_r^2(\Pi_1,\Pi_2))}. \tag{55}$$

Collecting (53) and (55) yields the final bound for (36). This completes the proof. □

4.3 Proof of Theorem 2

Theorem 2 is almost a direct consequence of Theorem 1. Recall that when $r \to \infty$, we have $\varepsilon_j(r) \to 0$, $U_r(\Pi_j) \to 0$ for $j = 1, 2$ and $T_r(\Pi_1, \Pi_2) \to T(\Pi_1, \Pi_2)$. For $r \to \infty$, $Q^1_r \to Q^{1'}$ and $Q^2_r \to Q^{2'}$, where

$$
\begin{aligned}
Q^{1'} &= 2\rho^2(x_1, x_2) + \frac{4}{\pi}\Big(\frac{3^{3/4}}{2}|a_1 - a_2| + (\sqrt{b_1} - \sqrt{b_2})^2 + \\
&\quad + (\pi + 3^{3/4})T^2(\Pi_1, \Pi_2) + 3^{3/4}\big(\Pi_1(|u| > 1) + \Pi_2(|u| > 1)\big)^{1/2} T(\Pi_1, \Pi_2)\Big), \\
Q^{2'} &= \frac{4}{\pi}\sqrt{3^{3/2}(\sqrt{b_1} - \sqrt{b_2})^2 + (2\pi)^2(T^2(\Pi_1, \Pi_2))}.
\end{aligned}
$$

In a last step we take the maximum of all prefactors. □

Acknowledgements

The first and the second author would like to express their gratitude to the International Research Training Group (IRTG) 1740 Berlin-São Paulo: "Dynamical Phenomena in Complex Networks: Fundamentals and Applications" and the mathematics department of UNICAMP, Brazil, in particular to Prof. Paulo Ruffino for their support. The second, the third and the fourth author would like to thank Universität Potsdam for special travel support, and DAAD for a travel grant of the "Leonhard-Euler" Program, Project Nr. 55518603, in May 2013. All the authors thank Prof. Sylvie Roelly at Universität Potsdam for all the moral, travel and infrastructure support.

References

[1] L. Arnold. Hasselmann's program revisited: The analysis of stochasticity in deterministic climate models. *in Imkeller, Peter (ed.) et al., Stochastic climate models. Basel: Birkhäuser. Prog. Probab.* 49:141-157, 2001.

[2] R. Benzi, G. Parisi, A. Sutera, A. Vulpiani. Stochastic resonance in climate change. *Tellus*, 34:10-16, 1982.

[3] N. Berglund, D. Landon. Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh-Nagumo model. *Nonlinearity*, 25:2303-2335, 2012.

[4] N. Berglund, B. Gentz. Metastability in simple climate models: Pathwise analysis of slowly driven Langevin equations. *Stoch. Dyn.*, 2:327-356, 2002.

[5] P.D. Ditlevsen. Observation of a stable noise induced millennial climate changes from an ice-core record. *Geophysical Research Letters*, 26 (10):1441–1444, 1999.

[6] R.M. Dudley. Real analysis and probability. Cambridge University Press, New York, 2004.

[7] J. Gairing, M. Högele, T. Kosenkova, A. Kulik. Coupling distances between Lévy measures and statistical inference for paleoclimate data. *In preparation.*

[8] J. Gairing, P. Imkeller. Stable CLTs and rates for power variation of α-stable Lévy processes. *Methodology and Computing in Applied Probability*, 1-18, 2013.

[9] C. Hein, P. Imkeller, and I. Pavlyukevich. Limit theorems for p-variations of solutions of SDEs driven by additive stable Lévy noise and model selection for paleo-climatic data. *Interdisciplinary Math. Sciences*, 8:137-150, 2009.

[10] A. Debussche, M. Högele, P. Imkeller. The dynamics of non-linear reaction-diffusion equations with small Lévy noise. *Springer Lecture Notes in Mathematics, Vol. 2085*, 2013.

[11] C. Doss, M. Thieullen. Oscillations and random perturbations of a FitzHugh-Nagumo system. *Preprint hal-00395284*, 2009.

[12] K. Hasselmann. Stochastic climate models: Part I. Theory. *Tellus*, 28:473–485, 1976.

[13] M. Högele, I. Pavlyukevich. The exit problem from the neighborhood of a global attractor for heavy-tailed Lévy diffusions. *Stochastic Analysis and Applications*, to appear, 2013.

[14] P. Imkeller, A. Monahan. Conceptual stochastic climate models. *Stochastics and Dynamics*, 2:311–326, 2002.

[15] P. Imkeller. Energy balance models: Viewed from stochastic dynamics. *Prog. Prob.*, 49:213–240, 2001.

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

[17] P. Imkeller and I. Pavlyukevich. Metastable behaviour of small noise Lévy-driven diffusions. *ESAIM: Probability and Statistics*, 12:412–437, 2008.

[18] P. Imkeller, I. Pavlyukevich, T. Wetzel. First exit times for Lévy-driven diffusions with exponentially light jumps. *Ann. Probab.*, 37 (2) 530-564, 2009.

[19] B.V. Gnedenko, A.N. Kolmogorov. Limit distributions for sums of independent random variables. Addison-Wesley, 1968.

[20] N. Ikeda and S. Watanabe. Stochastic differential equations and diffusion processes. North-Holland, Kodansha ltd., Tokyo, 1981.

[21] Members NGRIP. NGRRIP data. *Nature*, 431:147–151, 2004.

[22] V.S. Koroliuk, N. Limnios, I.V. Samoilenko. Lévy approximation of impulsive recurrent process with semi-Markov switching. *Theory of Probability and Mathematical Statistics*, 80:85–92, 2009.

[23] T. Kosenkova. Weak convergence of the series scheme of Markov chains to the solution of Lévy driven SDE. *Theory of Stochastic Processes*, 18(34): 86–100, 2012.

[24] T. Kosenkova, A. Kulik. Explicit bounds for the convergence rate of a sequence of Markov chains to a Lévy-type process. *In preparation.*

[25] I. Pavlyukevich. First exit times of solutions of stochastic differential equations with heavy tails. *Stochastics and Dynamics*, 11, Exp.No.2/3:1–25, 2011.

[26] S. T. Rachev and L. Rüschendorf. Mass Transportation Problems. Vol.I: Theory, Vol.II.: Applications. *Probability and its Applications*. Springer-Verlag, New York 1998.

[27] H.C. Tuckwell, R. Rodriguez, F.Y.M. Wan, Determination of firing times for the stochastic Fitzhugh-Nagumo neuronal model. *Neural Computation*, 15:143 – 159, 2003.