ALEA, Lat. Am. J. Probab. Math. Stat. 17, 643–674 (2020)
DOI: 10.30757/ALEA.v17-25

On the anisotropic stable JCIR process

Martin Friesen and Peng Jin

School of Mathematics and Natural Sciences, University of Wuppertal
Gaußstraße 20, 42119 Wuppertal, Germany.
E-mail address: friesen@math.uni-wuppertal.de

Department of Mathematics, Shantou University
243 Daxue Road, Shantou, Guangdong 515063, China.
E-mail address: pjin@stu.edu.cn

Abstract. We investigate the anisotropic stable JCIR process which is a multi-dimensional extension of the stable JCIR process but also a multi-dimensional analogue of the classical JCIR process. We prove that the heat kernel of the anisotropic stable JCIR process exists and it satisfies an a-priori bound in a weighted anisotropic Besov norm. Based on this regularity result we deduce the strong Feller property and prove, for the subcritical case, exponential ergodicity in total variation. Also, we show that in the one-dimensional case the corresponding heat kernel is smooth.

1. Introduction

The classical JCIR process is a commonly used building block for different models in mathematical finance, see Alfonsi (2015). For given $b, \sigma \geq 0$ and $\beta \in \mathbb{R}$ it is obtained as the unique $\mathbb{R}_+$-valued strong solution to

$$\mathrm{d}X^x(t) = (b + \beta X(t))\mathrm{d}t + \sqrt{\sigma X(t)}\mathrm{d}B(t) + \mathrm{d}J(t), \qquad X^x(0) = x \geq 0,$$

where $(B(t))_{t\geq 0}$ is a one-dimensional Brownian motion and $(J(t))_{t\geq 0}$ is a Lévy subordinator on $\mathbb{R}_+$ that is independent of the Brownian motion. For a particular choice of subordinator $(J(t))_{t\geq 0}$ such a process was first introduced in Duffie and Gârleanu (2001). Some of its specific properties were studied in Jin et al. (2019), see also the references therein. Replacing the Brownian motion $(B(t))_{t\geq 0}$ by a spectrally positive α-stable Lévy process $(Z^\alpha(t))_{t\geq 0}$ whose symbol is given, for

Received by the editors August 15th, 2019; accepted June 17th, 2020.
2010 *Mathematics Subject Classification.* 60H10, 60J25, 60J35, 37A25.
Key words and phrases. Stable JCIR Process, Affine Process, Heat Kernel, Anisotropic Besov Space, Strong Feller Property, Exponential Ergodicity.
The research of Peng Jin was supported by the STU Scientific Research Foundation for Talents (No. NTF18023).

$\alpha \in (1,2)$, by

$$\Psi_\alpha(\xi) = \int_0^\infty \left(\mathrm{e}^{\mathrm{i}\xi z} - 1 - \mathrm{i}\xi z\right) \mu_\alpha(\mathrm{d}z), \qquad \mu_\alpha(\mathrm{d}z) = \mathbb{1}_{\mathbb{R}_+}(z) \frac{1}{c(\alpha)} \frac{\mathrm{d}z}{z^{1+\alpha}}, \tag{1.1}$$

and replacing the square-root by $\sqrt[\alpha]{\cdot}$ one obtains the stable JCIR process

$$\mathrm{d}X^x(t) = (b + \beta X^x(t))\mathrm{d}t + \sqrt[\alpha]{\sigma X^x(t)}\mathrm{d}Z^\alpha(t) + \mathrm{d}J(t), \quad X^x(0) = x \geq 0. \tag{1.2}$$

Note that the normalization constant $c(\alpha) = \int_0^\infty (e^{-z} - 1 + z)\, z^{-1-\alpha}\mathrm{d}z$ is chosen in such a way that $\Psi_\alpha(\mathrm{i}\xi) = \xi^\alpha$ for $\xi \geq 0$. This process is a special case of the short-rate models used in Jiao et al. (2017, 2018); Chazal et al. (2018a). One important advantage of these models is their analytical tractability as many desired expressions (e.g. the Laplace transform) can be computed explicitly.

In this work we study the *anisotropic stable JCIR process*, i.e., the multi-dimensional analogue of the stable JCIR process (1.2), obtained as the unique $\mathbb{R}_+^m$-valued strong solution to the system of stochastic equations

$$\mathrm{d}X_k^x(t) = \left(b_k + \sum_{j=1}^m \beta_{kj} X_j^x(t)\right)\mathrm{d}t + \sqrt[\alpha_k]{\sigma_k X_k^x(t)}\mathrm{d}Z_k(t) + \mathrm{d}J_k(t), \tag{1.3}$$

where $k \in \{1, \ldots, m\}$, $X^x(0) = x \in \mathbb{R}_+^m$, $b = (b_1, \ldots, b_m), (\sigma_1, \ldots, \sigma_m) \in \mathbb{R}_+^m$ and $\beta = (\beta_{jk})_{j,k\in\{1,\ldots,m\}}$ is such that $\beta_{jk} \geq 0$ for all $j \neq k$. Here $Z_1, \ldots, Z_m$ are independent and each Z_k, $k = 1, \ldots, m$, is a one-dimensional spectrally positive α_k-stable Lévy process with symbol Ψ_{α_k} as in (1.1), where $\alpha_1, \ldots, \alpha_m \in (1,2)$. The process J, which is independent of $Z = (Z_1, \ldots, Z_m)$, is a Lévy subordinator on $\mathbb{R}_+^m$, i.e., its Lévy measure ν is supported on $\mathbb{R}_+^m$ and J has symbol

$$\Psi_J(\xi) = \int_{\mathbb{R}_+^m} \left(\mathrm{e}^{\mathrm{i}\langle \xi, z\rangle} - 1\right)\nu(\mathrm{d}z), \qquad \int_{\mathbb{R}_+^m} \min\{1, |z|\}\nu(\mathrm{d}z) < \infty.$$

It follows from Barczy et al. (2015) that (1.3) has a unique $\mathbb{R}_+^m$-valued strong solution. Moreover, this process is an affine process on state space $\mathbb{R}_+^m$ (see Duffie et al., 2003; Barczy et al., 2015) whose characteristic function satisfies

$$\mathbb{E}[\mathrm{e}^{\langle u, X^x(t)\rangle}] = \mathrm{e}^{\phi(t,u) + \langle x, \psi(t,u)\rangle}, \qquad x \in \mathbb{R}_+^m, \tag{1.4}$$

where $u \in \mathbb{C}^m$ is such that $\mathrm{Re}(u) \leq 0$. Here ϕ and $\psi = (\psi_1, \ldots, \psi_m)$ are the unique solutions to the generalized Riccati equations

$$\begin{cases} \partial_t \phi(t,u) = F(\psi(t,u)), & \phi(t,0) = 0, \\ \partial_t \psi(t,u) = R(\psi(t,u)), & \psi(t,0) = u, \end{cases} \tag{1.5}$$

where F and $R = (R_1, \ldots, R_m)$ are given by

$$F(u) = \langle b, u\rangle + \int_{\mathbb{R}_+^m} \left(\mathrm{e}^{\langle u, z\rangle} - 1\right)\nu(\mathrm{d}z),$$

$$R_j(u) = \sum_{k=1}^m \beta_{kj} u_k + \int_0^\infty \left(\mathrm{e}^{u_j z} - 1 - u_j z\right)\mu_{\alpha_j}(\mathrm{d}z).$$

Following the general theory of affine processes it can be shown that $(X^x(t))_{t\geq 0}$ is a Feller process, that its transition semigroup acts on the Banach space of continuous functions vanishing at infinity, and that its generator $(L, D(L))$ has core $C_c^\infty(\mathbb{R}_+^m)$ and for $f \in C_c^\infty(\mathbb{R}_+^m)$

$$\begin{aligned} Lf(x) = \langle b + \beta x, \nabla f(x)\rangle + \int_{\mathbb{R}_+^m} (f(x+z) - f(x))\,\nu(\mathrm{d}z) \\ + \sum_{j=1}^m \sigma_j x_j \int_0^\infty \left(f(x + e_j z) - f(x) - z\frac{\partial f(x)}{\partial x_j}\right) \mu_{\alpha_j}(\mathrm{d}z), \end{aligned}$$

where $e_1, \dots, e_m$ denote the canonical basis vectors in $\mathbb{R}^m$.

The purpose of this work is twofold. Firstly, we investigate regularity of the heat kernel including a very simple proof of the strong Feller property, and secondly, based on the obtained results we study the convergence to equilibrium in total variation. On the way proving these results we also obtain non-extinction for the anisotropic stable JCIR process in the spirit of Foucart and Uribe Bravo (2014); Duhalde et al. (2014); Friesen et al. (2019b).

One commonly used method to study existence and smoothness of heat kernels is based on Malliavin calculus, see e.g. Bass and Cranston (1986); De Marco (2011); Picard (1996) and the references therein. Concerning other analytical methods we refer to Bogdan et al. (2020); Chaker (2019); Knopova and Kulik (2018); Kulczycki and Ryznar (2018, 2019); Kulczycki et al. (2020+) where some interesting progress for stochastic equations driven by cylindrical Lévy processes has been obtained. Having in mind that the anisotropic stable JCIR process has no diffusion component, that the Lévy measure of the driving noise is singular and has no second moments, and finally that the volatility coefficients in (1.3) are merely Hölder continuous and degenerate at the boundary, it is not clear how the aforementioned techniques could be applied in the setting of this paper. Based on the affine structure of the process it is reasonable to study the heat kernel by Fourier methods similarly to Filipović et al. (2013), where affine processes with non-degenerate diffusion component were treated, or by spectral expansions in the spirit of Chazal et al. (2018b). While Fourier methods turn out to be adequate for proving existence of a smooth density for the one-dimensional stable JCIR process (see Section 3), it seems difficult to extend them to the anisotropic framework with absent diffusion component. Moreover, it would be interesting to extend the techniques developed in Chazal et al. (2018b) to this multi-dimensional setting. In contrast, our approach for the study of the multi-dimensional case is based on a suitable short-time approximation of the process combined with a discrete integration by parts in the spirit of Debussche and Fournier (2013), Romito (2018), Friesen et al. (2020) and Friesen et al. (2020+a). Since these methods do not use the affine structure of the process, they can be applied to other Markov processes as well.

The long-time behavior of one-dimensional affine processes with state space $\mathbb{R}_+$ was studied in Keller-Ressel and Mijatović (2012), Li (2011, Chapter 3) and Li and Ma (2015). Results applicable to a class of non-affine Markov processes on $\mathbb{R}_+$ have been recently obtained in Friesen et al. (2019a). The coupling method in Li and Ma (2015) is very effective for 1-dimensional continuous-state branching processes with immigration. However, it used the fact the the extinction time of a continuous-state branching process can be estimated via the Laplace transform of the process. It is not clear if this approach can be extended to higher dimensional cases. For subcritical OU-type processes and 1-dimensional continuous-state branching processes with immigration, the exponential ergodicity in total variation has been derived under rather general conditions, see Wang (2012); Li and Ma (2015) and Friesen et al. (2019a), all of which used coupling techniques. Other than these two cases, only very few results on ergodicity in total variation are available for multi-dimensional affine processes, except for the models treated in Barczy et al. (2014); Jin et al. (2017); Mayerhofer et al. (2020); Zhang and Glynn (2018). The reason is as follows: in the general case it is not clear if the powerful coupling technique (see Wang, 2012; Li and Ma, 2015) still works; also, it remains a difficult problem to verify the irreducibility of the process when applying the Meyn-Tweedie method (see Meyn and Tweedie, 2009). To overcome these difficulties we use instead a Harris-type theorem based on a local Dobrushin condition, see Theorem D.1 and Hairer (2016); Kulik (2018). In order to verify the local Dobrushin condition we use continuity (regularity) of the heat kernel combined with some weak form of irreducibility similarly to Peng and Zhang (2018). At this point it is worthwhile to mention that the verification of the local Dobrushin condition does not require the full strength of our regularity result. Indeed, one could apply Kulik (2018, Proposition 2.9.1 and Remark 2.9.2) for which the Besov regularity from Section 4 is sufficient. This work seems to provide the first result on ergodicity in total variation for multi-dimensional affine processes which does not rely on smoothing properties of the diffusion component. Moreover, the method of this paper can be also applied to non-affine Markov processes.

This paper is organized as follows. In Section 2 we state and discuss the main results of this work. Regularity of the heat kernel for the one-dimensional stable JCIR process as in (1.2) is discussed in Section 3. Regularity of the anisotropic stable JCIR process is studied in Section 4, while ergodicity in total variation is proved in Section 5. Finally, some auxiliary results and general theory on ergodicity of Markov processes are collected in the appendix.

2. Statement of results

2.1. *Existence and smoothness of the heat kernel in dimension $m = 1$.* Let $(X^x(t))_{t\geq 0}$ be the one-dimensional stable JCIR process, i.e., the unique $\mathbb{R}_+$-valued strong solution to (1.2) and $P_t(x, \mathrm{d}y)$ its transition probability kernel. The following is our first main result.

Theorem 2.1. *Suppose that there exist constants $C, M > 0$ and $\vartheta \in (\alpha - 1, 1]$ such that*

$$b\xi + \int_0^\infty \left(1 - \mathrm{e}^{-z\xi}\right) \nu(\mathrm{d}z) \geq C\xi^{\vartheta}, \qquad \xi \geq M. \tag{2.1}$$

Then for each $t > 0$ and each $x \geq 0$, the heat kernel $P_t(x, \mathrm{d}y)$ has density $p_t(x, y)$ which is jointly continuous in $(t, x, y) \in (0, \infty) \times [0, \infty)^2$. Moreover, for each $t > 0$, the function $\mathbb{R}_+ \times \mathbb{R}_+ \ni (x, y) \longmapsto p_t(x, y)$ is smooth and

$$\sup_{(x,y)\in\mathbb{R}_+\times\mathbb{R}_+} |\partial_x^n \partial_y^k p_t(x, y)| < \infty, \qquad \forall n, k \in \mathbb{N}_0.$$

Condition (2.1) is natural to guarantee that the process does not hit the boundary and hence $P_t(x, \mathrm{d}y)$ has no atom at the boundary, i.e. $p_t(x, y)$ is also continuous at $y = 0$. Let us refer to Foucart and Uribe Bravo (2014); Duhalde et al. (2014) for some related results. If $b > 0$, then (2.1) is satisfied with $\vartheta = 1$ and $C = b$. In the case $b = 0$ condition (2.1) is still satisfied provided that the subordinator J has sufficiently many small jumps (see the examples at the end of this section).

The proof of Theorem 2.1 is given in Section 3 and deeply relies on the affine structure of the process (see (1.4)), i.e. we exploit the fact that its characteristic function satisfies

$$\mathbb{E}[\mathrm{e}^{uX^x(t)}] = \mathrm{e}^{\phi(t,u)+x\psi(t,u)}, \qquad t \geq 0,\ x \geq 0, \tag{2.2}$$

where $u \in \mathbb{C}$ is such that $\mathrm{Re}(u) \leq 0$, and ϕ, ψ solve uniquely the generalized Riccati equations

$$\begin{cases} \partial_t \phi(t,u) = b\psi(t,u) + \int_0^\infty \left(\mathrm{e}^{z\psi(t,u)} - 1\right)\nu(\mathrm{d}z), & \phi(0,u) = 0 \\ \partial_t \psi(t,u) = \beta\psi(t,u) + \int_0^\infty \left(\mathrm{e}^{z\psi(t,u)} - 1 - \psi(t,u)z\right)\mu_\alpha(\mathrm{d}z), & \psi(0,u) = u. \end{cases} \tag{2.3}$$

We deduce the assertion by showing enough integrability for the characteristic function $\mathbb{R} \ni u \longmapsto \mathbb{E}[\mathrm{e}^{\mathrm{i}uX^x(t)}]$. For this purpose we adapt some ideas from the multi-dimensional diffusion case studied in Filipović et al. (2013), where a Hörmander-type condition on the drift and diffusion parameters is imposed.

2.2. *Existence of heat-kernel and strong Feller property in dimension $m \geq 1$.* Here and below we denote by $(X^x(t))_{t\geq 0}$ the anisotropic stable JCIR process obtained from (1.3) with initial condition $X^x(0) = x \in \mathbb{R}_+^m$, and recall that it depends on the parameters $b \in \mathbb{R}_+^m, \sigma_1, \ldots, \sigma_m \geq 0$, $\alpha_1, \ldots, \alpha_m \in (1,2)$, $(\beta_{kj})_{k,j=1,\ldots,m}$ with $\beta_{kj} \geq 0$ for $k \neq j$, and a Lévy subordinator $\nu(\mathrm{d}z)$ on $\mathbb{R}_+^m$. Finally, let us assume that $\sigma_1, \ldots, \sigma_m > 0$. The case where $\sigma_i = 0$ holds for some $i \in \{1, \ldots, m\}$ can be also studied by the methods of this paper provided we assume an additional "non-degeneracy" condition on the one-dimensional Lévy process $(J_i(t))_{t\geq 0}$. However, in order to keep the arguments simple and neat we decided to exclude these cases. The following condition is a multi-dimensional analogue of (2.1).

(A) There exist constants $C, M > 0$ and $\vartheta_1, \ldots, \vartheta_m$ such that for all $k = 1, \ldots, m$, $\vartheta_k \in (\alpha_k - 1, 1]$ and

$$b_k\xi + \int_{\mathbb{R}_+^m} \left(1 - \mathrm{e}^{-\xi z_k}\right)\nu(\mathrm{d}z) \geq C\xi^{\vartheta_k}, \qquad \forall \xi \geq M. \tag{2.4}$$

Remark 2.2. If $b \in \mathbb{R}_{++}^m = \{x \in \mathbb{R}_+^m \mid x_1, \ldots, x_m > 0\}$, then condition (A) is satisfied. If $b \in \partial\mathbb{R}_+^m$, then condition (A) is still satisfied provided that the Lévy process J has sufficiently many jumps in direction k with $b_k = 0$. Some particular examples satisfying condition (A) with $b \in \partial\mathbb{R}_+^m$ are given in the end of this section.

Condition (A) guarantees that the process has a sufficiently strong drift pointing inwards (i.e. in the interior $\mathbb{R}_{++}^m$) and hence does not hit the boundary of its state space, see Section 4 for additional details. The next remark states that (A) imposes essentially a condition that is independent of the big jumps of the subordinator J.

Remark 2.3. Let $b \in \mathbb{R}_+^m$ and let ν be a Lévy measure on $\mathbb{R}_+^m$. Then condition (A) is satisfied for b, ν if and only if it is satisfied for $b, \mathbb{1}_{\{|z|\leq 1\}}\nu(\mathrm{d}z)$.

The following is our main regularity result for the heat kernel of the anisotropic stable JCIR process.

Theorem 2.4. *Suppose that condition (A) is satisfied. Then $P_t(x, \mathrm{d}y) = p_t(x,y)\mathrm{d}y$ and*

$$\mathbb{R}_+^m \ni x \longmapsto p_t(x, \cdot) \in L^1(\mathbb{R}_+^m)$$

is continuous for each $t > 0$. In particular, the anisotropic stable JCIR process has the strong Feller property.

The proof of this result is given in Section 4 and is divided into 4 steps. Namely, we first prove existence of a heat kernel under an additional moment condition for ν and provide an estimate in a suitably weighted anisotropic Besov norm which takes also the behavior of the process at the boundary into account. Secondly, we estimate uniformly the probability that the process hits its boundary in positive time. Then, with the same moment condition for ν, we deduce the assertion from a compactness argument combined with previous two steps. Finally, we use a convolution trick to remove the extra moment assumption and prove the assertion in the general case. The same approach can also be applied to general affine (and non-affine) processes.

2.3. *Exponential ergodicity in total variation.* The anisotropic stable JCIR process is called *subcritical*, if $\beta = (\beta_{jk})_{j,k\in\{1,\dots,m\}}$ has only eigenvalues with negative real-parts. Assuming that the anisotropic stable JCIR process is subcritical and satisfies

$$\int_{\mathbb{R}_+^m} \mathbb{1}_{\{|z|>1\}} \log(1+|z|)\nu(\mathrm{d}z) < \infty, \tag{2.5}$$

existence, uniqueness and a representation of the characteristic function for the invariant measure π was first obtained in Jin et al. (2020) where stability for the corresponding Riccati equations was investigated. Then

$$\int_{\mathbb{R}_+^m} \log(1+|x|)\pi(\mathrm{d}x) < \infty \tag{2.6}$$

and an exponential rate of convergence for $P_t(x,\cdot) \longrightarrow \pi$ in different Wasserstein distances was shown in Friesen et al. (2020+b) where affine processes on the canonical state space have been obtained as unique strong solutions to a system of stochastic equations. For one-dimensional affine processes on $\mathbb{R}_+$ regularity (and other properties) of the invariant measure π was studied in Chazal et al. (2018b); Keller-Ressel and Mijatović (2012). Using the regularity for the heat kernel obtained in Theorem 2.4 we prove exponential ergodicity in the total variation norm

$$\|\rho\|_{\mathrm{TV}} = \sup_{A\in\mathcal{B}(\mathbb{R}_+^m)} |\rho|(A) = \sup_{\|f\|_\infty\le 1} \left| \int_{\mathbb{R}_+^m} f(x)\rho(\mathrm{d}x)\right|, \tag{2.7}$$

where $|\rho| = \rho^+ + \rho^-$ and $\rho^\pm$ denote the Hahn-Jordan decomposition of a signed Borel measure ρ on $\mathbb{R}_+^m$. Our last main result provides a sufficient condition for the exponential ergodicity in the stronger total variation distance.

Theorem 2.5. *Suppose that the anisotropic stable JCIR process is subcritical, satisfies condition (A) and (2.5). Then there exist constants $C, \delta > 0$ such that for all $t \ge 0$ and $x \in \mathbb{R}_+^m$*

$$\|P_t(x,\cdot) - \pi\|_{\mathrm{TV}} \le C\left(1 + \log(1+|x|) + \int_{\mathbb{R}_+^m} \log(1+|y|)\pi(\mathrm{d}y)\right) \mathrm{e}^{-\delta t}.$$

The proof of this theorem is based on a Harris-type theorem and is given in Section 5. It basically requires to check a local Dobrushin condition and a Foster-Lyapunov drift condition for the extended generator. The local Dobrushin condition is deduced from the regularity results from Theorem 2.4 combined with a weak form of irreducibility similar to Peng and Zhang (2018). Finally, the Foster-Lyapunov condition can be checked by direct computation combined with a convolution argument similar to Friesen et al. (2020+b); Jin et al. (2020).

2.4. *Examples for main conditions.* Recall that condition (A) is satisfied, if $b \in \mathbb{R}_{++}^m$. So let us consider the case $b \in \partial\mathbb{R}_+^m$. For simplicity, we suppose that $b = 0$ and provide conditions on ν such that (A) is still satisfied. Set

$$\alpha_{\max} := \max\{\alpha_1, \dots, \alpha_m\}, \quad \alpha_{\min} := \min\{\alpha_1, \dots, \alpha_m\}.$$

Example 2.6. Let ν be given by the spherical decomposition

$$\nu(A) = \int_0^\infty \int_{S_+^{m-1}} \mathbb{1}_A(r\sigma)\lambda(\mathrm{d}\sigma)\frac{\mathrm{d}r}{r^{1+\vartheta}} \tag{2.8}$$

where $\vartheta \in (\alpha_{\max} - 1, 1)$, $S_+^{m-1} = \{\sigma \in \mathbb{R}_+^m \mid |\sigma| = 1\}$, and λ is a measure on S_+^{m-1}. Then we obtain

$$\int_{\mathbb{R}_+^m} \left(1 - \mathrm{e}^{-\xi z_k}\right)\nu(\mathrm{d}z) = \xi^\vartheta \int_{S_+^{m-1}} \sigma_k^\vartheta \lambda(\mathrm{d}\sigma) \int_0^\infty \left(1 - \mathrm{e}^{-r}\right)\frac{\mathrm{d}r}{r^{1+\vartheta}}.$$

Hence (A) holds, if $\int_{S_+^{m-1}} \sigma_k^\vartheta \lambda(\mathrm{d}\sigma) > 0$. This includes the following cases:

(a) If $\lambda(\mathrm{d}\sigma) = \mathbb{1}_{S_+^{m-1}}(\sigma)\mathrm{d}\sigma$ is the uniform distribution on S_+^{m-1}, then

$$\nu(\mathrm{d}z) = \mathbb{1}_{\mathbb{R}_+^m}(z)\frac{\mathrm{d}z}{|z|^{d+\vartheta}}.$$

(b) If $\lambda(\mathrm{d}\sigma) = \sum_{k=1}^m \delta_{e_k}(\mathrm{d}\sigma)$, then

$$\nu(\mathrm{d}z) = \sum_{k=1}^m \mathbb{1}_{\mathbb{R}_+}(z_k)\frac{\mathrm{d}z_k}{z_k^{1+\vartheta}} \otimes \prod_{j\neq k}\delta_0(\mathrm{d}z_j).$$

The next example shows that the stability index ϑ appearing in (2.8) is also allowed to depend on the direction of the jump.

Example 2.7. Let $J(t) = (J_1(t), \dots, J_m(t))$ where $J_1, \dots, J_m$ are independent Lévy subordinators on $\mathbb{R}_+$ with Lévy measures $\mathbb{1}_{\mathbb{R}_+}(z_k)z_k^{-1-\vartheta_k}\mathrm{d}z_k$ with $\nu_k \in (\alpha_k - 1, 1)$and $k = 1, \dots, m$. Then J has Lévy measure

$$\nu(\mathrm{d}z) = \sum_{k=1}^m \mathbb{1}_{\mathbb{R}_+}(z_k)\frac{\mathrm{d}z_k}{z_k^{1+\vartheta_k}} \otimes \prod_{j\neq k}\delta_0(\mathrm{d}z_j)$$

and for $\xi \geq 0$ it holds that

$$\int_{\mathbb{R}_+^m} \left(1 - \mathrm{e}^{-\xi z_k}\right)\nu(\mathrm{d}z) = \xi^{\vartheta_k} \int_0^\infty \left(1 - \mathrm{e}^{-r}\right)\frac{\mathrm{d}r}{r^{1+\vartheta_k}}.$$

In particular condition (A) is satisfied.

We may also easily find examples where in some directions $b_k > 0$ while for other directions $b_k = 0$ and the Lévy measure ν has sufficiently many jumps (e.g. it is given by previous two examples). Our last example provides a deviation from anisotropic stable Lévy measures.

Example 2.8. Take $\vartheta_k \in (\alpha_k - 1, 1)$, $k = 1, \dots, m$, and let ν be given by

$$\nu(\mathrm{d}z) = \sum_{k=1}^{m} g_k(z_k) \frac{\mathrm{d}z_k}{z_k^{1+\vartheta_k}} \otimes \prod_{j \neq k} \delta_0(\mathrm{d}z_j),$$

where $g_k : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ are bounded. For each $\xi \geq 1$ we obtain

$$\begin{aligned}\int_{\mathbb{R}_+^m} \left(1 - \mathrm{e}^{-\xi z_k}\right) \nu(\mathrm{d}z) &= \xi^{\vartheta_k} \int_0^\infty \left(1 - \mathrm{e}^{-r}\right) g_k\left(\frac{r}{\xi}\right) \frac{\mathrm{d}r}{r^{1+\vartheta_k}} \\ &\geq \xi^{\vartheta_k} \int_0^1 \left(1 - \mathrm{e}^{-r}\right) \frac{\mathrm{d}r}{r^{1+\vartheta_k}} \cdot \inf_{x \in [0,1]} \{g_k(x)\}\end{aligned}$$

Hence condition (A) is satisfied, provided $\inf_{x \in [0,1]}\{g_k(x)\} > 0$ holds for all $k = 1, \dots, m$.

3. Regularity of the heat kernel for the one-dimensional stable JCIR process

In this section we suppose that the conditions of Theorem 2.1 are satisfied. Letting $f(t,u) = \mathrm{Re}(\psi(t,u))$ and $g(t,u) = \mathrm{Im}(\psi(t,u))$, where ψ is obtained from (2.3), we find that $f(t,\mathrm{i}y), g(t,\mathrm{i}y)$ are the unique solutions to

$$\begin{cases} \partial_t f = \beta f + \int_0^\infty \left(\mathrm{e}^{fz}\cos(gz) - 1 - fz\right) \mu_\alpha(\mathrm{d}z), & f(0,\mathrm{i}y) = 0, \\ \partial_t g = \beta g + \int_0^\infty \left(\mathrm{e}^{fz}\sin(gz) - gz\right) \mu_\alpha(\mathrm{d}z), & g(0,\mathrm{i}y) = y. \end{cases}$$

It follows from the general theory of affine processes (see Duffie et al., 2003, Theorem 2.7) that $f \leq 0$. This property will be frequently used. The following is our crucial estimate.

Proposition 3.1. *For each $t_0 > 0$, there exist constants $M, C_1, C_2 > 0$, which depend on t_0, such that for all $|y| \geq M$ and $t \geq t_0$,*

$$b \int_0^t f(s,\mathrm{i}y)\mathrm{d}s + \int_0^t \int_0^\infty \left(\mathrm{e}^{zf(s,\mathrm{i}y)} - 1\right) \nu(\mathrm{d}z) \leq -C_1|y|^{1+\vartheta-\alpha} + C_2. \tag{3.1}$$

Below we first prove Theorem 2.1 and then Proposition 3.1.

Proof of Theorem 2.1: Let $t > 0$ be fixed and choose $t_0 \in (0,t)$. Note that for $u \in \mathbb{R}$, we have $f(t,\mathrm{i}u) \leq 0$ and

$$\begin{aligned}\mathrm{Re}(\phi(t,\mathrm{i}u)) &= b\int_0^t f(s,\mathrm{i}u)ds + \int_0^t \int_0^\infty \left(\mathrm{e}^{zf(s,\mathrm{i}u)}\cos(zg(s,\mathrm{i}u)) - 1\right)\nu(\mathrm{d}z) \\ &\leq b\int_0^t f(s,\mathrm{i}u)ds + \int_0^t \int_0^\infty \left(\mathrm{e}^{zf(s,\mathrm{i}u)} - 1\right)\nu(\mathrm{d}z).\end{aligned} \tag{3.2}$$

By (3.2) and Proposition 3.1, there exist constants $M, C_1, C_2 > 0$ such that for all $|u| \geq M$ and $t \geq t_0$,

$$\begin{aligned}\left|\mathbb{E}[\mathrm{e}^{\mathrm{i}uX^x(t)}]\right| &= \left|\mathrm{e}^{\phi(t,\mathrm{i}u)+x\psi(t,\mathrm{i}u)}\right| \\ &= \mathrm{e}^{\mathrm{Re}(\phi(t,\mathrm{i}u))}\mathrm{e}^{xf(t,\mathrm{i}u)} \\ &\leq \mathrm{e}^{\mathrm{Re}(\phi(t,\mathrm{i}u))} \leq \exp\left\{-C_1|u|^{1+\vartheta-\alpha} + C_2\right\}.\end{aligned} \tag{3.3}$$

Hence

$$\int_{-\infty}^{\infty} |u|^p \left|\mathbb{E}[\mathrm{e}^{\mathrm{i}uX^x(t)}]\right| \mathrm{d}u < \infty$$

for all $p \geq 0$. So $P_t(x, \mathrm{d}y) = p_t(x,y)\mathrm{d}y$, where $p_t(x,y)$ is given by

$$\begin{aligned} p_t(x,y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}yu}\mathbb{E}[\mathrm{e}^{\mathrm{i}uX^x(t)}]\mathrm{d}u \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}yu}\mathrm{e}^{\phi(t,\mathrm{i}u)+x\psi(t,\mathrm{i}u)}\mathrm{d}u. \end{aligned} \tag{3.4}$$

It is clear that the integrand in (3.4) is jointly continuous in $(t,x,y) \in (0,\infty) \times [0,\infty)^2$, and in view of the estimate (3.3) we may apply dominated convergence to find that $p_t(x,y)$ is also jointly continuous in (t,x,y). Using formula (6.16) in the proof of Duffie et al. (2003, Proposition 6.1) we find a constant $C = C_t > 0$ such that $|\psi(t,\mathrm{i}u)| \leq C(1+|u|)$, $u \in \mathbb{R}$. Hence using (3.3) we may differentiate under the integral in (3.4) and find that $(x,y) \longmapsto p_t(x,y)$ is smooth with all derivatives being bounded. The assertion is proved. □

The rest of this section is devoted to the proof of Proposition 3.1. For the proof we use some ideas taken from Filipović et al. (2013). Namely, for $y \in \mathbb{R}$ with $|y| \neq 0$, introduce

$$\begin{cases} F(t,y) := \frac{1}{|y|} f\left(\frac{t}{|y|^{\alpha-1}}, \mathrm{i}y\right), & t \geq 0, \\ G(t,y) := \frac{1}{|y|} g\left(\frac{t}{|y|^{\alpha-1}}, \mathrm{i}y\right), & t \geq 0. \end{cases}$$

Using the substitution $z \longmapsto |y|z$ shows that (F,G) solve

$$\begin{cases} \partial_t F = \beta\frac{F}{|y|^{\alpha-1}} + \int_0^\infty \left(\mathrm{e}^{Fz}\cos(Gz) - 1 - Fz\right)\mu_\alpha(\mathrm{d}z), & F(0,y) = 0, \\ \partial_t G = \beta\frac{G}{|y|^{\alpha-1}} + \int_0^\infty \left(\mathrm{e}^{Fz}\sin(Gz) - Gz\right)\mu_\alpha(\mathrm{d}z), & G(0,y) = \frac{y}{|y|}. \end{cases}$$

We first prove the following lemma.

Lemma 3.2. *There exist constants $t_0, \delta > 0$ and $M > 1$ such that for all $t \in (0, t_0]$ and $|y| \geq M$,*

$$F(t,y) \leq -\delta t.$$

Proof: Note that $\partial_t F(0,y) = \int_0^\infty (\cos(z) - 1)\,\mu_\alpha(\mathrm{d}z) < 0$ and

$$\partial_t G(0,y) = \begin{cases} \frac{\beta}{|y|^{\alpha-1}} + \int_0^\infty (\sin(z) - z)\,\mu_\alpha(\mathrm{d}z), & y > 0, \\ -\frac{\beta}{|y|^{\alpha-1}} - \int_0^\infty (\sin(z) - z)\,\mu_\alpha(\mathrm{d}z), & y < 0. \end{cases}$$

Without loss of generality we suppose $y > 0$, which implies $G(0,y) = 1$.

By continuity, we find $a > 0$ small enough and $M > 0$ large enough such that for all $(F,G) \in D = [-a,0] \times [1-a, 1+a]$ and all $|y| \geq M$,

$$-2\delta \leq \beta\frac{F}{|y|^{\alpha-1}} + \int_0^\infty \left(\mathrm{e}^{Fz}\cos(Gz) - 1 - Fz\right)\mu_\alpha(\mathrm{d}z) \leq -\delta \tag{3.5}$$

and

$$\left|\beta\frac{G}{|y|^{\alpha-1}} + \int_0^\infty \left(\mathrm{e}^{Fz}\sin(Gz) - Gz\right)\mu_\alpha(\mathrm{d}z)\right| \leq K, \tag{3.6}$$

where $\delta, K > 0$ are constants.

Starting from $(0, 1)$, the solution $(F(t, y), G(t, y))$ will stay within D for some positive time, since the velocity vector field is bounded in D. More precisely, let

$$t_0 := \frac{a}{\sqrt{4\delta^2 + K^2}} > 0,$$

then (3.5) and (3.6) imply that for $t \in (0, t_0]$ and $|y| \geq M$,

$$(F(t, y), G(t, y)) \in D$$

and thus

$$F(t, y) = \int_0^t \partial_s F(s, y)\mathrm{d}s \leq -\int_0^t \delta \mathrm{d}s = -\delta t.$$

The lemma is proved. □

We are now prepared to provide a full proof of Proposition 3.1.

Proof of Proposition 3.1: Let $T > 1$ be such that $T^{-1} < t_0 < T$. In the following we first prove that there exist constants $K, C_1, C_2 > 0$ such that for all $|y| \geq K$ and $t \in [T^{-1}, T]$,

$$b\int_0^t f(s, \mathrm{i}y)\mathrm{d}s + \int_0^t \int_0^\infty \left(\mathrm{e}^{zf(s,\mathrm{i}y)} - 1\right) \nu(\mathrm{d}z) \leq -C_1|y|^{1+\vartheta-\alpha} + C_2. \tag{3.7}$$

Define

$$\tilde{\beta} := \begin{cases} \beta, & \text{if } \beta < 0, \\ -1, & \text{if } \beta \geq 0. \end{cases}$$

Using that $\cos(Gz) \leq 1$ combined with

$$\int_0^\infty \left(e^{Fz} - 1 - Fz\right) \mu_\alpha(\mathrm{d}z) = (-F)^\alpha,$$

we find that $F(s, y)$ satisfies

$$\begin{cases} \partial_s F \leq \tilde{\beta}\frac{F}{|y|^{\alpha-1}} + (-F)^\alpha, & s \geq t_1, \\ F(t_1, y) \leq -\rho, \end{cases} \tag{3.8}$$

for all $|y| \geq M > 1$. Here $t_1, \rho, M > 0$ are constants whose existence is guaranteed by Lemma 3.2, and t_1 can actually be made arbitrarily small such that

$$t_1 < T^{-1}. \tag{3.9}$$

Since for $\varkappa \in \mathbb{R}$ the solution to

$$\partial_s \bar{F} = \varkappa\bar{F} + (-\bar{F})^\alpha, \quad \bar{F}(0) = -\rho$$

is given by

$$\bar{F}(s) = -\left(\left(\rho^{1-\alpha} - \varkappa^{-1}\right) e^{-\varkappa(\alpha-1)s} + \varkappa^{-1}\right)^{\frac{1}{1-\alpha}},$$

by comparison theorem for 1-dimensional ODEs, we obtain

$$F(s, y) \leq -\left(\left(\rho^{1-\alpha} - \frac{|y|^{\alpha-1}}{\tilde{\beta}}\right) \exp\left(\frac{\tilde{\beta}(1-\alpha)}{|y|^{\alpha-1}}(s - t_1)\right) + \frac{|y|^{\alpha-1}}{\tilde{\beta}}\right)^{\frac{1}{1-\alpha}}, \quad s \geq t_1.$$

So

$$\begin{aligned} &f(s, \mathrm{i}y) \\ &\quad = |y|F(|y|^{\alpha-1}s, y) \end{aligned}$$

$$\leq -|y|\left(\left(\rho^{1-\alpha}-\frac{|y|^{\alpha-1}}{\tilde\beta}\right)\exp\left(\tilde\beta(1-\alpha)\left(s-\frac{t_1}{|y|^{\alpha-1}}\right)\right)+\frac{|y|^{\alpha-1}}{\tilde\beta}\right)^{\frac{1}{1-\alpha}}, \tag{3.10}$$

whenever $s \geq t_1/|y|^{\alpha-1}$.

"Case 1": Suppose $b>0$. Without loss of generality assume $b=1$. Note that (3.9) holds. For $|y|\geq M>1$ and $t\in[T^{-1},T]$, we have

$$\begin{aligned}
&\int_0^t f(s,\mathrm{i}y)ds\\
&\quad\leq \int_{t_1/|y|^{\alpha-1}}^t f(s,\mathrm{i}y)\mathrm{d}s\\
&\quad\leq -|y|\int_{\frac{t_1}{|y|^{\alpha-1}}}^t \left(\left(\rho^{1-\alpha}-\frac{|y|^{\alpha-1}}{\tilde\beta}\right)e^{\tilde\beta(1-\alpha)\left(s-\frac{t_1}{|y|^{\alpha-1}}\right)}+\frac{|y|^{\alpha-1}}{\tilde\beta}\right)^{\frac{1}{1-\alpha}}\mathrm{d}s.
\end{aligned}$$

Set

$$\eta(s):=\left(\left(\rho^{1-\alpha}-\frac{|y|^{\alpha-1}}{\tilde\beta}\right)\exp\left(\tilde\beta(1-\alpha)\left(s-\frac{t_1}{|y|^{\alpha-1}}\right)\right)+\frac{|y|^{\alpha-1}}{\tilde\beta}\right)^{\frac{1}{1-\alpha}}>0. \tag{3.11}$$

Then

$$\begin{aligned}
\eta'(s)&=\tilde\beta\left(\rho^{1-\alpha}-\frac{|y|^{\alpha-1}}{\tilde\beta}\right)\exp\left(\tilde\beta(1-\alpha)\left(s-\frac{t_1}{|y|^{\alpha-1}}\right)\right)\\
&\qquad\cdot\left(\left(\rho^{1-\alpha}-\frac{|y|^{\alpha-1}}{\tilde\beta}\right)\exp\left(\tilde\beta(1-\alpha)\left(s-\frac{t_1}{|y|^{\alpha-1}}\right)\right)+\frac{|y|^{\alpha-1}}{\tilde\beta}\right)^{\frac{\alpha}{1-\alpha}}\\
&=\tilde\beta\eta(s)^\alpha\left(\eta(s)^{1-\alpha}-\frac{|y|^{\alpha-1}}{\tilde\beta}\right).
\end{aligned}$$

Therefore, substituting $s\to\eta(s)=z$ yields

$$\begin{aligned}
\int_0^t f(s,\mathrm{i}y)\mathrm{d}s&\leq -|y|\int_{\eta(t_1/|y|^{\alpha-1})}^{\eta(t)}\frac{z}{\tilde\beta z^\alpha\left(z^{1-\alpha}-\frac{|y|^{\alpha-1}}{\tilde\beta}\right)}\mathrm{d}z\\
&=-|y|\int_\rho^{\eta(t)}\frac{1}{\left(\tilde\beta-|y|^{\alpha-1}z^{\alpha-1}\right)}\mathrm{d}z\\
&=-\int_{|y|\eta(t)}^{|y|\rho}\frac{1}{z^{\alpha-1}-\tilde\beta}\mathrm{d}z\leq-\frac{|y|\left(\rho-\eta(t)\right)}{\left(|y|\rho\right)^{\alpha-1}-\tilde\beta}.
\end{aligned} \tag{3.12}$$

Note that $\tilde\beta<0$ and the function

$$\begin{aligned}
&\left(\left(\rho^{1-\alpha}-\frac{r^{\alpha-1}}{\tilde\beta}\right)\exp\left(\tilde\beta(1-\alpha)\left(t-\frac{t_1}{r^{\alpha-1}}\right)\right)+\frac{r^{\alpha-1}}{\tilde\beta}\right)^{\frac{1}{1-\alpha}}\\
&\quad=\left(\rho^{1-\alpha}e^{\tilde\beta(1-\alpha)\left(t-\frac{t_1}{r^{\alpha-1}}\right)}-\frac{r^{\alpha-1}}{\tilde\beta}\left(e^{\tilde\beta(1-\alpha)\left(t-\frac{t_1}{r^{\alpha-1}}\right)}-1\right)\right)^{\frac{1}{1-\alpha}}
\end{aligned}$$

is monotone increasing in $r\in(0,\infty)$. Therefore, for $|y|\geq M$ and $t\in[T^{-1},T]$, we obtain

$$\rho\geq\rho-\eta(t)$$

$$
\begin{aligned}
&\geq \rho - \left(\left(\rho^{1-\alpha} - \frac{M^{\alpha-1}}{\tilde{\beta}} \right) \exp\left(\tilde{\beta}(1-\alpha)\left(t - \frac{t_1}{M^{\alpha-1}} \right) \right) + \frac{M^{\alpha-1}}{\tilde{\beta}} \right)^{\frac{1}{1-\alpha}} \\
&\geq \rho - \left(\left(\rho^{1-\alpha} - \frac{M^{\alpha-1}}{\tilde{\beta}} \right) \exp\left(\tilde{\beta}(1-\alpha)\left(T^{-1} - \frac{t_1}{M^{\alpha-1}} \right) \right) + \frac{M^{\alpha-1}}{\tilde{\beta}} \right)^{\frac{1}{1-\alpha}} \\
&=: c_1 > 0. \qquad (3.13)
\end{aligned}
$$

Combining (3.12) and (3.13) gives

$$\int_0^t f(s, \mathrm{i}y)\mathrm{d}s \leq -\frac{c_1|y|}{(|y|\rho)^{\alpha-1} - \tilde{\beta}}.$$

Obviously, we can choose a larger $M' > M$ such that

$$\int_0^t f(s, \mathrm{i}y)\mathrm{d}s \leq -\frac{c_1|y|}{2\,(|y|\rho)^{\alpha-1}} \leq -c_2|y|^{2-\alpha}, \quad \forall\ |y| \geq M',\ t \in [T^{-1}, T].$$

"Case 2": Suppose $b = 0$. Similarly as in Case 1, by (2.1), (3.2) and (3.10), we obtain, for $|y| \geq M > 1$ and $t \in [T^{-1}, T]$,

$$
\begin{aligned}
&\int_0^t \int_0^\infty \left(e^{zf(s,\mathrm{i}y)} - 1 \right) \nu(\mathrm{d}z)\mathrm{d}s \\
&\quad = -\int_0^t \int_0^\infty \left(1 - e^{zf(s,\mathrm{i}y)} \right) \nu(\mathrm{d}z)\mathrm{d}s \\
&\quad \leq -\int_{\frac{t_1}{|y|^{\alpha-1}}}^t \left[c_3 \left(-f(s, \mathrm{i}y) \right)^{\vartheta} - c_4 \right] \mathrm{d}s \\
&\quad \leq c_4 t - c_3|y|^{\vartheta} \int_{\frac{t_1}{|y|^{\alpha-1}}}^t \left(\left(\rho^{1-\alpha} - \frac{|y|^{\alpha-1}}{\tilde{\beta}} \right) e^{\tilde{\beta}(1-\alpha)\left(s - \frac{t_1}{|y|^{\alpha-1}}\right)} + \frac{|y|^{\alpha-1}}{\tilde{\beta}} \right)^{\frac{\vartheta}{1-\alpha}} \mathrm{d}s,
\end{aligned}
$$

where we have used (2.1) so that $\int_0^\infty (1 - \mathrm{e}^{-\xi z})\nu(\mathrm{d}z) \geq c_3\xi^{\vartheta} - c_4$ for all $\xi \geq 0$ and some constants $c_3, c_4 > 0$. Using again the change of variables $z = \eta(s)$, where η is defined in (3.11), we get

$$
\begin{aligned}
&\int_0^t \int_0^\infty \left(e^{zf(s,\mathrm{i}y)} - 1 \right) \nu(\mathrm{d}z)\mathrm{d}s \\
&\quad \leq c_4 t - c_3 \int_{|y|\eta(t)}^{|y|\rho} \frac{z^{\vartheta-1}}{z^{\alpha-1} - \tilde{\beta}} \mathrm{d}z \\
&\quad \leq c_4 T - \frac{c_3|y|\,(|y|\rho)^{\vartheta-1}}{(|y|\rho)^{\alpha-1} - \tilde{\beta}} \left(\rho - \eta(t) \right) \\
&\quad \overset{(3.13)}{\leq} c_4 T - \frac{c_1 c_3 |y|\,(|y|\rho)^{\vartheta-1}}{(|y|\rho)^{\alpha-1} - \tilde{\beta}} \leq -c_5|y|^{1+\vartheta-\alpha} + c_6, \quad \forall\ |y| \geq M'',\ t \in [T^{-1}, T],
\end{aligned}
$$

where $M'' > M$ is another large enough constant.

Summarizing Case 1 and 2, and noting that $2 - \alpha \geq 1 + \vartheta - \alpha$, we obtain (3.7). Now, for $t > T$ and $|y| \geq K$, it holds also that

$$
\begin{aligned}
&b\int_0^t f(s, \mathrm{i}y)\mathrm{d}s + \int_0^t \int_0^\infty \left(\mathrm{e}^{zf(s,\mathrm{i}y)} - 1 \right) \nu(\mathrm{d}z) \\
&\quad \leq b\int_0^T f(s, \mathrm{i}y)\mathrm{d}s + \int_0^T \int_0^\infty \left(\mathrm{e}^{zf(s,\mathrm{i}y)} - 1 \right) \nu(\mathrm{d}z)
\end{aligned}
$$

$$\leq -C_1|y|^{1+\vartheta-\alpha} + C_2.$$

The proposition is proved. □

4. Existence and regularity of the heat kernel

4.1. *Heat kernel and anisotropic Besov regularity.* In order to measure anisotropic smoothness related to the cylindrical Lévy process $Z = (Z_1, \dots, Z_m)$, we use an anisotropic analogue of classical Besov spaces. Corresponding to the regularity indices $(\alpha_1, \dots, \alpha_m)$ we define a mean order of smoothness $\overline{\alpha} > 0$ and an anisotropy $a = (a_1, \dots, a_m)$ by

$$\frac{1}{\overline{\alpha}} = \frac{1}{m}\left(\frac{1}{\alpha_1} + \cdots + \frac{1}{\alpha_m}\right), \qquad a_i = \frac{\overline{\alpha}}{\alpha_i}, \quad i = 1, \dots, m. \tag{4.1}$$

Then note that $0 < a_1, \dots, a_m < \infty$ and $a_1 + \cdots + a_m = m$. Take $\lambda > 0$ with $\lambda / a_k \in (0,1)$ for all $k \in \{1, \dots, m\}$. For a measurable function $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ introduce

$$\|f\|_{B^{\lambda,a}_{1,\infty}} := \|f\|_{L^1(\mathbb{R}^m)} + \sum_{k=1}^{m} \sup_{h \in [-1,1]} |h|^{-\lambda/a_k} \|\Delta_{he_k} f\|_{L^1(\mathbb{R}^m)}, \tag{4.2}$$

where $\Delta_h f(x) = f(x+h) - f(x)$, $h \in \mathbb{R}^m$. The anisotropic Besov space $B^{\lambda,a}_{1,\infty}(\mathbb{R}^m)$ is defined as the set of all $L^1(\mathbb{R}^m)$ functions f with $\|f\|_{B^{\lambda,a}_{1,\infty}} < \infty$ (see Dachkovski, 2003 and Triebel, 2006 for additional details and references). By studying estimates on the heat kernel weighted by

$$\rho_\delta(x) = \min\{\delta, x_1^{1/\alpha_1}, \dots, x_m^{1/\alpha_m}\} \mathbb{1}_{\mathbb{R}^m_+}(x), \qquad \delta \in (0,1],$$

we can also take the behavior of the process at the boundary into account. The following is our main result for the regularity of the heat kernel.

Theorem 4.1. *Suppose that condition (A) is satisfied and assume there exists $\tau > 0$ satisfying*

$$\int_{\mathbb{R}^m_+} \mathbb{1}_{\{|z|>1\}} |z|^{1+\tau} \nu(\mathrm{d}z) < \infty.$$

Then for each $t > 0$ and each $x \in \mathbb{R}^m_+$ the transition kernel $P_t(x, \mathrm{d}y)$ has density $p_t(x,y)$ with respect to the Lebesgue measure. Moreover, there exists some small constant $\lambda > 0$ such that for each $T > 0$, $\varkappa \in (0,1]$ and $\delta \in (0,1]$,

$$\|p_t^\delta(x,\cdot)\|_{B^{\lambda,a}_{1,\infty}(\mathbb{R}^m_+)} \leq C(1+|x|)^{\varkappa}(1 \wedge t)^{-1/\alpha_{\min}}, \quad t \in (0,T], \; x \in \mathbb{R}^m_+, \tag{4.3}$$

where $p_t^\delta(x,y) := \rho_\delta(y) p_t(x,y)$ and $C = C(\lambda, \tau, \varkappa, \delta, T) > 0$ is a constant.

The proof of this result follows the arguments given in Friesen et al. (2020) and Friesen et al. (2020+a) where general stochastic equations have been considered. Since we need the precise dependence on x in (4.3) and since the proofs are significantly simpler for (1.3) compared with the general case, we provide, for convenience of the reader, a full proof of Proposition 4.1 in the appendix.

4.2. *Boundary non-attainment.* In this section we prove the following estimate on the behavior of the anisotropic stable JCIR process at the boundary.

Proposition 4.2. *Suppose that condition (A) is satisfied. Then for each $t > 0$ and each $R > 0$, there exists $C > 0$ such that*

$$\mathbb{P}[\min\{X_1^x(t), \dots, X_m^x(t)\} \le \varepsilon] \le C\varepsilon, \qquad \varepsilon \in (0,1),\ x \in \mathbb{R}_+^m,\ |x| \le R.$$

In particular, $\mathbb{P}[X^x(t) \in \mathbb{R}_{++}^m] = 1$ holds for all $t > 0$ and all $x \in \mathbb{R}_+^m$.

Proof: Consider $x = (x_1, \dots, x_m) \in \mathbb{R}_+^m$ and let $(X^x(t))_{t \ge 0}$ be the anisotropic stable JCIR process obtained from (1.3). Moreover, let $Y^x(t) = (Y_1^{x_1}(t), \dots, Y_m^{x_m}(t))$ be the unique $\mathbb{R}_+^m$-valued strong solution to

$$\mathrm{d}Y_k^{x_k}(t) = (b_k + \beta_{kk} Y_k^{x_k}(t))\,\mathrm{d}t + \sqrt[\alpha_k]{\sigma_k Y_k^{x_k}(t)}\mathrm{d}Z_k(t) + \mathrm{d}J_k(t), \tag{4.4}$$

where $k \in \{1, \dots, d\}$ and $Y_k^{x_k}(0) = x_k$. Existence and uniqueness of such a process is again a direct consequence of Barczy et al. (2015), see also Fu and Li (2010). Moreover, $(Y^x(t))_{t \ge 0}$ is the anisotropic JCIR process where all off-diagonal drift terms equal to zero. Using the fact that $\beta_{kj} \ge 0$ whenever $k \ne j$ we may apply the comparison result established in Friesen et al. (2019b, Proposition 4.2) to deduce

$$\mathbb{P}[X_k^x(t) \ge Y_k^{x_k}(t),\ \ t \ge 0] = 1, \qquad k \in \{1, \dots, m\}.$$

Since $(J_k(t))_{t \ge 0}$ is a Lévy subordinator on $\mathbb{R}_+$ whose Lévy measure is given by $\nu_k = \nu \circ \mathrm{pr}_k^{-1}$, where $\mathrm{pr}_k(z) = z_k$ denotes the projection on the k-th coordinate, we can apply Theorem 2.1 for the process $(Y_k^{x_k}(t))_{t \ge 0}$. Let $p_t^k(x_k, y_k)$ be its heat kernel. Then

$$\begin{aligned} \mathbb{P}[\min\{X_1^x(t), \dots, X_m^x(t)\} \le \varepsilon] &\le \sum_{k=1}^m \mathbb{P}[X_k^x(t) \le \varepsilon] \\ &\le \sum_{k=1}^m \mathbb{P}[Y_k^{x_k}(t) \le \varepsilon] \\ &= \sum_{k=1}^m \int_0^\varepsilon p_t^k(x_k, y_k)\mathrm{d}y_k \le C\varepsilon, \end{aligned}$$

since $p_t^k(x_k, y_k)$ is jointly continuous in (x_k, y_k) and $|x| \le R$. This proves the assertion. □

4.3. *Proof of Theorem 2.4.* Let us first prove a slightly weaker assertion.

Lemma 4.3. *Assume the same assumptions as in Theorem 4.1. Then the mapping $\mathbb{R}_+^m \ni x \longmapsto p_t^\delta(x, \cdot) \in L^1(\mathbb{R}_+^m)$ is continuous for each $t > 0$ and each $\delta \in (0,1]$, where $p_t^\delta(x, y) = \rho_\delta(y) p_t(x, y)$.*

Proof: Using the Feller property (see Duffie et al., 2003, Proposition 8.2), we find that $x \longmapsto p_t(x, y)\mathrm{d}y$ is weakly continuous, i.e.

$$x \longmapsto \int_{\mathbb{R}_+^m} f(y) \rho_\delta(y) p_t(x, y)\mathrm{d}y \tag{4.5}$$

is continuous for each bounded continuous function f.

Let $x \in \mathbb{R}_+^m$ be fixed. Suppose $(x_n)_n$ is a sequence such that $x_n \to x$. Using (4.3) we find that

$$\sup_{n\in\mathbb{N}} \|p_t^\delta(x_n,\cdot)\|_{B_{1,\infty}^{\lambda,a}} < \infty.$$

Next observe that for each $K, R > 0$,

$$\sup_{|x|\le K} \int_{|y|\ge R} p_t^\delta(x,y)\mathrm{d}y \le \delta \sup_{|x|\le K} \int_{|y|\ge R} p_t(x,y)\mathrm{d}y \le \frac{\delta}{R} \sup_{|x|\le K} \mathbb{E}[|X_t^x|] \le \frac{\delta K C_t}{R},$$

where we have used Proposition A.1. Hence we may apply the Kolmogorov-Riesz compactness criterion which gives existence of a subsequence $(x_{n_k})_k$ such that $p_t^\delta(x_{n_k},\cdot)$ has a limit in $L^1(\mathbb{R}_+^m)$. By the weak continuity (4.5) this limit is exactly $p_t^\delta(x,\cdot)$.

So for each sequence $(x_n)_n$ with $x_n \to x$ we have found a subsequence $(x_{n_k})_k$ such that $p_t^\delta(x_{n_k},\cdot) \to p_t^\delta(x,\cdot)$ in L^1 as $k\to\infty$. This proves the desired L^1 continuity. □

We are now prepared to provide a full proof of Theorem 2.4. The affine structure of the anisotropic JCIR process allows us to study first the particular case with ν having no big jumps, i.e., $\nu(\mathrm{d}z) = \mathbb{1}_{\{|z|\le 1\}}\nu(\mathrm{d}z)$, and then the general case by a convolution argument. Namely, define $\nu_0(\mathrm{d}z) = \mathbb{1}_{\{|z|\le 1\}}\nu(\mathrm{d}z)$ and $\nu_1(\mathrm{d}z) = \mathbb{1}_{\{|z|>1\}}\nu(\mathrm{d}z)$, and write $F(u) = \langle b,u\rangle + F_0(u) + F_1(u)$ where for $i = 0,1$,

$$F_i(u) = \int_{\mathbb{R}_+^m} \left(\mathrm{e}^{\langle u,z\rangle} - 1\right)\nu_i(\mathrm{d}z), \qquad u\in\mathbb{C}^m \text{ with } \mathrm{Re}(u)\le 0.$$

Let $(Y^x(t))_{t\ge0}$ the unique strong solution to (1.3) with $\nu=\nu_0$ and let $(\widetilde{Y}^x(t))_{t\ge0}$ be the unique strong solution to (1.3) with $b=0$ and $\nu=\nu_1$, i.e.,

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u,Y^x(t)\rangle}\right] = \exp\left(\int_0^t \left(\langle b,\psi(s,\mathrm{i}u)\rangle + F_0(\psi(s,\mathrm{i}u))\right)\mathrm{d}s + \langle x,\psi(s,\mathrm{i}u)\rangle\right),$$
$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u,\widetilde{Y}^x(t)\rangle}\right] = \exp\left(\int_0^t F_1(\psi(s,\mathrm{i}u))\mathrm{d}s + \langle x,\psi(s,\mathrm{i}u)\rangle\right),$$

where ψ is obtained from (1.5). Denote by $Q_t^0(x,\cdot)$ the transition probabilities of $(Y^x(t))_{t\ge0}$ and by $Q_t^1(x,\cdot)$ the transition probabilities of $(\widetilde{Y}^x(t))_{t\ge0}$. Using (1.4) we find

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u,Y^x(t)\rangle}\right]\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u,\widetilde{Y}^0(t)\rangle}\right]\\
&= \exp\left(\int_0^t \left(\langle b,\psi(s,\mathrm{i}u)\rangle + F_0(\psi(s,\mathrm{i}u))\right)\mathrm{d}s + \langle x,\psi(s,\mathrm{i}u)\rangle\right)\exp\left(\int_0^t F_1(\psi(s,\mathrm{i}u))\mathrm{d}s\right)\\
&= \exp\left(\phi(t,\mathrm{i}u) + \langle x,\psi(t,\mathrm{i}u)\rangle\right)\\
&= \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u,X^x(t)\rangle}\right]
\end{aligned}$$

which yields

$$P_t(x,\cdot) = Q_t^0(x,\cdot) * Q_t^1(0,\cdot), \qquad t>0,\quad x\in\mathbb{R}_+^m, \tag{4.6}$$

where $*$ denotes the convolution of measures.

Proof of Theorem 2.4: According to Theorem 4.1, the kernel $Q_t^0(x,\cdot)$ has a density $q_t^0(x,\cdot)$ for $t>0$. In view of (4.6), for $t>0$ and $x\in\mathbb{R}_+^m$, the measure $P_t(x,\cdot)$

possesses also a density $p_t(x,\cdot)$ with respect to the Lebesgue measure and it is given by

$$p_t(x,y) = \int_{\mathbb{R}_+^m} q_t^0(x, y-z) Q_t^1(0,\mathrm{d}z), \quad y \in \mathbb{R}_+^m,$$

and $p_t(x,y) = 0$ for $y \notin \mathbb{R}_+^m$.

Fix $t > 0$ and $x \in \mathbb{R}_+^m$. Let $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}_+^m$ be such that $x_n \to x$. Our aim is to show that

$$p_t(x_n,\cdot) \to p_t(x,\cdot) \quad \text{in} \quad L^1\left(\mathbb{R}_+^m\right), \quad \text{as} \quad n\to\infty. \tag{4.7}$$

We will finish the proof in two steps:

"Step 1": We show that $q_t^0(x_n,\cdot)$ converges in $L^1\left(\mathbb{R}_+^m\right)$ to $q_t^0(x,\cdot)$ as $n\to\infty$. Let $\varepsilon \in (0,1)$. Write $\widetilde{\rho}(z) = \min\{z_1^{1/\alpha_1},\dots,z_m^{1/\alpha_m}\}$ so that $\rho_\varepsilon(z) = \min\{\varepsilon, \widetilde{\rho}(z)\}$. Then

$$\begin{aligned}
&\|q_t^0(x_n,\cdot) - q_t^0(x,\cdot)\|_{L^1\left(\mathbb{R}_+^m\right)} \\
&\le \int_{\mathbb{R}_+^m} \left(q_t^0(x_n,y) + q_t^0(x,y)\right) \mathbb{1}_{\{\widetilde{\rho}(y)\le\varepsilon\}}\mathrm{d}y \\
&\quad + \int_{\mathbb{R}_+^m} \left|q_t^0(x_n,y) - q_t^0(x,y)\right| \mathbb{1}_{\{\widetilde{\rho}(y)>\varepsilon\}}\mathrm{d}y \\
&= \mathbb{P}[\widetilde{\rho}(Y^{x_n}(t)) \le \varepsilon] + \mathbb{P}[\widetilde{\rho}(Y^{x}(t)) \le \varepsilon] \\
&\quad + \int_{\mathbb{R}_+^m} \left|q_t^0(x_n,y) - q_t^0(x,y)\right| \mathbb{1}_{\{\widetilde{\rho}(y)>\varepsilon\}}\mathrm{d}y.
\end{aligned}$$

Using Proposition 4.2 we can have

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \mathbb{P}[\widetilde{\rho}(Y^{x_n}(t)) \le \varepsilon] + \mathbb{P}[\widetilde{\rho}(Y^{x}(t)) \le \varepsilon] \\
&\quad \le \sup_{n\in\mathbb{N}} \mathbb{P}\left[\bigcup_{i=1}^m \left\{(Y_i^{x_n}(t))^{1/\alpha_i} \le \varepsilon\right\}\right] + \mathbb{P}\left[\bigcup_{i=1}^m \left\{(Y_i^{x}(t))^{1/\alpha_i} \le \varepsilon\right\}\right] \\
&\quad \le \sup_{n\in\mathbb{N}} m\mathbb{P}\left[\min\left\{Y_1^{x_n}(t),\dots,Y_m^{x_n}(t)\right\} \le \varepsilon^{\alpha_{\min}}\right] \\
&\qquad + m\mathbb{P}\left[\min\left\{Y_1^{x}(t),\dots,Y_m^{x}(t)\right\} \le \varepsilon^{\alpha_{\min}}\right] \\
&\quad \le C\varepsilon^{\alpha_{\min}} \le C\varepsilon.
\end{aligned}$$

For the third term we use

$$\begin{aligned}
&\int_{\mathbb{R}_+^m} \left|q_t^0(x_n,y) - q_t^0(x,y)\right| \mathbb{1}_{\{\widetilde{\rho}(y)>\varepsilon\}}\mathrm{d}y \\
&\quad = \varepsilon^{-1}\int_{\mathbb{R}_+^m} \left|\rho_\varepsilon(y) q_t^0(x_n,y) - \rho_\varepsilon(y)) q_t^0(x,y)\right| \mathbb{1}_{\{\widetilde{\rho}(y)>\varepsilon\}}\mathrm{d}y \\
&\quad \le \varepsilon^{-1}\int_{\mathbb{R}_+^m} \left|\rho_\varepsilon(y) q_t^0(x_n,y) - \rho_\varepsilon(y)) q_t^0(x,y)\right| \mathrm{d}y,
\end{aligned}$$

where the right-hand side tends by Lemma 4.3 to zero as $n\to\infty$. So

$$\limsup_{n\to\infty} \|q_t^0(x_n,\cdot) - q_t^0(x,\cdot)\|_{L^1\left(\mathbb{R}_+^m\right)} \le C\varepsilon.$$

Since $\varepsilon \in (0,1)$ is arbitrary, the desired convergence $q_t^0(x_n,\cdot) \to q_t^0(x,\cdot)$ in $L^1\left(\mathbb{R}_+^m\right)$ is proved.

"Step 2": We show that (4.7) is true in the general case. We have

$$\begin{aligned}
&\int_{\mathbb{R}_+^m} |p_t(x_n, y) - p_t(x, y)| \, \mathrm{d}y \\
&= \int_{\mathbb{R}^m} |p_t(x_n, y) - p_t(x, y)| \, \mathrm{d}y \\
&\leq \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \left| q_t^0(x_n, y - z) - q_t^0(x, y - z) \right| Q_t^1(0, \mathrm{d}z) \mathrm{d}y \\
&= \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \left| q_t^0(x_n, y - z) - q_t^0(x, y - z) \right| \mathrm{d}y Q_t^1(0, \mathrm{d}z) \\
&= \| q_t^0(x_n, \cdot) - q_t^0(x, \cdot) \|_{L^1(\mathbb{R}_+^m)} \to 0, \quad \text{as} \quad n \to \infty.
\end{aligned}$$

So (4.7) is true. The theorem is proved. □

5. Exponential ergodicity in total variation

5.1. *Regularity for the invariant measure.* As a consequence of our regularity results for the heat kernel (see Sections 3 and 4), we can also deduce similar results for the invariant measure. We start with the one-dimensional case.

Corollary 5.1. *Suppose that (2.1) is satisfied. Assume that $\beta < 0$ and $\int_{(1,\infty)} \log(1+z)\nu(\mathrm{d}z) < \infty$. Then the unique invariant measure $\pi(\mathrm{d}x)$ has a smooth density g which satisfies $g(x) = 0$ for all $x \leq 0$ and vanishes at infinity.*

Proof: It was shown in Keller-Ressel and Mijatović (2012), see also Jin et al. (2020), that for all $y \in \mathbb{R}$,

$$\lim_{t \to \infty} \mathrm{e}^{\phi(t, \mathrm{i}y) + x\psi(t, \mathrm{i}y)} = \mathrm{e}^{\phi(\infty, \mathrm{i}y)} = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}xy} \pi(\mathrm{d}x),$$

where

$$\phi(\infty, \mathrm{i}y) = \int_0^\infty \left(b\psi(s, \mathrm{i}y) + \int_{(0,\infty)} \left(\mathrm{e}^{z\psi(s,\mathrm{i}y)} - 1 \right) \nu(\mathrm{d}z) \right) \mathrm{d}s$$

and the integral against $\mathrm{d}s$ is absolutely convergent. Since the process is supported on $\mathbb{R}_+$ it is clear that $\pi((-\infty, 0)) = 0$. Using Proposition 3.1 we may take the limit $t \to \infty$ in (3.1) and find for $t_0 = 1$ constants $M, C > 0$ such that

$$\left| \mathrm{e}^{\phi(\infty, \mathrm{i}y)} \right| \leq \mathrm{e}^{-C|y|^{1+\vartheta-\alpha}}, \qquad |y| \geq M.$$

The assertion follows from classical properties of the Fourier transform. □

In the multi-dimensional case we may use (4.3) to deduce the same regularity for the unique invariant measure π.

Corollary 5.2. *Suppose that the anisotropic stable JCIR process is subcritical, satisfies condition (A) and ν satisfying*

$$\int_{\mathbb{R}_+^m} \mathbb{1}_{\{|z|>1\}} |z|^{1+\tau} \nu(\mathrm{d}z) < \infty$$

for some $\tau > 0$. Then π is absolutely continuous with respect to the Lebesgue measure, i.e., $\pi(\mathrm{d}x) = g(x)\mathrm{d}x$ and there exists a constant $C > 0$ such that

$$\|g^1\|_{B_{1,\infty}^{\lambda,a}} \leq C \int_{\mathbb{R}_+^m} (1 + |x|) \pi(\mathrm{d}x) < \infty,$$

where $g^1(x) = \rho_1(x)g(x)$ *and* λ, a *are given as in Theorem 4.1.*

Proof: Using the invariance of π we get

$$\pi(\mathrm{d}x) = \int_{\mathbb{R}_+^m} P_t(y, \mathrm{d}x)\pi(\mathrm{d}y) = \left(\int_{\mathbb{R}_+^m} p_t(y,x)\pi(\mathrm{d}y)\right)\mathrm{d}x,$$

i.e., $\pi(\mathrm{d}x)$ has a density $g(x)$ which satisfies

$$g(x) = \int_{\mathbb{R}_+^m} p_t(y,x)\pi(\mathrm{d}y).$$

Hence we obtain, for $p_t^1(x,y) = \rho_1(y)p_t(x,y)$,

$$\begin{aligned}\|\pi^1\|_{B_{1,\infty}^{\lambda,a}} &\leq \int_{\mathbb{R}_+^m} \|p_t^1(y,\cdot)\|_{B_{1,\infty}^{\lambda,a}}\pi(\mathrm{d}y) \\ &\leq C\int_{\mathbb{R}_+^m}(1+|y|)\pi(\mathrm{d}y).\end{aligned}$$

Using Proposition A.2 combined with the weak convergence $p_t(x,y)\mathrm{d}y \longrightarrow \pi(\mathrm{d}y)$ we find that $\int_{\mathbb{R}_+^m}|y|\pi(\mathrm{d}y) < \infty$. The assertion is proved. □

5.2. *Proof of Theorem 2.5.* Here and below we suppose that the conditions of Theorem 2.5 are satisfied. As in Section 4.3, let $\nu_0(\mathrm{d}z) = \mathbb{1}_{\{|z|\leq 1\}}\nu(\mathrm{d}z)$, $\nu_1(\mathrm{d}z) = \mathbb{1}_{\{|z|>1\}}\nu(\mathrm{d}z)$ and let $(Y^x(t))_{t\geq 0}$ be the solution to (1.3) with $\nu = \nu_0$ and $(\widetilde{Y}^x(t))_{t\geq 0}$ be the solution to (1.3) with $b = 0$ and $\nu = \nu_1$. Denote by $Q_t^0(x,\cdot)$ the transition probabilities of $(Y^x(t))_{t\geq 0}$ and by $Q_t^1(x,\cdot)$ the transition probabilities of $(\widetilde{Y}^x(t))_{t\geq 0}$. Recall that (4.6) holds. In order to prove Theorem 2.5 we first establish a similar statement for $(Y^x(t))_{t\geq 0}$.

Proposition 5.3. *There exists constants* $C, \delta > 0$ *such that*

$$\|Q_t^0(x,\cdot) - Q_t^0(y,\cdot)\|_{\mathrm{TV}} \leq C\min\left\{1, (1+|x|+|y|)\mathrm{e}^{-\delta t}\right\}, \tag{5.1}$$

for all $x, y \in \mathbb{R}_+^m$ *and all* $t \geq 0$.

Proof: Following Jin et al. (2020, Lemma 3.4) we define a new norm by

$$|x|_M = \langle x,x\rangle_M^{1/2} = \langle x, Mx\rangle^{1/2} \quad \text{where } M = \int_0^\infty \mathrm{e}^{t\beta^\top}\mathrm{e}^{t\beta}\mathrm{d}t. \tag{5.2}$$

Then note that M is symmetric, positive definite and satisfies $M\beta + \beta^\top M = -\mathbb{1}$. In view of Theorem D.1 it suffices to show that the following two properties are satisfied.

(i) The function $V(x) = (1+|x|_M^2)^{1/2}$ belongs to the domain of the extended generator and there exist constants $c_1, c_2 > 0$ such that

$$L_0V(x) \leq -c_1V(x) + c_2, \qquad x \in \mathbb{R}_+^m,$$

where L_0 denotes the extended generator of $(Y^x(t))_{t\geq 0}$.

(ii) For every $R > 0$ there exist $h > 0$ and $\delta \in (0,2)$ with

$$\|Q_h^0(x,\cdot) - Q_h^0(y,\cdot)\|_{\mathrm{TV}} \leq 2-\delta,$$

for all $x, y \in \mathbb{R}_+^m$ with $|x|, |y| \leq R$.

Property (i) can be shown by similar (but essentially simpler) arguments to Jin et al. (2020, Lemma 3.4 and Proposition 3.7). For the sake of completeness a proof is outlined in the appendix, see Lemma B.1. Let us now prove property (ii). Let $R > 0$ and take any $x, y \in \mathbb{R}_+^m$ with $|x|,\ |y| \le R$. Fix any bounded measurable function f on $\mathbb{R}_+^m$ satisfying $\|f\|_\infty \le 1$. Choose $h > 1$ and $\widetilde{R} > 0$ to be specified later on. Let $H = H_{h,x,y}$ be the joint distribution of $(Y^x(h-1), Y^y(h-1))$, i.e. $H(\mathrm{d}\widetilde{x}, \mathrm{d}\widetilde{y}) = \mathbb{P}[Y^x(h-1) \in \mathrm{d}\widetilde{x},\ Y^y(h-1) \in \mathrm{d}\widetilde{y}]$. Since $(Q_t^0)_{t\ge 0}$ satisfies the conditions of Theorem 2.4, $Q_t^0(x,\cdot)$ has density $q_t^0(x,\cdot)$ being continuous in x with respect to $L^1(\mathbb{R}^m)$. Hence we find $\eta(\widetilde{R}) > 0$ (independent of f) such that $\|q_1^0(\widetilde{x},\cdot) - q_1^0(\widetilde{y},\cdot)\|_{L^1(\mathbb{R}_+^m)} \le 1$ for $(\widetilde{x},\widetilde{y}) \in \Theta$, where

$$\Theta = \{(\widetilde{x},\widetilde{y}) \in \mathbb{R}_+^m \times \mathbb{R}_+^m \mid |\widetilde{x} - \widetilde{y}| \le \eta(\widetilde{R}),\ |\widetilde{x}|,\ |\widetilde{y}| \le \widetilde{R}\}.$$

It follows easily that $|Q_1^0 f(\widetilde{x}) - Q_1^0 f(\widetilde{y})| \le 1$ for $(\widetilde{x},\widetilde{y}) \in \Theta$. Then we obtain from $\|Q_1^0 f\|_\infty \le 1$

$$\begin{aligned}
|Q_h^0 f(x) - Q_h^0 f(y)| &\le \left| \int_\Theta (Q_1^0 f(\widetilde{x}) - Q_1^0 f(\widetilde{y})) H(\mathrm{d}\widetilde{x}, \mathrm{d}\widetilde{y}) \right| \\
&\quad + \left| \int_{\Theta^c} (Q_1^0 f(\widetilde{x}) - Q_1^0 f(\widetilde{y})) H(\mathrm{d}\widetilde{x}, \mathrm{d}\widetilde{y}) \right| \\
&\le H(\Theta) + 2H(\Theta^c) \\
&= 2 - H(\Theta).
\end{aligned}$$

Next we obtain

$$\begin{aligned}
H(\Theta^c) &\le \mathbb{P}[|Y^x(h-1) - Y^y(h-1)| > \eta(\widetilde{R})] + \mathbb{P}[|Y^x(h-1)| > \widetilde{R}] \\
&\quad + \mathbb{P}[|Y^y(h-1)| > \widetilde{R}] \\
&\le \eta(\widetilde{R})^{-1} \mathbb{E}[|Y^x(h-1) - Y^y(h-1)|] + \widetilde{R}^{-1} \mathbb{E}\left[|Y^x(h-1)| + |Y^y(h-1)|\right] \\
&\le \frac{m}{\eta(\widetilde{R})} |x - y| \mathrm{e}^{-c(h-1)} + \frac{C(1 + |x| + |y|)}{\widetilde{R}} \\
&\le \frac{2mR\mathrm{e}^c}{\eta(\widetilde{R})} \mathrm{e}^{-ch} + \frac{C(1+2R)}{\widetilde{R}},
\end{aligned}$$

where $c > 0$ and $C > 0$ are some constants given by Friesen et al. (2020+b, Proposition 6.1) and Proposition A.2. Take first $\widetilde{R} > 0$ and then $h > 1$ large enough such that

$$\frac{C(1+2R)}{\widetilde{R}} < \frac{1}{2} \quad \text{and} \quad \frac{2mR\mathrm{e}^c}{\eta(\widetilde{R})} \mathrm{e}^{-ch} < \frac{1}{2}.$$

Then

$$H(\Theta) = 1 - H(\Theta^c) \ge 1 - \frac{2mR\mathrm{e}^c}{\eta(\widetilde{R})} \mathrm{e}^{-ch} - \frac{C(1+2R)}{\widetilde{R}} =: \delta \in (0,1).$$

This proves the assertion. □

We are now prepared to give a proof for Theorem 2.5.

Proof of Theorem 2.5: Let π be the unique invariant measure and let H be a coupling of δ_x and π, i.e., a Borel probability measure over $\mathbb{R}_+^m \times \mathbb{R}_+^m$ whose marginals are δ_x and π, respectively. Then

$$\begin{aligned}\|P_t(x,\cdot)-\pi\|_{\mathrm{TV}} &\leq \int_{\mathbb{R}_+^m\times\mathbb{R}_+^m}\|P_t(y,\cdot)-P_t(\widetilde{y},\cdot)\|_{\mathrm{TV}}H(\mathrm{d}y,\mathrm{d}\widetilde{y})\\ &\leq \int_{\mathbb{R}_+^m\times\mathbb{R}_+^m}\|Q_t^0(y,\cdot)-Q_t^0(\widetilde{y},\cdot)\|_{\mathrm{TV}}H(\mathrm{d}y,\mathrm{d}\widetilde{y})\\ &\leq C\int_{\mathbb{R}_+^m\times\mathbb{R}_+^m}\min\left\{1,(1+|y|+|\widetilde{y}|)e^{-\delta t}\right\}H(\mathrm{d}y,\mathrm{d}\widetilde{y}),\end{aligned}$$

where we have used Friesen et al. (2019a, Lemma 2.3) and then (5.1) for the integrand. Note that for $a,b\geq 0$,

$$\begin{aligned}1\wedge(ab) &\leq C\log(1+ab)\\ &\leq C\min\{\log(1+a),\log(1+b)\}+C\log(1+a)\log(1+b)\\ &\leq C\log(1+a)(1+\log(1+b))\\ &\leq Ca(1+\log(1+b)),\end{aligned}$$

where the second inequality is proved in Friesen et al. (2020+b, Lemma 8.5). Choosing $a=\mathrm{e}^{-\delta t}$ and $b=1+|y|+|\widetilde{y}|$ in the last inequality gives

$$\begin{aligned}\|P_t(x,\cdot)-\pi\|_{\mathrm{TV}} &\leq C\mathrm{e}^{-\delta t}\int_{\mathbb{R}_+^m\times\mathbb{R}_+^m}(1+\log(2+|y|+|\widetilde{y}|))\,H(\mathrm{d}y,\mathrm{d}\widetilde{y})\\ &\leq C\mathrm{e}^{-\delta t}\int_{\mathbb{R}_+^m\times\mathbb{R}_+^m}(1+\log(1+|y|)+\log(1+|\widetilde{y}|))\,H(\mathrm{d}y,\mathrm{d}\widetilde{y})\\ &= C\mathrm{e}^{-\delta t}\left(1+\log(1+|x|)+\int_{\mathbb{R}_+^m}\log(1+|y|)\pi(\mathrm{d}y)\right),\end{aligned}$$

where we have used the subadditivity $\log(1+a+b)\leq\log(1+a)+\log(1+b)$. This completes the proof of Theorem 2.5. □

Appendix A. Moments of the anisotropic stable JCIR process

The following can be shown by rather standard arguments, see e.g. Friesen et al. (2020+b, Proposition 5.1).

Proposition A.1. *Let $\eta\in(0,\alpha_{\min})$ and suppose that*

$$\int_{|z|>1}|z|^{\eta}\nu(\mathrm{d}z)<\infty. \tag{A.1}$$

Then for each $T>0$, there exists a constant $C_T>0$ such that

$$\sup_{t\in[0,T]}\mathbb{E}[|X^x(t)|^{\eta}]\leq C_T(1+|x|)^{\eta},\qquad x\in\mathbb{R}_+^m.$$

In particular, if (A.1) holds for $\eta=1$, then $(X(t))_{t\geq 0}$ has finite first moment. This moment was computed in Barczy et al. (2015) where it was shown that

$$\mathbb{E}[X^x(t)]=\mathrm{e}^{\beta t}x+\int_0^t\mathrm{e}^{\beta s}\left(b+\int_{\mathbb{R}_+^m}z\nu(\mathrm{d}z)\right)\mathrm{d}s. \tag{A.2}$$

Actually in Barczy et al. (2015) the more general class of multi-type continuous-state branching processes were studied. An extension of such a formula to general affine processes on the canonical state space was obtained in Friesen et al. (2020+b). One simple consequence is the uniform boundedness of the first moment stated below.

Proposition A.2. *Suppose that β has only eigenvalues with negative real-parts and (A.1) holds for $\eta = 1$. Then there exists a constant $C > 0$ such that*

$$\sup_{t \geq 0} \mathbb{E}[|X^x(t)|] \leq C(1 + |x|), \qquad x \in \mathbb{R}_+^m.$$

Proof: Let $\varkappa > 0$ be such that $|\mathrm{e}^{\beta t} y| \leq e^{-\varkappa t}|y|$ for all $y \in \mathbb{R}^m$ and set

$$\widetilde{b} = b + \int_{\mathbb{R}_+^m} z \nu(\mathrm{d}z).$$

Using first the sub-additivity of the square-root, then the Cauchy-Schwartz inequality and finally (A.2) we find that

$$\begin{aligned}
\mathbb{E}[|X^x(t)|] &\leq \sum_{k=1}^m \mathbb{E}[X_k^x(t)] \\
&\leq \sqrt{m} |\mathbb{E}[X^x(t)]| \\
&\leq \sqrt{m} |\mathrm{e}^{\beta t} x| + \sqrt{m} \int_0^t |\mathrm{e}^{\beta s} \widetilde{b}| \mathrm{d}s \\
&\leq \sqrt{m} \mathrm{e}^{-\varkappa t} |x| + \sqrt{m} \int_0^t \mathrm{e}^{-\varkappa s} |\widetilde{b}| \mathrm{d}s \\
&\leq \sqrt{m} |x| + \sqrt{m} \frac{|\widetilde{b}|}{\varkappa},
\end{aligned}$$

which proves the assertion. □

Appendix B. Lyapunov estimate for the extended generator

Recall that $|x|_M$ is defined by (5.2), $V(x) = (1 + |x|_M^2)^{1/2}$ and observe that we can find constants $c^* \geq c_* > 0$ such that

$$c_* |x| \leq |x|_M \leq c^* |x|, \qquad x \in \mathbb{R}_+^m. \tag{B.1}$$

Let L_0 be the extended generator of the anisotropic stable JCIR process $(Y^x(t))_{t \geq 0}$ obtained from (1.3) whose subordinator ν has only small jumps, i.e., $\nu(\{|z| > 1\}) = 0$.

Lemma B.1. *Suppose that β has only eigenvalues with negative real-parts. Then V belongs to the domain of the extended generator L_0, one has*

$$\begin{aligned}
L_0 V(x) &= \langle b + \beta x, \nabla V(x) \rangle + \int_{\{|z| \leq 1\}} \left(V(x+z) - V(x) \right) \nu(\mathrm{d}z) \\
&\quad + \sum_{j=1}^m \sigma_j x_j \int_0^\infty \left(V(x + e_j z) - V(x) - z \frac{\partial V(x)}{\partial x_j} \right) \mu_{\alpha_j}(\mathrm{d}z), \qquad x \in \mathbb{R}_+^m
\end{aligned}$$

and there exists two constants $c_1, c_2 > 0$ such that

$$L_0 V(x) \leq -c_1 V(x) + c_2, \qquad x \in \mathbb{R}_+^m. \tag{B.2}$$

Proof: By direct computation one finds that

$$\nabla V(x) = \frac{Mx}{V(x)} \quad \text{and} \quad \frac{\partial^2 V(x)}{\partial x_j \partial x_k} = \frac{M_{jk}}{V(x)} - \frac{(Mx)_k (Mx)_j}{V(x)^3},$$

which, together with (B.1), imply $|\nabla V(x)| \le C$ and $\left|\frac{\partial^2 V(x)}{\partial x_j \partial x_k}\right| \le CV(x)^{-1}$ for all $k, j \in \{1, \dots, m\}$. Here and below we use C to denote a generic positive constant whose precise value is not important and may vary from time to time. By the mean-value theorem we obtain

$$|V(x+z) - V(x)| \le C|z|, \qquad x, z \in \mathbb{R}_+^m,$$

and applying the mean-value theorem twice gives

$$\begin{aligned} \left| V(x + e_j z) - V(x) - z\frac{\partial V(x)}{\partial x_j} \right| &= \left| z^2 \int_0^1 \int_0^1 \frac{\partial^2 V(x + e_j zsr)}{\partial x_j^2} \mathrm{d}r \mathrm{d}s \right| \\ &\le Cz^2 \int_0^1 \int_0^1 \frac{1}{V(x + e_j zrs)} \mathrm{d}r \mathrm{d}s \\ &\le C\frac{z^2}{V(x)}, \end{aligned}$$

where we have used $V(x+e_j zrs) \ge (1+c_*^2|x+e_j zrs|^2)^{1/2} \ge (1+c_*^2|x|^2)^{1/2} \ge CV(x)$. Hence all integrals in $L_0 V$ are well defined and one easily finds that $|L_0 V(x)| \le CV(x)$, $x \in \mathbb{R}_+^m$. Applying the Itó formula gives $V(Y_t^x) = V(x) + \int_0^t L_0 V(Y_s^x)\mathrm{d}s + M_t(V)$, where $(M_t(V))_{t\ge 0}$ is a local martingale. Using the fact that Y_t^x has finite first moment combined with the particular form of $M_t(V)$, one can easily show that $(M_t(V))_{t\ge 0}$ is, indeed, a true martingale. Hence taking expectations gives $\mathbb{E}[V(Y_t^x)] = V(x) + \int_0^t \mathbb{E}[L_0 V(Y_s^x)]\mathrm{d}s$, i.e., V belongs to the domain of the extended generator L_0 and has the desired form.

It remains to prove (B.2). By continuity, for $|x|_M \le 1$ one clearly has $L_0 V(x) \le |L_0 V(x)| \le C$. Take $x \in \mathbb{R}_+^m$ with $|x|_M > 1$. For the drift we obtain

$$\langle b, \nabla V(x) \rangle \le |\langle b, \nabla V(x) \rangle| \le C.$$

Likewise, using the identity $M\beta + \beta^\top M = -\mathbb{1}$ we find that

$$\begin{aligned} \langle \beta x, \nabla V(x) \rangle &= \frac{1}{2}\langle M\beta x + \beta^\top M x, x\rangle V(x)^{-1} \\ &\le -\frac{1}{2}|x|^2 V(x)^{-1} \\ &\le -\frac{1}{2(c^*)^2}|x|_M^2 V(x)^{-1} \\ &\le -\frac{c_*^2}{2\sqrt{2}(c^*)^2}|x|_M \\ &\le -\frac{c_*^2}{4(c^*)^2} V(x), \end{aligned}$$

where we have used (B.1) and $V(x) \le \sqrt{2}|x|_M$ since $|x|_M > 1$. For the state-independent jumps we obtain

$$\int_{\{|z|\le 1\}} (V(x+z) - V(x))\, \nu(\mathrm{d}z) \le C,$$

while the state-dependent jumps can be estimated by

$$\begin{aligned}
&\sum_{j=1}^{m}\sigma_j x_j\int_0^{\infty}\left(V(x+e_jz)-V(x)-z\frac{\partial V(x)}{\partial x_j}\right)\mu_{\alpha_j}(\mathrm{d}z)\\
&=\sum_{j=1}^{m}\sigma_j x_j\int_0^{R}\left(V(x+e_jz)-V(x)-z\frac{\partial V(x)}{\partial x_j}\right)\mu_{\alpha_j}(\mathrm{d}z)\\
&\quad+\sum_{j=1}^{m}\sigma_j x_j\int_R^{\infty}\left(V(x+e_jz)-V(x)-z\frac{\partial V(x)}{\partial x_j}\right)\mu_{\alpha_j}(\mathrm{d}z)\\
&\le C\sum_{j=1}^{m}\frac{x_j}{V(x)}\int_0^{R}z^2\mu_{\alpha_j}(\mathrm{d}z)+C\sum_{j=1}^{m}x_j\int_R^{\infty}z\mu_{\alpha_j}(\mathrm{d}z)\\
&\le C\max_{j\in\{1,\dots,m\}}\int_0^{R}z^2\mu_{\alpha_j}(\mathrm{d}z)+C\max_{j\in\{1,\dots,m\}}\int_R^{\infty}z\mu_{\alpha_j}(\mathrm{d}z)V(x),
\end{aligned}$$

where $R>0$ is some constant to be fixed below. Combining all estimates we obtain

$$\begin{aligned}
L_0V(x)\le C&\left(1+\max_{j\in\{1,\dots,m\}}\int_0^{R}z^2\mu_{\alpha_j}(\mathrm{d}z)\right)\\
&-\left(\frac{c_*^2}{4(c^*)^2}-C\max_{j\in\{1,\dots,m\}}\int_R^{\infty}z\mu_{\alpha_j}(\mathrm{d}z)\right)V(x).
\end{aligned}$$

Choosing R large enough such that

$$C\max_{j\in\{1,\dots,m\}}\int_R^{\infty}z\mu_{\alpha_j}(\mathrm{d}z)\le\frac{c_*^2}{8(c^*)^2},$$

the assertion is proved. □

Appendix C. Proof of Theorem 4.1

Let $\lambda>0$ and $(a_1,\dots,a_m)$ be the anisotropy defined in (4.1). The anisotropic Hölder-Zygmund space $C_b^{\lambda,a}(\mathbb{R}^m)$ is defined as the Banach space of functions ϕ with finite norm

$$\|\phi\|_{C_b^{\lambda,a}}=\|\phi\|_\infty+\sum_{k=1}^{m}\sup_{h\in[-1,1]}|h|^{-\lambda/a_k}\|\Delta_{he_k}\phi\|_\infty.$$

The following lemma provides our main technical tool for the proof of Theorem 4.1.

Lemma C.1. *Let $\lambda,\eta>0$ be such that $(\lambda+\eta)/a_k\in(0,1)$ for all $k=1,\dots,d$. Suppose that G is a finite measure over $\mathbb{R}^m$ and there exists $A>0$ such that for all $\phi\in C_b^{\eta,a}(\mathbb{R}^m)$, all $k=1,\dots,m$ and all $h\in[-1,1]$*

$$\left|\int_{\mathbb{R}^m}(\phi(x+he_k)-\phi(x))G(\mathrm{d}x)\right|\le A\|\phi\|_{C_b^{\eta,a}}|h|^{(\lambda+\eta)/a_k}. \tag{C.1}$$

Then there exists $g\in B_{1,\infty}^{\lambda,a}(\mathbb{R}^m)$ such that $G(\mathrm{d}x)=g(x)\mathrm{d}x$ and

$$\|g\|_{B_{1,\infty}^{\lambda,a}}\le G(\mathbb{R}^m)+3mA(2m)^{\eta/\lambda}\left(1+\frac{\lambda}{\eta}\right)^{1+\frac{\eta}{\lambda}}. \tag{C.2}$$

This lemma was first proved in Debussche and Fournier (2013, Lemma 2.1) and Debussche and Romito (2014) for the isotropic case $a_1 = \cdots = a_m$. Above anisotropic version is due to Friesen et al. (2020+a). The essential step in the process of proving Theorem 4.1 is based on a suitable application of this lemma to the finite measure $G_t(\mathrm{d}y) = \rho_\delta(y)P_t(x, \mathrm{d}y)$. In order to prove (C.1) for G_t, we use an approximation similar to Debussche and Fournier (2013), Friesen et al. (2020), and Friesen et al. (2020+a). From now on we fix $x \in \mathbb{R}_+^m$ and let $X^x = (X^x(t))_{t\geq 0}$ be the unique solution to (1.3) with $X^x(0) = x$ and ν satisfying

$$\int_{\mathbb{R}_+^m} \mathbb{1}_{\{|z|>1\}}|z|^{1+\tau}\nu(\mathrm{d}z) < \infty$$

for some $\tau > 0$. To simplify the notation we let $(X(t))_{t\geq 0}$ stand for $(X^x(t))_{t\geq 0}$.

C.1. *Short time approximation.* For $\varepsilon \in (0, 1\wedge t)$, define the approximation $X^\varepsilon(t) = (X_1^\varepsilon(t), \ldots, X_m^\varepsilon(t))$ by

$$\begin{aligned} X_i^\varepsilon(t) = X_i(t-\varepsilon) + \left(b_i + \sum_{k=1}^m \beta_{ik}X_k(t-\varepsilon)\right)\varepsilon \qquad \text{(C.3)} \\ + \sigma_i^{1/\alpha_i}X_i(t-\varepsilon)^{1/\alpha_i}(Z_i(t) - Z_i(t-\varepsilon)) + (J_i(t) - J_i(t-\varepsilon)), \end{aligned}$$

where $i = 1, \ldots, m$. Define $\kappa_1, \ldots, \kappa_m > 0$ by

$$\kappa_i = \min\left\{1 + \frac{1}{\alpha_{\max}}, \frac{1}{\alpha_i} + \frac{1}{\alpha_i^2}\right\}, \qquad i = 1, \ldots, m. \qquad \text{(C.4)}$$

The next proposition shows that the convergence rate for $X_i^\varepsilon(t) \to X_i(t)$ as $\varepsilon \to 0$ is precisely given by κ_i.

Proposition C.2. *Let $i \in \{1, \ldots, m\}$ be arbitrary. The following assertions hold:*

(a) *For each $\eta \in (0, (1+\tau)\wedge\alpha_{\min})$ and $T \geq 1$, there exists a constant $C = C(\eta, T) > 0$ such that, for all $0 \leq s \leq t \leq s+1 \leq T$, it holds that*

$$\mathbb{E}[|X_i(t) - X_i(s)|^\eta] \leq C(1+|x|)^\eta(t-s)^{\eta/\alpha_i}.$$

(b) *For each $\eta \in (0,1)$ and $T > 0$, there exists a constant $C = C(\eta, T) > 0$ such that*

$$\mathbb{E}[|X_i(t) - X_i^\varepsilon(t)|^\eta] \leq C(1+|x|)^\eta\varepsilon^{\eta\kappa_i}, \quad t \in (0,T], \quad \varepsilon \in (0, 1\wedge t).$$

Proof: Fix constants $\gamma_1, \ldots, \gamma_m$, satisfying for each $i = 1, \ldots, m$

$$\gamma_i \in (\alpha_i, 2) \quad \text{and} \quad \frac{\gamma_i}{\alpha_i} < \min\{1+\tau, \alpha_{\min}\}.$$

In the following we will use C to denote a positive constant, whose exact value is not important and may change from time to time.

(a) Write $\mathbb{E}[|X_i(t) - X_i(s)|^\eta] \leq R_1 + R_2 + R_3$, where

$$\begin{aligned} R_1 &= C\mathbb{E}\left[\left|\int_s^t \left(b_i + \sum_{k=1}^m \beta_{ik}X_k(u)\right)\mathrm{d}u\right|^\eta\right], \\ R_2 &= C\mathbb{E}\left[\left|\int_s^t X_i(u-)^{1/\alpha_i}\mathrm{d}Z_i(u)\right|^\eta\right], \\ R_3 &= C\mathbb{E}\left[|J_i(t) - J_i(s)|^\eta\right]. \end{aligned}$$

If $\eta \in (0,1]$, then we use the Jensen inequality and Proposition A.1 to obtain

$$\begin{aligned} R_1 &\le C\left(\mathbb{E}\left[\left|\int_s^t \left(b_i + \sum_{k=1}^m \beta_{ik} X_k(u)\right) \mathrm{d}u\right|\right]\right)^\eta \\ &\le C(t-s)^\eta + C(t-s)^\eta \sup_{u\in[s,t]} (\mathbb{E}[|X(u)|])^\eta \\ &\le C(1+|x|)^\eta (t-s)^\eta. \end{aligned}$$

If $\eta \in (1, \alpha_{\min})$, then we use the Hölder inequality with $\frac{1}{\eta} + \frac{1}{\frac{\eta}{\eta-1}} = 1$ to obtain

$$\begin{aligned} R_1 &\le C(t-s)^{\eta-1} \int_s^t \mathbb{E}\left[\left|b_i + \sum_{k=1}^m \beta_{ik} X_k(u)\right|^\eta\right] \mathrm{d}u \\ &\le C(t-s)^{\eta-1}(t-s)\left(1 + \sup_{u\in[s,t]} \mathbb{E}[|X(u)|^\eta]\right) \\ &\le C\,(1+|x|)^\eta\,(t-s)^\eta. \end{aligned}$$

Combining both cases $\eta \in (0,1]$ and $\eta \in (1,\alpha_{\min})$ we find that $R_1 \le C\,(1+|x|)^\eta\,(t-s)^\eta$. For the second term we apply Lemma E.2 to obtain

$$R_2 \le C(t-s)^{\eta/\alpha_i} \sup_{u\in[s,t]} \left(\mathbb{E}[X_i(u)^{\gamma_i/\alpha_i}]\right)^{\eta/\gamma_i}. \tag{C.5}$$

Since γ_i also satisfies $\gamma_i/\alpha_i < \min\{1+\tau, \alpha_{\min}\}$, we may apply Proposition A.1 to find that

$$\sup_{u\in[s,t]} \mathbb{E}[X_i(u)^{\gamma_i/\alpha_i}] \le \sup_{u\in[s,t]} \mathbb{E}[|X(u)|^{\gamma_i/\alpha_i}] \le C\,(1+|x|)^{\gamma_i/\alpha_i}.$$

Inserting this into (C.5) gives $R_2 \le C(1+|x|)^{\eta/\alpha_i}(t-s)^{\eta/\alpha_i}$. For the last term we may apply the estimates for the stochastic integrals from Friesen et al. (2020, appendix) to find that $R_3 = C\mathbb{E}[J_i(t-s)^\eta] \le C(t-s)^{\eta\wedge 1}$. Combining all estimates for R_1, R_2, R_3 yields

$$\begin{aligned} \mathbb{E}[|X_i(t) - X_i(s)|^\eta] &\le C(1+|x|)^\eta (t-s)^{\eta\wedge 1} + C(1+|x|)^{\eta/\alpha_i}(t-s)^{\eta/\alpha_i} \\ &\le C(1+|x|)^\eta (t-s)^{\eta/\alpha_i}. \end{aligned}$$

This proves the assertion.

(b) Write $\mathbb{E}[|X_i(t) - X_i^\varepsilon(t)|^\eta] \le R_1 + R_2$, where

$$\begin{aligned} R_1 &= \mathbb{E}\left[\left|\int_{t-\varepsilon}^t \left(\sum_{k=1}^m \beta_{ik}(X_k(u) - X_k(t-\varepsilon))\right) \mathrm{d}u\right|^\eta\right], \\ R_2 &= \sigma_i^{\eta/\alpha_i} \mathbb{E}\left[\left|\int_{t-\varepsilon}^t (X_i(u-)^{1/\alpha_i} - X_i(t-\varepsilon)^{1/\alpha_i}) \mathrm{d}Z_i(u)\right|^\eta\right]. \end{aligned}$$

For the first term we use part (a) and the fact that $\eta \in (0,1)$ to obtain

$$\begin{aligned} R_1 &\le \left(\mathbb{E}\left[\int_{t-\varepsilon}^t \sum_{k=1}^m |\beta_{ik}||X_k(u) - X_k(t-\varepsilon)| \mathrm{d}u\right]\right)^\eta \\ &\le C\varepsilon^\eta \sum_{k=1}^m \sup_{u\in[t-\varepsilon,t]} \left(\mathbb{E}\left[|X_k(u) - X_k(t-\varepsilon)|\right]\right)^\eta \end{aligned}$$

$$\leq C\varepsilon^{\eta}\sum_{k=1}^{m}(1+|x|)^{\eta}\varepsilon^{\eta/\alpha_k}$$
$$\leq C(1+|x|)^{\eta}\varepsilon^{\eta+\eta/\alpha_{\max}}.$$

Let us turn to the second term. We use Lemma E.2 and then $|y^{1/\alpha_i} - z^{1/\alpha_i}| \leq |y-z|^{1/\alpha_i}$ for $y, z \geq 0$ to find that

$$R_2 \leq C\varepsilon^{\eta/\alpha_i} \sup_{u\in[t-\varepsilon,t]} \left(\mathbb{E}\left[|X_i(u)^{1/\alpha_i} - X_i(t-\varepsilon)^{1/\alpha_i}|^{\gamma_i}\right]\right)^{\eta/\gamma_i}$$
$$\leq C\varepsilon^{\eta/\alpha_i} \sup_{u\in[t-\varepsilon,t]} \left(\mathbb{E}[|X_i(u) - X_i(t-\varepsilon)|^{\gamma_i/\alpha_i}]\right)^{\eta/\gamma_i}.$$

Since $\gamma_i/\alpha_i < \min\{1+\tau, \alpha_{\min}\}$, we may apply part (a) which gives

$$\sup_{u\in[t-\varepsilon,t]} \mathbb{E}[|X_i(u) - X_i(t-\varepsilon)|^{\gamma_i/\alpha_i}] \leq C(1+|x|)^{\gamma_i/\alpha_i}\varepsilon^{\gamma_i/\alpha_i^2}$$

and hence

$$R_2 \leq C(1+|x|)^{\eta/\alpha_i}\varepsilon^{\frac{\eta}{\alpha_i}(1+1/\alpha_i)}.$$

This proves the assertion. □

C.2. *The key estimate.* Recall that $\kappa_1, \ldots, \kappa_m$ are defined in (C.4). Based on the previous approximation we show the following.

Proposition C.3. *Let $t > 0$ be arbitrary and fixed. Take $\varkappa \in (0, 1/\alpha_{\max}]$ and let $\eta \in (0, \varkappa a_{\min})$. Then there exists a constant $C > 0$ such that, for any $\varepsilon \in (0, 1\wedge t)$, $h \in [-1, 1]$, $\phi \in C_b^{\eta,a}(\mathbb{R}^m)$ and $i \in \{1, \ldots, m\}$,*

$$|\mathbb{E}\left[\rho_\delta(X(t))\Delta_{he_i}\phi(X(t))\right]|$$
$$\leq C\|\phi\|_{C_b^{\eta,a}}(1+|x|)^{\varkappa}\left(|h|^{\eta/a_i}\varepsilon^{\varkappa/\alpha_{\max}} + |h|\varepsilon^{-1/\alpha_i} + \max_{j\in\{1,\ldots,m\}}\varepsilon^{\eta\kappa_j/a_j}\right).$$

Proof: For $\varepsilon \in (0, 1\wedge t)$ let $X^\varepsilon(t) = (X_1^\varepsilon(t), \ldots, X_m^\varepsilon(t))$ be given as in (C.3). Then

$$|\mathbb{E}\left[\rho_\delta(X(t))\Delta_{he_i}\phi(X(t))\right]| \leq R_1 + R_2 + R_3,$$

where R_1, R_2, R_3 are given by

$$R_1 = \left|\mathbb{E}\left[\Delta_{he_i}\phi(X(t))\left(\rho_\delta(X(t)) - \rho_\delta(X(t-\varepsilon))\right)\right]\right|,$$
$$R_2 = \mathbb{E}\left[|\Delta_{he_i}\phi(X(t)) - \Delta_{he_i}\phi(X^\varepsilon(t))|\rho_\delta(X(t-\varepsilon)\right],$$
$$R_3 = \left|\mathbb{E}\left[\rho_\delta(X(t-\varepsilon))\Delta_{he_i}\phi(X^\varepsilon(t))\right]\right|.$$

For the first term we use $\rho_\delta \leq 1$ and $\varkappa \leq 1/\alpha_k$ to find

$$|\rho_\delta(x) - \rho_\delta(y)| \leq 2 \wedge \left(\sum_{k=1}^{m}|x_k^{1/\alpha_k} - y_k^{1/\alpha_k}|\right)$$
$$\leq \sum_{k=1}^{m} 2\wedge|x_k - y_k|^{1/\alpha_k}$$
$$\leq C\sum_{k=1}^{m}|x_k - y_k|^{\varkappa},$$

and hence deduce from Proposition C.2.(a) that

$$\begin{aligned} R_1 &\le C\|\phi\|_{C_b^{\eta,a}} |h|^{\eta/a_i} \sum_{k=1}^m \mathbb{E}[|X_k(t) - X_k(t-\varepsilon)|^{\varkappa}] \\ &\le C\|\phi\|_{C_b^{\eta,a}} |h|^{\eta/a_i} \sum_{k=1}^m (1+|x|)^{\varkappa} \varepsilon^{\varkappa/\alpha_k} \\ &\le C\|\phi\|_{C_b^{\eta,a}} |h|^{\eta/a_i} (1+|x|)^{\varkappa} \varepsilon^{\varkappa/\alpha_{\max}}. \end{aligned}$$

For R_2 we first use that $\phi \in C_b^{\eta,a}(\mathbb{R}^m)$, then $\rho_\delta \le 1$ and finally Proposition C.2.(b) to obtain

$$\begin{aligned} R_2 &\le C\|\phi\|_{C_b^{\eta,a}} \max_{j\in\{1,\dots,m\}} \mathbb{E}\left[|X_j(t) - X_j^{\varepsilon}(t)|^{\eta/a_j}\right] \\ &\le C\|\phi\|_{C_b^{\eta,a}} \max_{j\in\{1,\dots,m\}} \varepsilon^{\eta\kappa_j/a_j}(1+|x|)^{\eta/a_j} \\ &\le C\|\phi\|_{C_b^{\eta,a}} (1+|x|)^{\varkappa} \max_{j\in\{1,\dots,m\}} \varepsilon^{\eta\kappa_j/a_j}. \end{aligned}$$

Let us turn to R_3. Define $\sigma(x) = \mathrm{diag}((\sigma_1 x_1)^{1/\alpha_1}, \dots, (\sigma_m x_m)^{1/\alpha_m})$. Let f_t^Z be the density of $Z(t) = (Z_1(t), \dots, Z_m(t))$. Using (C.3) we find that

$$X^{\varepsilon}(t) = U^{\varepsilon}(t) + \sigma(X(t-\varepsilon))(Z(t) - Z(t-\varepsilon)),$$

with $U^{\varepsilon}(t) = (U_1^{\varepsilon}(t), \dots, U_m^{\varepsilon}(t))$ being given by

$$U_i^{\varepsilon}(t) = X_i(t-\varepsilon) + \left(b_i + \sum_{k=1}^m \beta_{ik} X_k(t-\varepsilon)\right)\varepsilon + (J_i(t) - J_i(t-\varepsilon)).$$

Finally note that $\sigma(X(t-\varepsilon))^{-1} = \mathrm{diag}((\sigma_1 X_1(t-\varepsilon))^{-1/\alpha_1}, \dots, (\sigma_m X_m(t-\varepsilon))^{-1/\alpha_m})$ is well-defined, since $\mathbb{P}[X(t-\varepsilon) \in \mathbb{R}_{++}^m] = 1$ holds by Proposition 4.2. Then we obtain for each $\varepsilon \in (0, 1\wedge t)$,

$$\begin{aligned} R_3 &= \left|\mathbb{E}\left[\int_{\mathbb{R}^m} \rho_\delta(X(t-\varepsilon))(\Delta_{he_i}\phi)(U^{\varepsilon}(t) + \sigma(X(t-\varepsilon))z) f_{\varepsilon}^Z(z)\mathrm{d}z\right]\right| \\ &= \left|\mathbb{E}\left[\int_{\mathbb{R}^m} \rho_\delta(X(t-\varepsilon))\phi(U^{\varepsilon}(t) + \sigma(X(t-\varepsilon))z)(\Delta_{-h\sigma(X(t-\varepsilon))^{-1}e_i} f_{\varepsilon}^Z)(z)\mathrm{d}z\right]\right| \\ &\le \|\phi\|_{\infty} \mathbb{E}\left[\rho_\delta(X(t-\varepsilon)) \int_{\mathbb{R}^m} |(\Delta_{-h\sigma(X(t-\varepsilon))^{-1}e_i} f_{\varepsilon}^Z)(z)|\mathrm{d}z\right] \\ &\le \|\phi\|_{\infty} |h| \sigma_i^{-1/\alpha_i} \mathbb{E}\left[\rho_\delta(X(t-\varepsilon)) X_i(t-\varepsilon)^{-1/\alpha_i} \int_{\mathbb{R}^m} \left|\frac{\partial f_{\varepsilon}^Z(z)}{\partial z_i}\right| \mathrm{d}z\right] \\ &\le C\|\phi\|_{C_b^{\eta,a}} |h| \varepsilon^{-1/\alpha_i}, \end{aligned}$$

where we have used Lemma E.1 and $\rho_\delta(x) x_i^{-1/\alpha_i} \le 1$. Summing up the estimates for R_1, R_2, R_3 yields the assertion. □

C.3. *Concluding the proof of Theorem 4.1.* Below we provide the proof of Theorem 4.1. Fix $t > 0$ and $x \in \mathbb{R}_+^m$. We will show that Lemma C.1 applies to the finite measure $G_t(x, \mathrm{d}y) = \rho_\delta(y) P_t(x, \mathrm{d}y)$. Using the particular form of κ_j we obtain $\kappa_j \alpha_j > 1$ and hence $\kappa_j / a_j > 1/\overline{\alpha}$ for all $j \in \{1, \dots, m\}$. This implies

$$\frac{a_j}{\kappa_j} \frac{1}{a_i} < \frac{\overline{\alpha}}{a_i} = \alpha_i, \quad i, j \in \{1, \dots, m\}.$$

Hence we find $\eta \in (0,1)$ and $c_1, \dots, c_m > 0$ such that, for all $i, j \in \{1, \dots, m\}$,

$$0 < \eta < a_i \varkappa, \qquad \frac{a_j}{\kappa_j} \frac{1}{a_i} < c_i < \alpha_i \left(1 - \frac{\eta}{a_i}\right).$$

Define

$$\lambda = \min_{i,j \in \{1,\dots,m\}} \left\{ \varkappa c_i a_i / \alpha_{\max},\ a_i - \eta - \frac{a_i c_i}{\alpha_i},\ \eta \left(c_i a_i \frac{\kappa_j}{a_j} - 1 \right) \right\} > 0.$$

Let $\phi \in C_b^{\eta,a}(\mathbb{R}^m)$. By Proposition C.3 we obtain, for $h \in [-1,1]$, $\varepsilon = |h|^{c_i}(1 \wedge t)$ and $i \in \{1, \dots, m\}$,

$$\begin{aligned}
&|\mathbb{E}\left[\rho_\delta(X(t))\Delta_{he_i}\phi(X(t))\right]| \\
&\le C\|\phi\|_{C_b^{\eta,a}}(1+|x|)^{\varkappa}\left(|h|^{\eta/a_i}\varepsilon^{\varkappa/\alpha_{\max}} + |h|\varepsilon^{-1/\alpha_i} + \max_{j\in\{1,\dots,m\}} \varepsilon^{\eta\kappa_j/a_j}\right) \\
&\le \frac{C\|\phi\|_{C_b^{\eta,a}}}{(1\wedge t)^{1/\alpha_i}}(1+|x|)^{\varkappa}\left(|h|^{\eta/a_i + c_i\varkappa/\alpha_{\max}} + |h|^{1-c_i/\alpha_i} + \max_{j\in\{1,\dots,m\}} |h|^{c_i\eta\kappa_j/a_j}\right) \\
&= \frac{C\|\phi\|_{C_b^{\eta,a}}}{(1\wedge t)^{1/\alpha_i}}|h|^{\eta/a_i}(1+|x|)^{\varkappa}\bigg(|h|^{c_i\varkappa/\alpha_{\max}} + |h|^{1-\eta/a_i - c_i/\alpha_i} \\
&\qquad + \max_{j\in\{1,\dots,m\}} |h|^{c_i\eta\kappa_j/a_j - \eta/a_i}\bigg) \\
&\le \frac{C\|\phi\|_{C_b^{\eta,a}}}{(1\wedge t)^{1/\alpha_i}}(1+|x|)^{\varkappa}|h|^{(\eta+\lambda)/a_i}.
\end{aligned}$$

This shows that Lemma C.1 is applicable to $G_t(x, \mathrm{d}y)$. Hence $G_t(x, \mathrm{d}y)$ has a density $g_t(x,y)$, i.e., $\rho_\delta(y)P_t(x,\mathrm{d}y) = g_t(x,y)\mathrm{d}y$. In view of (C.2) this density satisfies

$$\begin{aligned}
\|g_t(x,\cdot)\|_{B_{1,\infty}^{\lambda,a}(\mathbb{R}_+^m)} &\le \int_{\mathbb{R}_+^m} \rho_\delta(y)P_t(x,\mathrm{d}y) + C(t)(1+|x|)^{\varkappa}(1\wedge t)^{-1/\alpha_{\min}} \\
&\le C(t)(1+|x|)^{\varkappa}(1\wedge t)^{-1/\alpha_{\min}},
\end{aligned}$$

where we have used $\rho \le 1$ and $C(t)$ is a generic constant which is locally bounded in $t \ge 0$. Since $\rho_\delta(y) > 0$ for $y \in \mathbb{R}_{++}^m$, $P_t(x,\mathrm{d}y)$ has also a density $p_t(x,y)$ on $\mathbb{R}_{++}^m$ which gives $P_t(x,\mathrm{d}y) = p_t(x,y)\mathrm{d}y + P_t^{\mathrm{sing}}(x,\mathrm{d}y)$, where $P_t^{\mathrm{sing}}(x,\mathrm{d}y)$ is supported on $\partial\mathbb{R}_+^m$. Using Proposition 4.2 we conclude that $P_t^{\mathrm{sing}}(x,\mathrm{d}y) = 0$ and hence $p_t^\delta(x,y) = g_t(x,y)$. This proves the assertion of Theorem 4.1.

Appendix D. Some results on ergodicity in total variation norm

In this section we briefly summarize some results on geometric ergodicity in the total variation distance for continuous-time Markov processes. For additional details we refer to Hairer (2016) and Kulik (2018). Let E be a Polish space and let $(X_t)_{t\ge 0}$ be a Feller process on E. Denote by $(P_t(x,\mathrm{d}y))_{t\ge 0}$ its transition probabilities and by L the extended generator.

Theorem D.1. *Suppose that the following conditions are satisfied:*

(a) *There exists a continuous function $V : E \longmapsto [1,\infty)$ which belongs to the domain of the extended generator such that*

$$LV(x) \le -aV(x) + M, \qquad x \in E,$$

where $a, M > 0$ are some constants. Moreover, for each $R > 0$ the level sets $\{(x,y) \in E^2 \mid V(x) + V(y) \le R\}$ are compact.

(b) *For each $R > 0$, there exists $h > 0$ and $\delta \in (0,2)$ such that*

$$\|P_h(x,\cdot) - P_h(y,\cdot)\|_{\mathrm{TV}} \le 2 - \delta$$

holds for all $x, y \in E$ with $V(x) + V(y) \le R$.

Then there exists constants $C, \beta > 0$ such that for all $t \ge 0$ and $x, y \in E$,

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{\mathrm{TV}} \le Ce^{-\beta t}\left(V(x) + V(y)\right).$$

Moreover, there exists a unique invariant probability measure π. This measure satisfies

$$\int_E V(x)\pi(\mathrm{d}x) < \infty \tag{D.1}$$

and for all $t \ge 0$ and $x \in E$ one has

$$\|P_t(x,\cdot) - \pi\|_{\mathrm{TV}} \le Ce^{-\beta t}\left(V(x) + \int_E V(y)\pi(\mathrm{d}y)\right).$$

A proof of this Theorem is given in Hairer (2016, Theorem 4.1). Moreover, the same result can also be obtained from a combination of Corollary 2.8.3 and Theorem 3.2.3 in Kulik (2018), where also additional comments and examples are given.

Appendix E. Some properties of the cylindrical Lévy process $(Z_1, \dots, Z_m)$

Observe that the Lévy process $Z = (Z_1, \dots, Z_m)$ has symbol

$$\Psi_Z(\xi) = \int_{\mathbb{R}^m_+}\left(\mathrm{e}^{i\langle \xi, z\rangle} - 1 - \mathrm{i}\langle \xi, z\rangle\right)\mu(\mathrm{d}z) = \Psi_{\alpha_1}(\xi_1) + \cdots + \Psi_{\alpha_m}(\xi_m),$$

where the Lévy measure μ is given by

$$\mu(\mathrm{d}z) = \sum_{k=1}^m \mu_{\alpha_k}(\mathrm{d}z_k) \otimes \prod_{j \ne k}\delta_0(\mathrm{d}z_j).$$

The next lemma is standard and follows from the scaling property $Z_j(t) = t^{1/\alpha_j}Z_j(1)$, $j = 1, \dots, m$, where equality holds in the sense of distributions.

Lemma E.1. *$Z(t)$ has for each $t > 0$ a smooth density f_t^Z on $\mathbb{R}^m$. Moreover, there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}^m}\left|\frac{\partial f_t^Z(z)}{\partial z_j}\right|\mathrm{d}z \le Ct^{-1/\alpha_j}, \qquad t > 0.$$

Below we state some useful estimates on stochastic integrals with respect to the Lévy processes $Z_1, \dots, Z_m$ due to Debussche and Fournier (2013, Lemma A.2).

Lemma E.2. *Let $0 < \eta \le \alpha_j < \gamma \le 2$. Then there exists a constant $C = C(\eta, \gamma) > 0$ such that, for any predictable process $H(u)$ and $0 \le s \le t \le s+1$,*

$$\mathbb{E}\left[\left|\int_s^t H(u)dZ_j(u)\right|^\eta\right] \le C(t-s)^{\eta/\alpha_j}\sup_{u\in[s,t]}\mathbb{E}\left[|H(u)|^\gamma\right]^{\eta/\gamma}.$$

Acknowledgements

The authors would like to thank the anonymous referees for their careful reading of the manuscript. We would also like to thank Alexei Kulik for some interesting discussions on ergodicity of Markov processes and, in particular, on the meaning and sufficiency of condition (b) in Theorem D.1.

References

Alfonsi, A. *Affine diffusions and related processes: simulation, theory and applications*, volume 6 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan (2015). ISBN 978-3-319-05220-5; 978-3-319-05221-2. MR3363174.

Barczy, M., Döring, L., Li, Z., and Pap, G. Stationarity and ergodicity for an affine two-factor model. *Adv. in Appl. Probab.*, **46** (3), 878–898 (2014). MR3254346.

Barczy, M., Li, Z., and Pap, G. Stochastic differential equation with jumps for multi-type continuous state and continuous time branching processes with immigration. *ALEA Lat. Am. J. Probab. Math. Stat.*, **12** (1), 129–169 (2015). MR3340375.

Bass, R. F. and Cranston, M. The Malliavin calculus for pure jump processes and applications to local time. *Ann. Probab.*, **14** (2), 490–532 (1986). MR832021.

Bogdan, K., Sztonyk, P., and Knopova, V. Heat kernel of anisotropic nonlocal operators. *Doc. Math.*, **25**, 1–54 (2020).

Chaker, J. The martingale problem for a class of nonlocal operators of diagonal type. *Math. Nachr.*, **292** (11), 2316–2337 (2019). MR4033009.

Chazal, M., Loeffen, R., and Patie, P. Option pricing in a one-dimensional affine term structure model via spectral representations. *SIAM J. Financial Math.*, **9** (2), 634–664 (2018a). MR3805842.

Chazal, M., Loeffen, R., and Patie, P. Smoothness of continuous state branching with immigration semigroups. *J. Math. Anal. Appl.*, **459** (2), 619–660 (2018b). MR3732548.

Dachkovski, S. Anisotropic function spaces and related semi-linear hypoelliptic equations. *Math. Nachr.*, **248/249**, 40–61 (2003). MR1950714.

De Marco, S. Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions. *Ann. Appl. Probab.*, **21** (4), 1282–1321 (2011). MR2857449.

Debussche, A. and Fournier, N. Existence of densities for stable-like driven SDE's with Hölder continuous coefficients. *J. Funct. Anal.*, **264** (8), 1757–1778 (2013). MR3022725.

Debussche, A. and Romito, M. Existence of densities for the 3D Navier-Stokes equations driven by Gaussian noise. *Probab. Theory Related Fields*, **158** (3-4), 575–596 (2014). MR3176359.

Duffie, D., Filipović, D., and Schachermayer, W. Affine processes and applications in finance. *Ann. Appl. Probab.*, **13** (3), 984–1053 (2003). MR1994043.

Duffie, D. and Gârleanu, N. Risk and Valuation of Collateralized Debt Obligations. *Financial Analysts Journal*, **57** (1), 41–59 (2001). DOI: 10.2469/faj.v57.n1.2418.

Duhalde, X., Foucart, C., and Ma, C. On the hitting times of continuous-state branching processes with immigration. *Stochastic Process. Appl.*, **124** (12), 4182–4201 (2014). MR3264444.

Filipović, D., Mayerhofer, E., and Schneider, P. Density approximations for multivariate affine jump-diffusion processes. *J. Econometrics*, **176** (2), 93–111 (2013). MR3084047.

Foucart, C. and Uribe Bravo, G. Local extinction in continuous-state branching processes with immigration. *Bernoulli*, **20** (4), 1819–1844 (2014). MR3263091.

Friesen, M., Jin, P., Kremer, J., and Rüdiger, B. Exponential ergodicity for stochastic equations of nonnegative processes with jumps. *ArXiv Mathematics e-prints* (2019a). arXiv: 1902.02833.

Friesen, M., Jin, P., and Rüdiger, B. Boundary behavior of multi-type continuous-state branching processes with immigration. *ArXiv Mathematics e-prints* (2019b). arXiv: 1902.01162.

Friesen, M., Jin, P., and Rüdiger, B. Existence of densities for multi-type continuous-state branching processes with immigration. *Stochastic Process. Appl.* (2020). DOI: 10.1016/j.spa.2020.03.012.

Friesen, M., Jin, P., and Rüdiger, B. Existence of densities for stochastic differential equations driven by Lévy processes with anisotropic jumps (2020+a). To appear in *Ann. Inst. Henri Poincaré Probab. Stat.*

Friesen, M., Jin, P., and Rüdiger, B. Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes (2020+b). To appear in *Ann. Appl. Probab.*

Fu, Z. and Li, Z. Stochastic equations of non-negative processes with jumps. *Stochastic Process. Appl.*, **120** (3), 306–330 (2010). MR2584896.

Hairer, M. Convergence of Markov Processes (2016). Lecture notes, available online at `www.hairer.org/notes/Convergence.pdf`.

Jiao, Y., Ma, C., and Scotti, S. Alpha-CIR model with branching processes in sovereign interest rate modeling. *Finance Stoch.*, **21** (3), 789–813 (2017). MR3663644.

Jiao, Y., Ma, C., Scotti, S., and Zhou, C. The Alpha-Heston stochastic volatility model. *ArXiv Mathematics e-prints* (2018). arXiv: 1812.01914.

Jin, P., Kremer, J., and Rüdiger, B. Exponential ergodicity of an affine two-factor model based on the α-root process. *Adv. in Appl. Probab.*, **49** (4), 1144–1169 (2017). MR3732190.

Jin, P., Kremer, J., and Rüdiger, B. Moments and ergodicity of the jump-diffusion CIR process. *Stochastics*, **91** (7), 974–997 (2019). MR4018158.

Jin, P., Kremer, J., and Rüdiger, B. Existence of limiting distribution for affine processes. *J. Math. Anal. Appl.*, **486** (2), 123912, 31pp (2020). MR4060087.

Keller-Ressel, M. and Mijatović, A. On the limit distributions of continuous-state branching processes with immigration. *Stochastic Process. Appl.*, **122** (6), 2329–2345 (2012). MR2922631.

Knopova, V. and Kulik, A. Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise. *Ann. Inst. Henri Poincaré Probab. Stat.*, **54** (1), 100–140 (2018). MR3765882.

Kulczycki, T. and Ryznar, M. Transition density estimates for diagonal systems of SDEs driven by cylindrical α-stable processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, **15** (2), 1335–1375 (2018). MR3877025.

Kulczycki, T. and Ryznar, M. Semigroup properties of solutions of SDEs driven by Lévy processes with independent coordinates. *ArXiv Mathematics e-prints* (2019). arXiv: 1906.07173.

Kulczycki, T., Ryznar, M., and Sztonyk, P. Strong Feller property for SDEs driven by multiplicative cylindrical stable noise (2020+). To appear in *Potential Anal.*

Kulik, A. *Ergodic behavior of Markov processes. With applications to limit theorems*, volume 67 of *De Gruyter Studies in Mathematics*. De Gruyter, Berlin (2018). ISBN 978-3-11-045870-1; 978-3-11-045871-8. MR3791835.

Li, Z. *Measure-valued branching Markov processes*. Probability and its Applications (New York). Springer, Heidelberg (2011). ISBN 978-3-642-15003-6. MR2760602.

Li, Z. and Ma, C. Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. *Stochastic Process. Appl.*, **125** (8), 3196–3233 (2015). MR3343292.

Mayerhofer, E., Stelzer, R., and Vestweber, J. Geometric ergodicity of affine processes on cones. *Stochastic Process. Appl.*, **130** (7), 4141–4173 (2020). MR4102262.

Meyn, S. and Tweedie, R. L. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition (2009). ISBN 978-0-521-73182-9. MR2509253.

Peng, X. and Zhang, R. Exponential ergodicity for SDEs under the total variation. *J. Evol. Equ.*, **18** (3), 1051–1067 (2018). MR3859440.

Picard, J. On the existence of smooth densities for jump processes. *Probab. Theory Related Fields*, **105** (4), 481–511 (1996). MR1402654.

Romito, M. A simple method for the existence of a density for stochastic evolutions with rough coefficients. *Electron. J. Probab.*, **23**, Paper no. 113, 43 (2018). MR3885546.

Triebel, H. *Theory of function spaces. III*, volume 100 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel (2006). ISBN 978-3-7643-7581-2; 3-7643-7581-7. MR2250142.

Wang, J. On the exponential ergodicity of Lévy-driven Ornstein-Uhlenbeck processes. *J. Appl. Probab.*, **49** (4), 990–1004 (2012). MR3058984.

Zhang, X. and Glynn, P. W. Affine jump-diffusions: stochastic stability and limit theorems. *ArXiv Mathematics e-prints* (2018). arXiv: 1811.00122.