

Geometric Ergodicity of Affine Processes on Cones.

Eberhard Mayerhofer*, Robert Stelzer† and Johanna Vestweber‡

June 14, 2019

Abstract

For affine processes on finite-dimensional cones, we give criteria for geometric ergodicity - that is exponentially fast convergence to a unique stationary distribution. Ergodic results include both the existence of exponential moments of the limiting distribution, where we exploit the crucial affine property, and finite moments, where we invoke the polynomial property of affine semigroups. Furthermore, we elaborate sufficient conditions for aperiodicity and irreducibility. Our results are applicable to Wishart processes with jumps on the positive semidefinite matrices, continuous-time branching processes with immigration in high dimensions, and classical term-structure models for credit and interest rate risk.

AMS Subject Classification 2010:

Primary: 60J25 *Secondary:* 60G10, 60G51, 60J60, 60J75

Keywords: Affine process, geometric ergodicity, Feller process, Foster-Lyapunov drift condition, Harris recurrence, irreducibility, Lévy process, Wishart process.

1 Introduction

Affine processes are very popular, due to their flexibility and high analytical tractability: While they enjoy all classical features of general Markov processes in continuous time (e.g., state-dependent diffusive and jump behaviour), the Backward Kolmogorov PDEs simplify to non-linear ODEs for exponential initial data. On the canonical state spaces $\mathbb{R}_+^m \times \mathbb{R}^n$ they have been fully characterized by Duffie, Filipović and Schachermayer, in the seminal paper [28]. Particular affine models are well-known from financial applications (see e.g. [5, 14, 20, 27, 49]). They have been generalized to the time inhomogeneous case [33] and to other state spaces [15, 16, 18, 55, 80]. Of particular interest as state space are cones, especially the cone of positive semi-definite matrices (cf. [1, 16, 65, 69]). Such processes are used to model financial markets, where several assets exhibit stochastic (co)volatilities. The typical models derive stochastic volatility from the positive-definite Wishart process (see e.g. [2, 3, 12, 19, 22, 23, 42, 43, 66] and references therein), the Wishart

*Department of Mathematics & Statistics, University of Limerick, Limerick, V94 T9PX, Ireland; eberhard.mayerhofer@ul.ie

†Institute of Mathematical Finance, Ulm University, Helmholtzstrasse 18, 89075 Ulm, Germany; robert.stelzer@uni-ulm.de

‡Institute of Mathematical Finance, Ulm University, Helmholtzstrasse 18, 89075 Ulm, Germany; johanna2912@googlemail.com

model enriched with jumps ([10, 59, 61]) and the Ornstein-Uhlenbeck type stochastic volatility model (see e.g. [6, 7, 44, 74]). Wishart and general positive semi-definite affine models have also been successfully used to model interest rates (e.g. [8, 13, 41, 45]). For stochastic processes the existence and uniqueness of stationary solutions as well as convergence to the stationary solution is of high interest. Geometric ergodicity means that convergence to a unique stationary distribution happens in total variation exponentially fast. It ensures fast convergence to the stationary regime in simulations and paves the way for statistical inference. By the same argument as in [63, Proof of Theorem 4.3, Step 2] geometric ergodicity and the existence of some p -moments of the stationary distribution guarantee exponential β -mixing (and thus strong mixing) for Markov processes. This, in turn, implies functional central limit theorems for the process (see for instance [24]), and thus yields asymptotic normality of estimators.¹

In this paper we investigate geometric ergodicity of affine processes on proper, closed convex cones K in finite dimensional vector spaces, by using the stability concepts for Markov processes (cf. [25, 70–72]). In particular, we construct natural exponential and polynomial test functions to establish a Foster-Lyapunov drift condition, and we are able to characterize those affine processes which satisfy the latter condition in terms of a sufficiently fast decay of the linear drift term. To this end, we develop a new result for linear maps A with positive resolvent, i.e. $(\lambda - A)^{-1}(K) \subseteq K$, extending a result by Ky Fan, [31, Theorem 5'] for the natural cones \mathbb{R}_+^n . Thus, for aperiodic and irreducible affine processes, we find easy-to-check conditions for geometric ergodicity. For affine jump-diffusions on symmetric cones whose jumps are of compound-Poisson type with state-dependent jump intensity, we provide sufficient conditions for irreducibility and aperiodicity by relating them to diffusion processes killed at an exponential rate. Likewise, we establish easy to check sufficient conditions for irreducibility and aperiodicity in the finite activity pure jump case. Finally, we apply our general results to the special case of Wishart diffusions (with jumps) as well as the standard cone \mathbb{R}_+^m and consider a simple case for the general state space $K \times \mathbb{R}^n$, for which no general theory on affine processes yet exists.

The literature on geometric ergodicity for affine processes on cones is sparse and either limited to the canonical state space, or to Ornstein-Uhlenbeck type processes:

- For $K = \mathbb{R}_+$ [52, 53] prove Harris recurrence and exponential ergodicity for the basic affine jump-diffusion, which arises as a default intensity model in credit risk [26], and slight extensions. These processes may not have state-dependent jumps. The proofs of [52] are technical and use the explicit form of the transition density. The result is a special case of our theory (see Remark 5.6).
- On the state space $\mathbb{R}_+ \times \mathbb{R}$ [4] studies geometric ergodicity of a mean-reverting two-factor affine jump-diffusion process on a half-space. We give a generalization to affine diffusion models on $K \times \mathbb{R}^n$, where K is a proper closed convex cone (See Theorem 5.8 and Remark 5.9 (iv)).
- Ornstein-Uhlenbeck processes driven by Lévy noise belong to the affine class. They are cone-valued when driven by appropriate cone-valued Lévy processes and the drift ensures that the cone cannot be left. For these processes existence of and convergence to stationary distributions is well-understood (see [77] and [63] for geometric ergodicity).
- In the Wishart case only criteria for convergence in distribution to the stationary case and ergodicity, in the sense that laws of large numbers hold, are known (see [2]).

- While finalizing the present paper, [85] became available. This paper considers geometric ergodicity on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$. It requires finite activity jumps, strong solutions to the SDEs associated to affine processes and diagonal diffusion terms ensuring a kind of uniform ellipticity on the whole state space. If there are no state dependent jumps, it establishes positive Harris recurrence under logarithmic moment conditions and geometric ergodicity under p -th moment conditions with $p > 0$. In the case of state dependent jumps it requires at least first moment conditions.

The present paper is primarily dedicated to affine processes on high-dimensional cones, and we allow for general, possibly infinite variation, jump behaviour and state-dependent compensators. We do not need the affine processes to be strong solutions to SDEs. Moreover, we also obtain results on exponential moments which are particularly relevant and natural in an affine context. To ensure irreducibility and aperiodicity we use appropriate controllability conditions which are weaker than uniform-ellipticity like conditions.

Program of the paper

The preliminary Section 2 is divided into three subsections. In Section 2.1 we develop a characterization of linear maps whose eigenvalues have strictly negative parts (Theorem 2.7). This result plays a crucial part in the construction of suitable test functions for which the generator of affine processes satisfies the Foster-Lyapunov drift condition (Section 3). In Section 2.2 we recall the definition of affine processes on cones, as provided by [18], and summarize some properties. Section 2.3 recalls the definition of Harris recurrence and geometric ergodicity, and sufficient criteria for the latter property.

The main ergodicity results are developed in Section 3. We split this section into two parts, for exponentially affine test functions (Section 3.1, Theorem 3.1) and polynomial ones (Section 3.2, Theorem 3.6).

Section 4 provides criteria for aperiodicity and irreducibility for affine jump-diffusions with hypoelliptic main symbol and pure jump processes.

The final Section 5 is dedicated to apply the theory of Sections 3 and 4 to particularly relevant special cases: Wishart processes (with jumps) on the cone of positive semi-definite $d \times d$ matrices, affine processes on the natural cone \mathbb{R}_+^n , and affine term-structure models on $K \times \mathbb{R}^n$, where K is a proper closed convex cone. By attaching a linear space to the cone K , the state space is not proper anymore and thus falls outside the cones considered before.

2 Preparatory statements

2.1 Order preserving maps

Throughout the paper, we use the notation \mathbb{R}_+ for the non-negative real line, and \mathbb{C} for the complex numbers.

Let V be a finite dimensional linear space V with inner product $\langle \cdot, \cdot \rangle$. We identify, by virtue of Riesz' representation theorem, each element φ of the dual V^* by the unique $u \in V$ such that $\varphi(x) = \langle u, x \rangle$ for any $x \in V$. Hence we shall not distinguish between V and V^* in the following.

Let $K \subset V$ be a closed convex cone. We denote by \preceq the induced partial order, that is for $x, y \in V$, $x \preceq y$ if and only if $y - x \in K$. K is called proper, if $K \cap \{-K\} = \{0\}$, and generating, if $K - K = V$. The dual cone of K is defined by $K^* := \{u \in V \mid \langle u, x \rangle \geq 0 \text{ for all } x \in K\}$.

Lemma 2.1. *Let K be a generating proper closed convex cone. Then*

(i) K^* has non-empty interior $(K^*)^\circ$, and

$$(K^*)^\circ = \{u \in V \mid \langle u, x \rangle > 0, \quad \forall x \in K \setminus \{0\}\}. \quad (2.1)$$

(ii) K^* is a generating proper closed convex cone.

(iii) K has non-empty interior and

$$K^\circ = \{x \in V \mid \langle u, x \rangle > 0, \quad \forall u \in K^* \setminus \{0\}\}.$$

Proof. The first property holds in view of [32, Proposition I.1.4].

K^* is generating, because K is proper ([32, Corollary I.1.3]). Since $(K^*)^* = \overline{K}$, and K is a closed generating cone, it follows that K^* is proper (see [32, Corollary I.1.3]). Hence (ii) is proved.

Property (iii) is an application of Property (i) applied to the dual cone, taking into account that $(K^*)^* = \overline{K} = K$ and using Property (ii). \square

In dimensions higher than one, comparison arguments for solutions of ODEs only hold, when the vector fields have a special property, namely quasimonotonicity (besides enough regularity to allow uniqueness). We recall this property next:

Definition 2.2. *A map $A : V \supseteq U \rightarrow V$ is quasimonotone increasing (qmi) with respect to K , if*

$$x, y \in U, \quad y - x \in K, \quad v \in K^*, \quad \langle v, y - x \rangle = 0 \quad \Rightarrow \quad \langle v, A(y) - A(x) \rangle \geq 0. \quad (2.2)$$

Remark 2.3. (i) *Note that for linear maps $A : V \rightarrow V$, this property simplifies to*

$$x \in K, \quad v \in K^*, \quad \langle v, x \rangle = 0 \quad \Rightarrow \quad \langle v, Ax \rangle \geq 0.$$

(ii) *Clearly, a linear map $A : V \rightarrow V$ is qmi with respect to K if and only if A^\top is qmi with respect to K^* , where A^\top is the adjoint of A , defined by*

$$\langle x, A^\top v \rangle = \langle Ax, v \rangle, \quad x, y \in V.$$

For a linear map $A : V \rightarrow V$, let $\sigma(A)$ denote its spectrum. We use the notation $\tau(A) = \max\{\Re(\lambda), \lambda \in \sigma(A)\}$ for the spectral bound, $\rho(A) := \max\{|\lambda|, \lambda \in \sigma(A)\}$ for the spectral radius and e^{tA} for the matrix exponential. Clearly $\tau(A) = \tau(A^\top)$, because A and A^\top have the same eigenvalues. Furthermore, $\lambda > \tau(A)$ implies that $\lambda \notin \sigma(A)$, whence $\lambda - A$ is an isomorphism. The following is a fundamental characterization of qmi maps in terms of their resolvent:

Proposition 2.4. *Let $A : V \rightarrow V$ be a linear map, and K be a generating, proper closed convex cone. The following are equivalent:*

(i) A is qmi with respect to K .

(ii) $e^{tA}(K) \subseteq K$.

(iii) For any $\lambda > \tau(A)$, $(\lambda - A)^{-1}(K) \subseteq K$.

(iv) For any $\lambda > \tau(A)$, $(\lambda - A^\top)^{-1}(K^*) \subseteq K^*$.

Proof. The equivalence of the first three conditions follows from [29, Satz 1, (2), (3) and (5), $\alpha = 1$]. Note that [29, (3)] is our definition of qmi, whereas [29] uses an equivalent one. Furthermore, the statement was shown for matrices, but we use the obvious modification for linear maps.

By Remark 2.3, A is qmi with respect to K , if and only if A^\top is qmi with respect to K^* . Therefore, statement (iv) is equivalent to (i), and can be proved similarly, by replacing the role of K and K^* , realizing that due to Lemma 2.1 the dual cone K^* has non-empty interior and thus K^* satisfies the standing assumption (e) of [29]. \square

Another ingredient in our main statement of this section (Theorem 2.7) is the celebrated theorem by Krein-Rutman [57]:

Theorem 2.5. *Let X be a Banach space and $C \subset X$ be a convex cone such that $C - C$ is dense in X . Let $T : X \rightarrow X$ be a compact operator which is positive, meaning $T(C) \subseteq C$, and assume that its spectral radius $\rho(T)$ is strictly positive. Then there exists $x \in C \setminus \{0\}$ such that $Tx = \rho(T)x$.*

We require an elementary lemma:

Lemma 2.6. *Let $A : V \rightarrow V$ be qmi with respect to a generating, proper closed convex cone. Then $\tau(A) = \max\{\Re(\lambda) \mid \lambda \in \sigma(A)\}$ is an eigenvalue of A .*

Proof. Assume, for a contradiction, $\tau(A) \notin \sigma(A)$. Then there exist $r \in \mathbb{N}, \mu_1, \dots, \mu_r \neq 0$ such that

$$\sigma_A^T := \sigma(A) \cap \{\tau(A) + i\mathbb{R}\} = \{\tau(A) + i\mu_i, \quad 1 \leq i \leq r\}.$$

Hence there exists $t_0 > 0$ such that $t_0\mu_i \notin \mathbb{Z}$ for any $1 \leq i \leq r$. By the Jordan normal form of A , all eigenvalues of $P := e^{t_0 A}$ are of the form $e^{t_0\lambda}$, where $\lambda \in \sigma(A)$. Therefore, the spectral radius of P equals $\rho_P := e^{t_0\tau(A)}$, but since $e^{t_0(\tau(A)+i\mu_i)} \neq e^{t_0\tau(A)} = \rho_P$ for any $1 \leq i \leq r$, the spectral radius is not an eigenvalue of P . On the other hand, since A is qmi, P is a positive map, hence by Theorem 2.5 $\rho(P)$ must be an eigenvalue, a contradiction. We conclude that indeed $\tau(A) \in \sigma(A)$. \square

The main statement of this section is the following extension of [31, Theorem 5', (a), (c) for $K = \mathbb{R}_+^n$] to general cones K :

Theorem 2.7. *Let K be a generating, proper closed convex cone, and $A : V \rightarrow V$ be a qmi linear map with respect to K . The following are equivalent:*

- (i) $\tau(A) < 0$.
- (ii) *There exists $v \in (K^*)^\circ$ such that $A^\top v \in -(K^*)^\circ$.*

Proof. (i) \Rightarrow (ii):

A^\top is qmi with respect to K^* , and clearly $\tau(A) = \tau(A^\top)$. Pick $\tau(A) < \lambda < 0$. By Proposition 2.4, $(\lambda - A^\top)^{-1}(K^*) \subseteq K^*$, hence by Lemma 2.1, there exists $v \in (K^*)^\circ$ such that $\eta^* := (\lambda - A^\top)^{-1}v \in K^*$, and clearly $v \neq 0$. Therefore, $\lambda\eta^* - A^\top\eta^* = (\lambda - A^\top)\eta^* = v$. We conclude that $-A^\top\eta^* = v - \lambda\eta^* \in K^* + (K^*)^\circ = (K^*)^\circ$.

(ii) \Rightarrow (i):

Pick a $\lambda > \tau(A)$. By Proposition 2.4, the operator $P := (\lambda - A)^{-1}$ is positive on K , and since 0 is not an eigenvalue of a linear isomorphism, the spectral radius of P is strictly larger than 0. Hence

by Theorem 2.5, there exists $x \in K \setminus \{0\}$ such that $(\lambda - A)^{-1}x = \rho_\lambda x$, and $\rho_\lambda := \max\{|\mu| \mid \mu \in \sigma(\lambda - A)^{-1}\}$. A straightforward check reveals that $\tau := \lambda - \frac{1}{\rho_\lambda} \in \sigma(A)$ and $Ax = \tau x$. Furthermore,

$$\tau = \lambda - \frac{1}{\rho_\lambda} = \max\{\lambda - |\lambda - \mu| \mid \mu \in \sigma(A)\}.$$

The two functions $\sigma(A) \rightarrow \mathbb{R}$,

$$\mu \mapsto \lambda - |\lambda - \mu|, \quad \mu \mapsto \Re(\mu),$$

agree on the real line, and the first one has its maximum at the real value $\tau \in \sigma(A)$, while the second one has its maximum at $\tau(A) \in \sigma(A)$ (see Lemma 2.6). Therefore the maxima of the two functions agree, i.e.,

$$\begin{aligned} \tau &= \lambda - \frac{1}{\rho_\lambda} = \max\{\lambda - |\lambda - \mu| \mid \mu \in \sigma(A)\} \\ &= \max\{\lambda - |\lambda - \mu| \mid \mu \in \sigma(A) \cap \mathbb{R}\} = \max\{\Re(\mu) \mid \mu \in \sigma(A) \cap \mathbb{R}\} \\ &= \max\{\Re(\mu) \mid \mu \in \sigma(A)\} = \tau(A). \end{aligned}$$

It thus remains to show that $\tau(A) < 0$: To this end, let $v \in (K^*)^\circ$ such that $A^\top v \in -(K^*)^\circ$. Then

$$\langle A^\top v, x \rangle = \langle v, Ax \rangle = \langle v, \tau(A)x \rangle = \tau(A)\langle v, x \rangle.$$

Since $\langle v, x \rangle > 0$ and $\langle A^\top v, x \rangle < 0$ due to Lemma 2.1, $\tau(A)$ must be strictly negative, and we are done. \square

2.2 Affine processes.

In the following we give the definition of affine processes on cones, as provided by [18], and summarize some properties. The state space K will be a generating, proper closed convex cone in a finite dimensional vector space V , as in the previous section.² Adjoined to the state space K is a point $\Delta \notin K$, the cemetery state, and we set $K_\Delta = K \cup \{\Delta\}$ for the one-point compactification. For a Markov process X on K , its transition function is denoted by $(p_t(x, \cdot))_{t \geq 0, x \in K}$. We can extend to K_Δ by introducing

$$p_t(x, \{\Delta\}) = 1 - p_t(x, K), \quad p_t(\Delta, \{\Delta\}) = 1,$$

for all $t \in \mathbb{R}_+$ and $x \in K$, with the convention $f(\Delta) = 0$ for any function f on K .

Definition 2.8. *A time-homogeneous Markov process X relative to some filtration (\mathcal{F}_t) with state space K (augmented by Δ) and transition kernels $(p_t(x, d\xi))_{t \geq 0, x \in K}$ is called affine if*

- (i) *it is stochastically continuous, that is, $\lim_{s \rightarrow t} p_s(x, \cdot) = p_t(x, \cdot)$ weakly on K for every $t \geq 0$ and $x \in K$, and*
- (ii) *its Laplace transform has exponential-affine dependence on the initial state. This means that there exist functions $\phi : \mathbb{R}_+ \times K^* \rightarrow \mathbb{R}_+$ and $\psi : \mathbb{R}_+ \times K^* \rightarrow V$ such that*

$$\int_K e^{-\langle u, \xi \rangle} p_t(x, d\xi) = e^{-\phi(t, u) - \langle \psi(t, u), x \rangle}, \quad (2.3)$$

for all $x \in K$ and $(t, u) \in \mathbb{R}_+ \times K^*$.

The following statement recalls the basic properties of affine processes on cones, in particular the description of (ϕ, ψ) as solutions for so-called generalized Riccati differential equations (cf. [18, Theorem 2.4, Part I])

Theorem 2.9. *Let X be an affine process on K . Then X is a Feller process, the functions ϕ and ψ given in (2.3) are differentiable with respect to time and satisfy the generalized Riccati equations for $u \in K^*$, that is,*

$$\frac{\partial \phi(t, u)}{\partial t} = F(\psi(t, u)), \quad \phi(0, u) = 0, \quad (2.4a)$$

$$\frac{\partial \psi(t, u)}{\partial t} = R(\psi(t, u)), \quad \psi(0, u) = u \in K^*, \quad (2.4b)$$

where $F(u) = \partial_t \phi(t, u)|_{t=0}$ and $R(u) = \partial_t \psi(t, u)|_{t=0}$. Moreover, relative to some truncation function³ χ , there exists a parameter set $(Q, b, B, c, \gamma, m, \mu)$ such that the functions F and R are of the form

$$F(u) = \langle b, u \rangle + c - \int_K \left(e^{-\langle u, \xi \rangle} - 1 \right) m(d\xi), \quad (2.5a)$$

$$R(u) = -\frac{1}{2}Q(u, u) + B^\top(u) + \gamma - \int_K \left(e^{-\langle u, \xi \rangle} - 1 + \langle \chi(\xi), u \rangle \right) \mu(d\xi), \quad (2.5b)$$

where

(i) $b \in K$,

(ii) $c \in \mathbb{R}_+$,

(iii) m is a Borel measure on K satisfying $m(\{0\}) = 0$ and

$$\int_K (\|\xi\| \wedge 1) m(d\xi) < \infty,$$

(iv) $Q : V \times V \rightarrow V$ is a symmetric bilinear function such that for all $v \in V$, $Q(v, v) \in K^*$ and $\langle x, Q(u, v) \rangle = 0$, whenever $\langle u, x \rangle = 0$ for $u \in K^*$ and $x \in K$,

(v) $\gamma \in K^*$,

(vi) μ is a K^* -valued σ -finite Borel measure on K satisfying $\mu(\{0\}) = 0$, $\int_K (\|\xi\|^2 \wedge 1) \langle x, \mu(d\xi) \rangle < \infty$ for all $x \in K$, and

$$\int_K \langle \chi(\xi), u \rangle \langle x, \mu(d\xi) \rangle < \infty \text{ for all } u \in K^* \text{ and } x \in K \text{ with } \langle u, x \rangle = 0,$$

(vii) $B^\top : V \rightarrow V$ is a linear map, satisfying

$$\langle x, B^\top(u) \rangle - \int_K \langle \chi(\xi), u \rangle \langle x, \mu(d\xi) \rangle \geq 0 \text{ for all } u \in K^* \text{ and } x \in K \text{ with } \langle u, x \rangle = 0.$$

By the defining property (2.3) and the generalized Riccati differential equations (2.4), for any $x \in K$, and $u \in K^*$ we have for $f_u(x) := e^{-\langle u, x \rangle}$ that

$$\lim_{t \downarrow 0} \frac{P_t f_u(x) - f_u(x)}{t} = (-F(u) - \langle R(u), x \rangle) f_u(x). \quad (2.6)$$

This limit also holds in supremum norm, due to the Feller property of X (we shall denote by $\mathcal{D}(\mathcal{A})$ the domain of the infinitesimal generator \mathcal{A} of X , that is the generator of the associated Feller semigroup).

Proposition 2.10. *For any $u \in (K^*)^\circ$, we have $f_u \in \mathcal{D}(\mathcal{A})$, and*

$$\mathcal{A}f_u(x) = (-F(u) - \langle R(u), x \rangle) f_u(x). \quad (2.7)$$

Proof. The statement is a slight adaption of the proof of [16, Proposition 4.12] to cones K , the only difference being that our cones are not necessarily symmetric⁴: By compactness of the unit sphere $\{x \in K \mid \|x\| = 1\}$, for any $u \in (K^*)^\circ \subset K^*$, there exists a constant $\lambda_u > 0$ such that $\langle u, x \rangle \geq \lambda_u \|x\|$ for any $x \in K$. Hence $f_u \in C_0(K)$. The rest of the proof is exactly the same as [16, Proposition 4.12]. \square

We end this section by giving a characterization of conservative affine processes, and an easy-to-check sufficient condition for non-explosion.⁵

Proposition 2.11. *Let X be an affine process on K . The following are equivalent:*

- (i) X is conservative.
- (ii) $c = 0$, $\gamma = 0$, and 0 is the only K^* -valued solution of the generalized Riccati differential equation

$$\partial_t \psi(t) = R(\psi(t)), \quad \psi(0) = 0. \quad (2.8)$$

In particular, if $c = 0$, $\gamma = 0$ and for some $x_0 \in K^\circ$,

$$\int_{\|\xi\| \geq 1} \|\xi\| \langle \mu(d\xi), x_0 \rangle < \infty, \quad (2.9)$$

then (ii) is satisfied, and thus X is non-explosive.

Proof. The characterization (i) \Leftrightarrow (ii) can be proved along the lines [68, Theorem 3.4], using the quasi-monotonicity of the function $u \rightarrow R(u)$ with respect to K^* (see [18, Proposition 3.12]).

Note that the positive measure $(\|\xi\|^2 \wedge 1) \langle \mu(d\xi), x_0 \rangle$ is finite, whence by [28, Lemma A.1 for $k = 1$] condition (2.9) implies that

$$\varphi_x : V \rightarrow \mathbb{C} : \quad v \mapsto \langle R(iv), x \rangle$$

is C^1 for $x = x_0$, and thus in particular, locally Lipschitz. Since $x_0 \in K^\circ$, for any $x \in K$ there exists $\lambda_x > 0$ such that $x \preceq \lambda_x x_0$, whence (2.9) indeed holds for any $x \in K$, and therefore φ_x is locally Lipschitz, for any $x \in K$. Since $K - K = V$ we infer that $V \rightarrow \mathbb{C}, v \mapsto R(iv)$ is locally Lipschitz. Suppose, $\psi(t)$ is a K^* -valued solution of (2.8). Then $\chi(t) = -i\psi(t)$ solves the complex-valued generalized Riccati differential equation

$$\partial_t \chi = R(i\chi), \quad \chi(0) = 0.$$

By local Lipschitz continuity of R on iV , $\chi \equiv 0$, thus $\psi \equiv 0$. Therefore, (ii) is satisfied and thus X is conservative.⁶ \square

2.3 Markov processes and ergodicity

In this section, $X = (X_t)_{t \geq 0}$ is a continuous-time Markov process (please see the standard literature like [25, 71–73] for definitions of notions for Markov processes not explained here) on a locally compact, separable metric space D with transition probabilities $p_t(x, A) = \mathbb{P}^x[X_t \in A]$ for $x \in D, A \in \mathcal{B}(D)$. We assume that X is a non-explosive Borel right process. We denote with π the unique invariant measure of X , if it exists. Exponential ergodicity means that the total variation distance of the transition semigroup and the invariant measure converge to zero independent of the initial state x and with an exponential rate. A “seemingly stronger” concept⁷ of this is the f -uniform ergodicity, which demands the convergence in the f -norm⁸ instead of the total variation norm:

Definition 2.12 (f -uniform ergodicity, [25, Chapter 3]). *X is called f -uniformly ergodic, if a measurable function $f : D \rightarrow [1, \infty)$ exists such that for all $x \in D$*

$$\|p_t(x, \cdot) - \pi\|_f \leq f(x)C\rho^t, \quad t \geq 0 \quad (2.10)$$

holds for some $C < \infty, \rho < 1$.

Another concept is a probabilistic form of stability, the so called Harris recurrence.

Definition 2.13 (Harris recurrence, [71, Chapter 2.2]). *Let $\eta_A := \int_0^\infty 1_{\{X_t \in B\}} dt$ be the occupation time and $\tau_A := \inf\{t \geq 0 : X_t \in A\}$ be the first hitting time of A .*

(i) *X is called Harris recurrent, if either*

- $\mathbb{P}^x[\eta_A = \infty] = 1$ whenever $\phi(A) > 0$ for some σ -finite measure ϕ , or
- $\mathbb{P}^x[\tau_A < \infty] = 1$ whenever $\mu(A) > 0$ for some σ -finite measure μ .

(ii) *Suppose that X is Harris recurrent with finite invariant measure π , then X is called positive Harris recurrent.*

Definition 2.14 ([25], Chapter 3). *For a σ -finite measure μ on $\mathcal{B}(D)$ we call the process X μ -irreducible if for any $A \in \mathcal{B}(X)$ with $\mu(A) > 0$*

$$\mathbb{E}^x(\eta_A) > 0, \quad \forall x \in X. \quad (2.11)$$

This is obviously the same as requiring $\int_0^\infty \mathbb{P}^x(X_t \in B) dt > 0, \forall x \in X$. If X is μ -irreducible, there exists a maximal irreducibility measure ψ such that every other irreducibility measure ν is absolutely continuous with respect to ψ (see [71, p. 493]). We write $\mathcal{B}^+(D)$ for the collection of all measurable subsets $A \in \mathcal{B}(X)$ with $\psi(A) > 0$.

Definition 2.15 ([25], Chapter 3). (i) *A nonempty set $C \in \mathcal{B}(D)$ is called small, if there exists an $m > 0$ and a nontrivial measure ν_m on $\mathcal{B}(D)$ such that for all $x \in C, A \in \mathcal{B}(D)$*

$$\mathbb{P}^x[X_t \in A] \geq \nu_m(A). \quad (2.12)$$

(ii) *A ψ -irreducible Markov process is called aperiodic if for some small set $C \in \mathcal{B}^+(X)$ there exists a T such that $\mathbb{P}^x[X_t \in C] > 0$ for all $t \geq T$ and all $x \in C$.*

For irreducible and aperiodic Markov processes a sufficient condition for f -uniform ergodicity is the so-called Foster-Lyapunov drift condition:

Theorem 2.16 ([25], Theorem 5.2). *Let $(X_t)_{t \geq 0}$ be a μ -irreducible and aperiodic Markov process. If there exist constants $b, c > 0$ and a petite set C in $\mathcal{B}(D)$ as well as a measurable function $f : D \rightarrow [1, \infty)$ (referred to as Lyapunov function) such that*

$$\mathcal{A}f \leq -bf + c1_C, \quad (2.13)$$

where \mathcal{A} is the extended generator, then X is f -uniformly ergodic.

Definition 2.17 (Extended generator, [25, Chapter 4]). $\mathcal{D}(\mathcal{A})$ denotes the set of all functions $f : D \times \mathbb{R}_+ \rightarrow \mathbb{R}$ for which a function $g : D \times \mathbb{R}_+ \rightarrow \mathbb{R}$ exists, such that $\forall x \in D, t > 0$

$$\mathbb{E}^x(f(X_t, t)) = f(x, 0) + \mathbb{E}^x \left(\int_0^t g(X_s, s) ds \right), \quad (2.14)$$

$$\mathbb{E}^x \left(\int_0^t |g(X_s, s)| ds \right) < \infty \quad (2.15)$$

holds. We write $\mathcal{A}f := g$ and call \mathcal{A} the extended generator of X . $\mathcal{D}(\mathcal{A})$ is called the domain of \mathcal{A} .

The above definition of the extended generator is the usual one in the papers on geometric ergodicity of Markov processes. Note that in other strands of the literature the definition is often different. For example, many papers only demand that $f(X_t, t) - f(x, 0) - \int_0^t g(X_s, s) ds$ is a local martingale. This is the case e.g. in [17]. Their Theorem 2.7 and the preceding discussion, however, ensures that their extended generator for m -polynomial processes with $m \geq 2$ is also the extended generator in our sense and all polynomials of order at most m are in its domain.

For some $\Delta > 0$ we call the Markov chain $(X_{k\Delta})_{k \in \mathbb{N}}$ a (Δ) -skeleton chain.

We skip the definition of a petite set since we use the following result, which ensures petiteness of compact sets (note that only irreducibility is assumed, but not aperiodicity):

Proposition 2.18 ([73, Theorem 6.0.1]). *If Y is a Ψ -irreducible Feller chain with $\text{supp}(\Psi)$ having non-empty interior, every compact set is petite.*

3 Geometric ergodicity of affine processes

The following two sections offer two different sets of sufficient conditions for geometric ergodicity of affine processes on K . The main difference between the two statements lies in the assumption of finite moments in the second one, instead of exponential ones in the first one; therefore, also the conclusions concerning the moments of the limiting distributions are different.

Due to the characterization of affine processes via the exponential-affine form of the Fourier/Laplace transform, considering exponential moments is particularly natural and relevant.

3.1 Exponential moments

Theorem 3.1. *Let X be an affine process on K with $c = 0, \gamma = 0$. Assume that*

- (i) X is ν -irreducible with the support of ν having non-empty interior and aperiodic,

(ii) there exists $\eta_0 \in (K^*)^\circ$ such that

$$B^\top \eta_0 + \int_{K \setminus \{0\}} \langle \eta_0, \xi - \chi(\xi) \rangle \mu(d\xi) \in -(K^*)^\circ \quad (3.1)$$

and for all $t \geq 0$, and for some $x \in K^\circ$, $\mathbb{E}^x[e^{\langle \eta_0, X_t \rangle}] < \infty$.

Then X is conservative, and for any $\eta \preceq \eta_0$, and all $x \in K$

$$\int_{\{\|\xi\| \geq 1\}} e^{\langle \eta, \xi \rangle} (m(d\xi) + \langle \mu(d\xi), x \rangle) < \infty \quad (3.2)$$

and X is geometrically ergodic and positive Harris recurrent. Furthermore, the stationary distribution π has finite exponential moments which can be computed with the affine transform formula, i.e. there exists $h_0 > 0$ such that for all $v \preceq h_0 \eta_0$,

$$\int_{K \setminus \{0\}} e^{\langle v, \xi \rangle} \pi(d\xi) = e^{-\int_0^\infty F(\psi(s, -v)) ds}. \quad (3.3)$$

Proof. Since η_0 is in the interior of the dual cone, Condition (3.1) implies that Condition (2.9) of Proposition 2.11 is satisfied. Hence by Proposition 2.11, X is conservative. Since $\mathbb{E}^x[e^{\langle \eta_0, X_t \rangle}] < \infty$, the monotonicity of the exponential implies $\mathbb{E}^x[e^{\langle \eta, X_t \rangle}] < \infty$ for any $\eta \preceq \eta_0$. We conclude by [54, Theorem 2.14 (a)] that (3.2) is satisfied for any $\eta \preceq \eta_0$ and that for any $x \in K$, any $\eta \preceq \eta_0$, and any $t \geq 0$

$$\mathbb{E}^x[e^{\langle \eta, X_t \rangle}] = e^{p(t, \eta) + \langle q(t, \eta), x \rangle}, \quad (3.4)$$

where q is a (not necessarily unique⁹) solution of the generalized Riccati differential equation $\partial_t q = -R(-q)$ with $q(0) = \eta$ and $\partial_t p = -F(-q)$, $p(0) = 0$. In particular, $p(t, \eta) = -\int_0^t F(-q(s, \eta)) ds$.¹⁰

Now $u \mapsto R(u)$ and $u \mapsto F(u)$ are analytic functions in a neighborhood of 0, due to (3.2) and their explicit definitions, see (2.5). Pick some $x_0 \in K^\circ$. It follows directly from the parametric restrictions (see Theorem 2.9) that the map $u \mapsto -R(-u)$ is qmi (with respect to the dual cone K^*) on the domain

$$\mathcal{U} := \left\{ u \in V \mid \int_{\{\|x\| \geq 1\}} e^{\langle u, \xi \rangle} (m(d\xi) + \langle \mu(d\xi), x_0 \rangle) < \infty \right\}.^{11}$$

Furthermore, by assumption (3.1),

$$\frac{\partial R}{\partial u} \Big|_{u=0}(\eta_0) = B^\top \eta_0 + \int_{K \setminus \{0\}} \langle \eta_0, \xi - \chi(\xi) \rangle \mu(d\xi) \in -(K^*)^\circ.$$

Thus Theorem 2.7 is applicable and guarantees that $\tau(\frac{\partial R}{\partial u} \Big|_{u=0}) < 0$. The latter implies by a well known result of Lyapunov ([60]) that 0 is an asymptotically stable equilibrium of the ODE $\partial_t q = -R(-q)$. In fact, there exists $\delta > 0$ such that $B_\delta(0) := \{u \in V \mid \|u\| \leq \delta\} \subset \mathcal{U}^\circ$ (where R is analytic). Hence, for any $u \in B_\delta(0)$ the (unique) solution $q(t, u)$ exists for all times $t \geq 0$ and satisfies $\delta \geq \|q(t, u)\| \rightarrow 0$, as $t \rightarrow \infty$. Furthermore, it is geometrically stable (see e.g. [83, Theorem 7.1]), that is, for all $\|u\| \leq \delta$, and all $t \geq 0$, $\|q(t, u)\| \leq ce^{-dt}$ for some constants $c, d > 0$. Therefore also $p(t, v) \rightarrow \int_0^\infty -F(-q(s, u)) ds$ as $t \rightarrow \infty$. Therefore, for $\eta \in B_\delta(0)$, the affine transform formula (3.4) holds, where (p, q) is the unique, global solution of the generalized Riccati equations. Due to

asymptotic stationarity of the ODE, we have therefore shown (3.3) for $v \in B_\delta(0)$ (this fact is used next, while the extension to any $v \preceq h\eta_0$ is done in the last paragraph of this proof).

By Taylor's formula, we have $-R(-h\eta_0) = hDR_{\eta_0}(0) + o(h^2)$, hence there exists $0 < h_0 < 1$ such that for $\eta = h_0\eta_0$, $-R(-\eta) \in -(K^*)^\circ$. Pick this h_0 sufficiently small such that also $h_0\eta_0 \in B_\delta(0)$, and set $\eta := h_0\eta_0$. Since η is in the interior of the dual cone, the map $H_\eta : K \rightarrow [1, \infty)$, $H_\eta(x) := e^{\langle \eta, x \rangle}$ has range equal to $[1, \infty)$. Since by assumption, $\mathbb{E}^x[H_{\eta_0}(X_t)] < \infty$, for any $t > 0$, also $\mathbb{E}^x[H_\eta(X_t)] < \infty$ and since $\eta_0 \in (K^*)^\circ$, there exists a constant $L > 0$ such that for all $x \in K$

$$\|x\|e^{\langle \eta, x \rangle} \leq Le^{\langle \eta_0, x \rangle},$$

which implies that

$$\mathbb{E}^x \left[\int_0^t H_\eta(X_s) |F(-\eta) + \langle R(-\eta), X_s \rangle| ds \right] < \infty.$$

Therefore, to show that H_η belongs to the domain of the extended generator, it suffices to establish that

$$\mathbb{E}^x[H_\eta(X_t)] = H_\eta(x) + \mathbb{E}^x \left[\int_0^t H_\eta(X_s) (-F(-\eta) + \langle -R(-\eta), X_s \rangle) ds \right].$$

Taking into account the affine transform formula (3.4), it is therefore enough to prove that for any $u \in B_\delta(0)$

$$e^{-\phi(t,u) - \langle \psi(t,u), x \rangle} - H_{-u}(x) = \mathbb{E}^x \left[\int_0^t H_{-u}(X_s) (-F(u) - \langle R(u), X_s \rangle) ds \right]. \quad (3.5)$$

Note that both left and right side of (3.5) are analytic functions on $B_\delta(0)$. Since any open set is a set of uniqueness for analytic functions, the analytic functions on the left and right sides of (3.5) must agree, if they agree only on $B_\delta(0) \cap (K^*)^\circ$ (cf. [21, (9.4.2)]). And indeed, the identity (3.5) holds for any $u \in (K^*)^\circ$, because in this case $H_{-u} \in \mathcal{D}(\mathcal{A})$, which first implies¹² that $H_{-\psi(s,u)} \in \mathcal{D}(\mathcal{A})$ for any $s \geq 0$, and as the semigroup action commutes with the action of the infinitesimal generator ([30, Proposition 1.5]), we have

$$\begin{aligned} -F(\psi(s,u) - \langle R(\psi(s,u), x) \rangle e^{\phi(s,u)} H_{\psi(s,u)} x) &= \mathcal{A}(\mathbb{E}^x[H_{-u}(X_s)]) = \mathbb{E}^x[\mathcal{A}H_{-u}(X_s)] \\ &= \mathbb{E}^x[(-F(u) - \langle R(u), X_s \rangle)H_{-u}(X_s)]. \end{aligned} \quad (3.6)$$

Using the fundamental theorem of calculus, the generalized Riccati differential equations (2.4), identity (3.6) and Fubini's theorem, we thus have for any $u \in (K^*)^\circ$,

$$\begin{aligned} e^{-\phi(t,u) - \langle \psi(t,u), x \rangle} - H_{-u}(x) &= \int_0^t (-\partial_s \phi(s,u) - \langle \partial_s \psi(s,u), x \rangle) e^{-\phi(s,u) - \langle \psi(s,u), x \rangle} ds \\ &= - \int_0^t (F(\psi(s,u) - \langle R(\psi(s,u), x) \rangle) e^{-\phi(s,u) - \langle \psi(s,u), x \rangle} ds \\ &= -\mathbb{E}^x \left[\int_0^t (F(u) + \langle R(u), X_s \rangle) H_{-u}(X_s) ds \right]. \end{aligned}$$

which is (3.5). This finishes the proof of H_η being in the domain of the extended generator.

Since $-R(-\eta) \in -(K^*)^\circ$, there exists $\lambda_0 > 0$ such that $\langle R(-\eta), x \rangle \geq \lambda_0 \|x\|$ (because, the continuous map $x \mapsto \langle R(-\eta), x \rangle$ attains its minimum over the set $\{x \in K \mid \|x\| = 1\}$). Hence,

$$\mathcal{A}H_\eta(x) \leq H_\eta(x)(-F(-\eta) - \lambda_0 \|x\|).$$

Select $k > 0$ such that $-F(-\eta) - \lambda_0 k = -c < 0$. Then for any $x \in K \setminus B$, where $B := \{x \in K \mid \|x\| \leq k\}$ we have

$$\mathcal{A}H_\eta(x) \leq -cH_\eta(x),$$

and since $\mathcal{A}H_\eta$ is bounded on the compact set B , there exists a constant $d > 0$ such that

$$\mathcal{A}H_\eta(x) \leq -cH_\eta(x) + d1_K(x),$$

that is, the Foster-Lyapunov drift condition is satisfied for H_η , because B is a compact set in K and therefore by Proposition 2.18 petite.

We continue the proof by applying Theorem 2.16: By assumption X is irreducible and aperiodic and we have shown that the Foster-Lyapunov drift condition holds. Further we know that X is a Feller process, see Theorem 2.9, hence it is a Borel right process, see [62]. Since X is conservative by assumption it is non-explosive. Thus all assumptions of Theorem 2.16 are fulfilled and thus X is H_η -uniformly ergodic (whence, in particular geometrically ergodic).

Next we show that X is positive Harris recurrent and that the stationary distribution has the demanded moments. For discrete time Markov chains in [9, Theorem 3.12] it is shown that the Foster-Lyapunov drift condition for (\mathcal{A}, H) implies besides geometric ergodicity also positive Harris recurrence and for the stationary distribution $\pi(H) < \infty$. Applying [25, Theorem 5.1] we get that the proven Foster-Lyapunov drift condition implies the discrete version for every skeleton chain of X . By the definition of positive Harris recurrence it holds also for the process X and with the chosen test functions H_η in the proof above the η -th exponential moment of the stationary distribution π exists.

It remains to prove (3.3) (we have shown above that it holds for $v \in B_\delta(0)$, and since $\eta = h_0\eta_0 \in B_\delta(0)$, it holds for η , in particular). Since X is H_η -uniformly ergodic, we know that for any continuous function f satisfying $|f| \leq H_\eta$, $\mathbb{E}^x[f(X_t)] \rightarrow \int f(\xi)\pi(d\xi)$, as $t \rightarrow \infty$, where π is the unique stationary limit of X . Since for any $v \preceq h_0\eta_0 = \eta$ we have $f_v(x) = \exp(\langle v, x \rangle) \leq H_\eta(x)$, it follows that for any $v \preceq h_0\eta_0$

$$e^{p(t,v) + \langle x, q(t,v) \rangle} = \mathbb{E}^x[e^{\langle v, X_t \rangle}] \rightarrow \int_K e^{\langle v, \xi \rangle} \pi(d\xi), \quad \text{as } t \rightarrow \infty.$$

Since the right hand side does not depend on the initial state x , we see $e^{\langle x, q(t,v) \rangle} \rightarrow 1$, i.e. $\langle x, q(t,v) \rangle \rightarrow 0$ as $t \rightarrow \infty$, for any $x \in K$. Since K is generating ($K - K = V$), and the inner product is continuous, we obtain $\lim_{t \rightarrow \infty} q(t,v) = 0$. Thus $e^{p(t,v)} = e^{\int_0^t -F(-q(s,v))ds}$ converges, as $s \rightarrow \infty$. Since the Laplace transform of the measure π must be strictly positive, the limit is non-zero, and thus $\lim_{t \rightarrow \infty} e^{p(t,v)} = e^{\lim_{t \rightarrow \infty} p(t,v)} = e^{\int_0^\infty -F(-q(t,v))dt}$. \square

The assumption of finite exponential moments for all times in Theorem 3.1 is implied by the existence of solutions to the generalized Riccati equations for all times, i.e. no finite blow-up ([54, Theorem 2.14 (b)]). Therefore we can restate the Theorem as follows:

Corollary 3.2. *Let X be a an affine process on K with $c = 0$, $\gamma = 0$. Assume that*

- (i) *X is ν -irreducible with the support of ν having non-empty interior and aperiodic,*
- (ii) *There exists $\eta_0 \in (K^*)^\circ$ such that*

$$B^\top \eta_0 + \int_{K \setminus \{0\}} \langle \eta_0, \xi - \chi(\xi) \rangle \mu(d\xi) \in -(K^*)^\circ \quad (3.7)$$

and assume that

$$\int_{\{\|x\| \geq 1\}} e^{\langle \eta_0, \xi \rangle} (m(d\xi) + \langle \mu(d\xi), x \rangle) < \infty \quad (3.8)$$

for some $x \in K^\circ$.

Then X is conservative, (3.8) holds for any $\eta \preceq \eta_0$, and all $x \in K$ and X is geometrically ergodic and positive Harris recurrent. Furthermore, the stationary distribution π has finite exponential moments which can be computed with the affine transform formula, i.e. there exists $h_0 > 0$ such that for all $v \preceq h_0 \eta_0$

$$\int_{K \setminus \{0\}} e^{\langle v, \xi \rangle} \pi(d\xi) = e^{-\int_0^\infty F(\psi(s, -v)) ds}.$$

Proof. By monotonicity of the exponential, of course (3.8) also holds for any $\eta \preceq \eta_0$. Therefore it holds in a neighborhood of the origin. Hence the generalized Riccati differential equations are well-posed for initial data near zero. Furthermore, by (3.7) and Theorem 2.7, the Frechet differential of $R(u)$ at zero has negative spectral bound τ . Therefore for small enough initial data, the generalized Riccati differential equations have global solutions which tend to zero as $t \rightarrow \infty$. According to [54, Theorem 2.14 (b)], for any $x \in K$, $\mathbb{E}^x[e^{\langle \eta, X_t \rangle}] < \infty$ for initial data η small enough. Thus the assumptions of Theorem 3.1 are satisfied. \square

Remark 3.3. (i) *It is easy to see that condition (3.1) is independent of the choice of the truncation function. The appearance of the truncation function implies that we can indeed cover cases with not only infinite activity, but even infinite variation state dependent jumps.*

(ii) *The drift condition (3.1) involves the linear drift and the state dependent jumps.*

(iii) *The state-independent jumps do not matter at all, (3.1) and (3.7) only imply that they need to have finite exponential moments. In the case of an affine diffusion the Assumption (ii) of Corollary 3.2 boils down to the existence of an $\eta_0 \in (K^*)^\circ$ such that $B^\top \eta_0 \in -(K^*)^\circ$. Note that in the diffusive case the above reasoning based on [54] implies that we automatically have some exponential moments.*

(iv) *For compound Poisson jumps and the cone \mathbb{R}_+^m Condition (3.1) is identical to the drift condition of [85] which shows that it is close to being necessary as well. The latter can also be seen from the special case of Ornstein-Uhlenbeck type processes (see e.g. [63, 77]).*

3.2 Finite moments

In this section, polynomial functions are used to obtain the Foster-Lyapunov drift condition for the extended generator of X . This requires to know the precise action of the generator on polynomials. To this end we extend Proposition 2.10 to a larger function space.

For the following paragraph we borrow some introductory material and notation from [18, Section 4.3]¹³. We shall consider the tensor product $V \otimes V^*$, which we identify via the canonical isomorphism

$$(u \otimes v)x = \langle x, v \rangle u, \quad x \in V,$$

with the vector space of linear maps on V denoted by $\mathcal{L}(V)$. Moreover, for an element $A \in \mathcal{L}(V)$, we denote its trace by $\text{Tr}(A)$. Observe that $\text{Tr}(A(u \otimes u)) = \langle u, Au \rangle$. Indeed, by choosing an orthonormal basis $\{e_\beta\}$ of V , we have

$$\text{Tr}(A(u \otimes u)) = \sum_{\beta} \langle A^\top e_\beta, (u \otimes u)e_\beta \rangle = \sum_{\beta} \langle A^\top e_\beta, \langle u, e_\beta \rangle u \rangle = \sum_{\beta} \langle u, e_\beta \rangle \langle A^\top e_\beta, u \rangle = \langle u, Au \rangle.$$

Let $A : K \rightarrow \mathcal{L}(V)$ be the linear diffusion coefficient, identified by

$$\langle u, A(x)v \rangle = \langle x, Q(u, v) \rangle, \quad u, v \in V, \quad (3.9)$$

where $Q : V \times V \rightarrow V$ is the symmetric bilinear function in Definition 2.9 (iv). Next, we introduce the following linear operator acting on $C_b^2(K)$ -functions¹⁴

$$\begin{aligned} \mathcal{A}^\sharp f(x) &= \frac{1}{2} \text{Tr} \left(A(x) \left(\frac{\partial}{\partial x} \otimes \frac{\partial}{\partial x} \right) \right) f(x) + \langle b + B(x), \nabla f(x) \rangle - (c + \langle \gamma, x \rangle) f(x) \\ &\quad + \int_K (f(x + \xi) - f(x)) m(d\xi) + \int_K (f(x + \xi) - f(x) - \langle \chi(\xi), \nabla f(x) \rangle) \langle x, \mu(d\xi) \rangle. \end{aligned} \quad (3.10)$$

A straightforward check reveals that for any $u \in K^*$, the functions $f_u(x) = \exp(-\langle u, x \rangle)$ satisfy

$$\mathcal{A}^\sharp f_u(x) = f_u(x)(-F(u) - \langle x, R(u) \rangle) = \mathcal{A} f_u(x), \quad (3.11)$$

where the last identity holds by virtue of Proposition 2.10. Let now \mathcal{S}_+ be the restriction of rapidly decreasing C^∞ functions on V to K . Using (3.11) and similar arguments as in the proof of [16, Proposition 5.9]¹⁵,

Proposition 3.4. *\mathcal{S} is part of the domain of the infinitesimal generator \mathcal{A} and for any $f \in \mathcal{S}_+$, we have $\mathcal{A}f(x) = \mathcal{A}^\sharp f(x)$.*

For the remainder of this section, we identify $V = \mathbb{R}^n$ to be able to introduce the spaces \mathcal{P}_k of polynomials in the variables x_1, \dots, x_n , with real coefficients of order $\leq k$ ¹⁶.

Theorem 3.5. *Let $m \geq 2$. Suppose $c = 0$, $\gamma = 0$, and for some $x_0 \in K^\circ$,*

$$\int_{K \setminus \{0\}} \|\xi\|^m (m(d\xi) + \langle x_0, \mu(d\xi) \rangle) < \infty. \quad (3.12)$$

Then X is a conservative process with infinitesimal generator \mathcal{A} acting on \mathcal{S}_+ as

$$\begin{aligned} \mathcal{A}^\sharp f(x) &= \frac{1}{2} \text{Tr} \left(A(x) \left(\frac{\partial}{\partial x} \otimes \frac{\partial}{\partial x} \right) \right) f(x) \\ &\quad + \langle b + \tilde{B}(x), \nabla f(x) \rangle + \int_K (f(x + \xi) - f(x) - \langle \nabla f(x), \xi \rangle) (m(d\xi) + \langle x, \mu(d\xi) \rangle), \end{aligned} \quad (3.13)$$

where $\tilde{B}x = Bx + \int_K (\xi - \chi(\xi)) \langle x, \mu(d\xi) \rangle$. Furthermore, X is an m -polynomial process, that is, for any $p \in \mathcal{P}_m$, any $x \in K$ and for all $t \geq 0$ we have

$$\mathbb{E}^x[p(X_t)] < \infty$$

and the extended generator of X reduces to a (finite dimensional) vector space endomorphism $\mathcal{A} : \mathcal{P}_m \rightarrow \mathcal{P}_m$ given by (the pointwise limit)

$$\lim_{t \downarrow 0} \frac{\mathbb{E}^x[p(X_t)] - p(x)}{t} = \mathcal{A}^\# p(x),$$

where $\mathcal{A}^\#$ is defined in (3.13).

Proof. Since $c = 0$ and $\gamma = 0$ and by (3.12), X is conservative due to Proposition 2.11. Furthermore, Condition (3.12) implies $\int_{\|\xi\| \geq 1} \|\xi\| \langle x, \mu(d\xi) \rangle < \infty$. Therefore the modified drift $\tilde{B} = \tilde{B} - \int_K (\xi - \chi(\xi)) \langle x, \mu(d\xi) \rangle$ is well-defined (since finite), and the special form of the generator (3.13) holds.

The second part can be proved by using [17, Theorem 2.15] and the realization of X as a K -valued càdlàg semimartingale (using, e.g., the canonical realization). \square

We are prepared to state and prove the following ergodicity result:

Theorem 3.6. *Let X be an affine process on K with $c = 0$, $\gamma = 0$ and let $m \geq 2$. Assume that*

- (i) X is ν -irreducible with the support of ν having non-empty interior and aperiodic.
- (ii) There exists $\eta_0 \in (K^*)^\circ$ such that

$$B^\top \eta_0 + \int_{K \setminus \{0\}} \langle \eta_0, \xi - \chi(\xi) \rangle \mu(d\xi) \in -(K^*)^\circ. \quad (3.14)$$

- (iii) The jump measures have finite m -th moment, that is, there exists $x_0 \in K^\circ$ such that

$$\int_K \|\xi\|^m (m(d\xi) + \langle x_0, \mu(d\xi) \rangle) < \infty. \quad (3.15)$$

Then X is a conservative process. Furthermore, it is geometrically ergodic, positive Harris recurrent and the stationary distribution π has finite m -th moment, that is, $\int_K (\|\xi\|^m + 1) \pi(d\xi) < \infty$.

Remark 3.7. *In the below proof, we rely on the theory of [17] which requires $m \geq 2$ in (3.15). However, we conjecture that a similar result holds using the first moment.*

Proof of Theorem 3.6. First, by Theorem 3.5, X is conservative and m -polynomial, and its extended generator is of the form (3.13).

For $\eta \preceq \eta_0$ define the functions $H_\eta : K \rightarrow [1, \infty)$, $H_\eta(x) := \langle \eta, x \rangle^m + 1$, which by construction have range in $[1, \infty)$. Let $D_v f(x)$ be the Fréchet differential of a function f at a point x in the direction v . We clearly have

$$\begin{aligned} D_u H_{\eta_0}(x) &= m \langle \eta_0, x \rangle^{m-1} \langle \eta_0, u \rangle, \\ D_{(u,v)}^2 H_{\eta_0}(x) &= m(m-1) \langle \eta_0, x \rangle^{m-2} \langle \eta_0, u \rangle \langle \eta_0, v \rangle = O(\|x\|^{m-2}). \end{aligned}$$

Therefore, a straightforward computation yields

$$\mathcal{A}H_{\eta_0}(x) = m \langle \eta_0, x \rangle^{m-1} \left(\langle x, B^\top \eta_0 + \int_{K \setminus \{0\}} \langle \eta_0, \xi - \chi(\xi) \rangle \mu(d\xi) \right) + O(\|x\|^{m-1}).$$

Moreover, by assumption (ii) and the fact that $\eta_0 \in (K^*)^\circ$, there exists $\lambda > 0$ such that

$$\tilde{B}^\top \eta_0 = B^\top \eta_0 + \int_{K \setminus \{0\}} \langle \eta_0, \xi - \chi(\xi) \rangle \mu(d\xi) \preceq -\lambda \eta_0.$$

Hence,

$$\mathcal{A}H_{\eta_0}(x) \leq -\lambda m \langle x, \eta_0 \rangle^m + O(\|x\|^{m-1}).$$

Thus, there exists $d \in \mathbb{R}$ such that for sufficiently large r , we have for the compact ball $B = \{x \in K \mid \|x\| \leq r\}$,

$$\mathcal{A}H_{\eta_0}(x) \leq -cH_{\eta_0}(x) + d \cdot 1_K,$$

for some $c < \lambda m$. By Assumption (i) and Proposition 2.18, every compact set is a petite set, whence (\mathcal{A}, H) satisfies the Foster-Lyapunov drift condition. By Theorem 2.16 we may conclude that X is H -uniformly ergodic.

Harris recurrence and the finiteness of m moments is proved along the corresponding lines of the Proof of Theorem 3.1. \square

4 Aperiodicity and irreducibility

The main statements of the previous Section, Theorems 3.1 and 3.6 assume away aperiodicity and irreducibility of the underlying affine process, since the focus was on establishing the needed Foster-Lyapunov drift condition. Establishing (Lebesgue-)irreducibility is typically linked to showing absolute continuity of transition probabilities and usually even to establishing strict positivity of the density. This is a research topic on its own. In this section, we show first that for a large class of diffusion processes, adding finite activity jumps does not change aperiodicity and irreducibility. As for diffusions (see Remark 4.4), the existence of (strictly positive) transition densities is much better understood than for the general affine case with jumps, this immediately establishes Lebesgue-irreducibility and aperiodicity for a large set of affine processes (purely diffusive or with finite activity jumps) on irreducible symmetric cones. Finally, we discuss some finite activity pure jump cases where irreducibility and aperiodicity can be obtained.

For the entire section, we work with the càdlàg modification of the canonical realization of $X \equiv (X_t)_{t \geq 0}$, for each initial law \mathbb{P}^x . Moreover, we assume that $c = 0$, $\gamma = 0$, and the jump-behavior is of compound Poisson type, in the sense that the jump measures m, μ are finite and μ admits a first moment,

$$m(K) < \infty, \quad \mu(K) \in K^*, \tag{4.1}$$

$$\int_{\|\xi\| \geq 1} \|\xi\| \mu(d\xi) \in K^*. \tag{4.2}$$

This implies the following:

- for any $x \in K$, $\int_{\|\xi\| \geq 1} \|\xi\| \langle x, \mu(d\xi) \rangle < \infty$, and therefore X is conservative. Note that $\int_{\|\xi\| \geq 1} \|\xi\| \langle x, \mu(d\xi) \rangle < \infty$ is also implied by the assumptions of Theorems 3.1 and 3.6.
- The process X has finite activity jumps, and therefore $\sum_{s \leq t} \Delta X_s := \sum_{s \leq t} (X_s - X_{s-})$ is well-defined.

- We may take $\chi \equiv 0$ as truncation function.

By enlarging the state-space, we can disentangle the process from its discontinuous component in a tractable way:

Lemma 4.1. *For each $x \in K$, $y \in K$, the process $(X_t, Y_t)_{t \geq 0}$, where $Y_t := y + \sum_{s \leq t} \Delta X_s$, is an affine process on $K \times K$, with transition probability*

$$\mathbb{P}^{x,y}[X_t \in A, Y_t \in B] = \mathbb{P}^x[X_t \in A, y + \sum_{s \leq t} \Delta X_s \in B].$$

Its Laplace transform for any $u \in (K^*)^\circ$, $v \in K^*$ is given by

$$\mathbb{E}^{x,y}[e^{-\langle u, X_t \rangle - \langle v, Y_t \rangle}] = e^{-\phi(t,u,v) - \langle \psi(t,u,v), x \rangle - \langle y, v \rangle}, \quad (4.3)$$

where $\phi = \phi(t, u, v)$, $\psi = \psi(t, u, v)$ are the unique, global solutions of the generalized Riccati differential equations

$$\begin{aligned} \partial_t \phi &= \langle b, \psi \rangle - \int_{K \setminus \{0\}} \left(e^{-\langle \psi, \xi \rangle - \langle v, \xi \rangle} - 1 \right) m(d\xi), \\ \partial_t \psi &= -\frac{1}{2} Q(\psi, \psi) + B^\top \psi - \int_{K \setminus \{0\}} \left(e^{-\langle \psi, \xi \rangle - \langle v, \xi \rangle} - 1 \right) \mu(d\xi), \end{aligned}$$

with initial conditions

$$\phi(0, (u, v)) = 0, \quad \psi(0, (u, v)) = u \in (K^*)^\circ.$$

Proof. Due to (4.2), the map

$$R(u) := -\frac{1}{2} Q(u, u) + B^\top u - \int_{K \setminus \{0\}} \left(e^{-\langle u, \xi \rangle - \langle v, \xi \rangle} - 1 \right) \mu(d\xi)$$

is locally Lipschitz on $(K^*)^\circ$. Therefore, a unique local solution exists, and we denote by $t_+(u, v) > 0$ the explosion time of $\psi(t, u, v)$.

As the jumps of X are of finite total variation, we can write $X = X^c + \sum_{s \leq t} \Delta X_s$, with the continuous part X^c . Clearly, the components of the extended process (X, Y) jump simultaneously and have jumps of equal size ($\Delta X_s = \Delta Y_s$, almost surely), hence for a bounded, continuous function $f(t, x, y)$, the process

$$\begin{aligned} & \sum_{s \leq t} (f(t, X_{s-} + \Delta X_s, Y_{s-} + \Delta Y_s) - f(t, X_{s-}, Y_{s-})) \\ & - \int_0^t (f(t, X_{s-} + \xi, Y_{s-} + \xi) - f(t, X_{s-}, Y_{s-})) (m(d\xi) + \langle \mu(d\xi), X_{s-} \rangle) ds \end{aligned}$$

is a uniformly integrable martingale. Therefore, by Itô's formula, for any $t < t_+(u, v)$ the process

$$e^{-\phi(t-s) - \langle \psi(t-s), X_s \rangle - \langle v, Y_s \rangle}, \quad 0 \leq s \leq t$$

is a uniformly integrable martingale and thus the affine transform formula (4.3) holds for $t < t_+(u, v)$. Since X is an affine process, by assumption, the generalized Riccati differential equations

$$\begin{aligned} \partial_t \phi^0 &= \langle b, \psi^0 \rangle - \int_{K \setminus \{0\}} \left(e^{-\langle \psi^0, \xi \rangle} - 1 \right) m(d\xi), \\ \partial_t \psi^0 &= -\frac{1}{2} Q(\psi^0, \psi^0) + B^\top \psi^0 - \int_{K \setminus \{0\}} \left(e^{-\langle \psi^0, \xi \rangle} - 1 \right) \mu(d\xi), \end{aligned}$$

with initial conditions

$$\phi^0(0, u) = 0, \quad \psi^0(0, u) = u$$

have global solutions for any $u \in K^*$, and, due to the local Lipschitz property mentioned above, the solutions are unique for $u \in (K^*)^\circ$. Note, $\psi^0(t, u)$ remains in $(K^*)^\circ$ for all times $t \geq 0$, due to [18, Lemma 3.2]. Let \preceq be the partial order induced by the dual cone. Furthermore, since $v \in K^*$, we have

$$\partial_t \psi(t, u, v) \succeq -\frac{1}{2}Q(\psi, \psi) + B^\top \psi - \int_{K \setminus \{0\}} \left(e^{-\langle \psi, \xi \rangle} - 1 \right) \mu(d\xi).$$

Hence for $u \in (K^*)^\circ$ and $v \in K^*$, we have by Volkmann's comparison result [84] that $\psi(t, u, v) \succeq \psi^0(t, u)$ for any $t < t_+(u, v)$. Due to the afore mentioned property $\psi^0(t, u) \in (K^*)^\circ$, $\psi(t, u, v)$ does not touch the boundary of the dual cone K^* in finite time. We conclude that if $t_+(u, v) < \infty$, then explosion in norm occurred as $t \uparrow t_+(u, v)$. But this is impossible, as it would imply, due to (4.3), that the Laplace transform $\mathbb{E}^{x, y}[e^{-\langle u, X_t \rangle - \langle v, Y_t \rangle}] = 0$ for $t = t_+(u, v)$. Hence we conclude that $t_+(u, v) = \infty$ and the affine transform formula (4.3) holds for all $t \geq 0$, $u \in (K^*)^\circ$ and $v \in K^*$. \square

The next statement shows the intuitive fact that before a process jumps, it cannot be distinguished from a killed diffusion process \hat{X} with the same characteristics as X , except having vanishing jump measures m and μ :

Proposition 4.2. *Suppose \hat{X} is an affine processes with parameters $\hat{Q} = Q, \hat{B} = B, \hat{b} = b$, but $\hat{m}(d\xi)$ and $\hat{\mu}(d\xi)$ vanish. Then for any $t > 0$, $x \in K$ and any $A \in \mathcal{B}(K)$ it holds that*

$$\mathbb{P}^x[X_t \in A, \tau > t] = \hat{\mathbb{E}}^x \left(e^{-\int_0^t (l + \langle \Lambda, \hat{X}_s \rangle) ds} 1_{\{\hat{X}_t \in A\}} \right), \quad (4.4)$$

where $l := m(K) \geq 0$ and $\Lambda = \mu(K) \in K^*$ and $\tau = \inf\{s \geq 0 \mid \Delta X_s \neq 0\}$.

Proof. In order to show equation (4.4) we have to show that the images of the measures $1_{\{\tau > t\}} d\mathbb{P}^x$ and $e^{-\int_0^t (l + \langle \Lambda, \hat{X}_s \rangle) ds} d\hat{\mathbb{P}}^x$ under X and \hat{X} , respectively, are equal. But for that it is enough to prove the identity of the Laplace transformations, i.e. that for any $t \geq 0$, $x \in K$ and $u \in -K^*$,

$$\mathbb{E}^x[e^{\langle u, X_t \rangle} 1_{\{\tau > t\}}] = \hat{\mathbb{E}}^x[e^{-\int_0^t (l + \langle \Lambda, \hat{X}_s \rangle) ds + \langle u, \hat{X}_t \rangle}].$$

We consider the K^2 -valued process (X, Y) as in Lemma 4.1. Clearly,

$$\mathbb{P}^x[X_t \in A, \tau > t] = \mathbb{P}^x[X_t \in A, \sum_{s \leq t} X_s = 0] = \mathbb{P}^x[X_t \in A, Y_t = y].$$

For a random variable $Z \geq 0$ it holds that $\mathbb{P}[Z = 0] = \lim_{w \rightarrow -\infty} \mathbb{E}[e^{wZ}]$. Applying this trick to $Z := \langle w_0, Y_t \rangle$ for $y = 0$, for some $w_0 \in (K^*)^\circ$ we get $\mathbb{P}^x[\tau > t] = \mathbb{P}^x[Y_t = 0] = \lim_{w \rightarrow -\infty} \mathbb{E}[e^{w \langle w_0, Y_t \rangle}]$.

Setting $y = 0$ in Lemma 4.1 we obtain by continuity for $w \rightarrow -\infty$

$$\begin{aligned} \partial_t \phi^\infty &= \langle b, \psi^\infty \rangle + l, \\ \partial_t \psi^\infty &= -\frac{1}{2}Q(\psi^\infty, \psi^\infty) + B^\top \psi^\infty + \Lambda, \end{aligned}$$

subject to

$$\phi^\infty(0) = 0, \quad \psi^\infty(0) = u.$$

Note that

$$e^{\phi^\infty(t,u)+\langle x,\psi^\infty(t,u)\rangle} = \hat{\mathbb{E}}^x(e^{\langle u,X_t^*\rangle} ds) = \hat{\mathbb{E}}^x(e^{\langle u,\hat{X}_t\rangle} e^{-\int_0^t(l+\langle \Lambda,\hat{X}_s\rangle)ds}),$$

where X^* is an affine process with parameters $Q, b, B, c = l, \gamma = \Lambda$ but with vanishing jump measures $\hat{m}(d\xi), \hat{\mu}(d\xi)$.

Summarizing, we have

$$\begin{aligned} \mathbb{E}^x[e^{\langle u,X_t\rangle} 1_{\{\tau>t\}}] &= \lim_{w\rightarrow-\infty} e^{\phi(t,(u,wv_0))+\langle x,\psi(t,(u,wv_0))\rangle} \\ &= e^{\psi^\infty(t,u)+\langle x,\psi^\infty(t,u)\rangle} = \hat{\mathbb{E}}^x[e^{\langle u,\hat{X}_t\rangle} e^{-\int_0^t(l+\langle \Lambda,\hat{X}_s\rangle)ds}]. \end{aligned}$$

□

The following essentially states that adding finite activity jumps to a diffusion with strictly positive transition densities maintains irreducibility and aperiodicity.

Proposition 4.3. *Suppose \hat{X} is an affine processes with parameters $\hat{Q} = Q, \hat{B} = B, \hat{b} = b$, but $\hat{m}(d\xi)$ and $\hat{\mu}(d\xi)$ vanish.*

Suppose for any $t > 0, x \in K$, the transition probabilities $\hat{p}_t(x, d\xi)$ of \hat{X} have a Lebesgue density which is strictly positive on K° . Then both X and \hat{X} are aperiodic and irreducible (with respect to the Lebesgue measure λ on K).

Proof. Let $\tau := \inf\{t > 0 \mid \Delta X_t \neq 0\}$ be again the first jump time of X . By the law of total probability, for any Borel set A , and any $t > 0$ and for any $x \in K$ we have

$$\mathbb{P}^x[X_t \in A] \geq \mathbb{P}^x[X_t \in A, \tau > t] = \mathbb{P}^x[X_t \in A \mid \tau > t] \cdot \mathbb{P}^x[\tau > t].$$

We show first that for any $A \in \mathcal{B}(K)$ with $\lambda(A) > 0$ and any $t > 0$ it holds that

$$\mathbb{P}^x[X_t \in A, \tau > t] > 0, \quad \forall x \in K. \quad (4.5)$$

By assumption, for any $A \in \mathcal{B}(K)$ with $\lambda(A) > 0$ it holds that $\hat{\mathbb{P}}^x[\hat{X}_t \in A] > 0 \forall x \in K$ (note that the boundary of K is a null set, see e.g. [58]). Fix $t > 0, x \in K$ and $A \in \mathcal{B}(K)$ with $\lambda(A) > 0$. Let $\varepsilon := \frac{1}{2}\mathbb{P}^x[\hat{X}_t \in A]$. Due to the continuity of the path of X there exists $\eta_\varepsilon > 0$ such that

$$\mathbb{P}^x\left[\sup_{0 \leq s \leq t} \langle \hat{X}_s, \Lambda \rangle \leq \eta_\varepsilon\right] \geq 1 - \varepsilon.$$

Thus for $B_\varepsilon := \{\sup_{0 \leq s \leq t} \langle X_s, \Lambda \rangle \leq \eta_\varepsilon\}$, we obtain

$$\mathbb{P}^x[(\hat{X}_t \in A) \cap B_\varepsilon] \geq 2\varepsilon + 1 - \varepsilon - \mathbb{P}^x[\hat{X}_t \in A \cup B_\varepsilon] \geq \varepsilon.$$

Using $\mathbb{P}^x[\hat{X}_t \in A] \geq \mathbb{P}^x[(\hat{X}_t \in A) \cap B_\varepsilon]$ we can conclude

$$\begin{aligned} \mathbb{P}^x[X_t \in A, \tau > t] &= \mathbb{E}^x\left[e^{-\int_0^t(l+\langle \Lambda,\hat{X}_s\rangle)ds} 1_{\{\hat{X}_t \in A\}}\right] \\ &\geq e^{-(l+\eta_\varepsilon)t} \mathbb{E}^x[1_{\{(\hat{X}_t \in A) \cap B_\varepsilon\}}] \\ &\geq e^{-(l+\eta_\varepsilon)t} \varepsilon > 0, \end{aligned}$$

and estimate (4.5) is shown.

Thus, for any $A \in \mathcal{B}(K)$ for which $\lambda(A) > 0$, it follows that $\hat{p}_t(x, A) > 0$ and $\hat{p}_t(x, A) > 0$ and thus the processes are λ -irreducible.

Finally, we show aperiodicity, i.e., that there exists a small set \mathcal{C} and a time $T > 0$ such that for any $t > T$,

$$\mathbb{P}^x[X_t \in \mathcal{C}] > 0, \quad x \in \mathcal{C}. \quad (4.6)$$

By strict positivity of the transition densities (see first part), the compact set $\mathcal{C} := \{\xi \in K \mid \|\xi\| \leq 1\}$ satisfies (4.6); it therefore suffices to show that \mathcal{C} is small for X : By Proposition 2.18, every compact set is petite for any T -skeleton chain. Furthermore, by λ -irreducibility of X , every T -skeleton chain is irreducible and thus also aperiodic, see [82, Proposition 1.2]. Therefore every petite set is small for every T -skeleton chain, see [73, Proposition 5.5.7]. We have thus established that \mathcal{C} is small for any T -skeleton chain ($T > 0$). In other words, there exists $n \in \mathbb{N}$ and a non-trivial measure $\nu_n(d\xi)$ such that for all $x \in \mathcal{C}$ we have $p_{nT}(x, \cdot) = \int p_{kT}(x, \cdot) a(dt) \geq \nu_n(\cdot)$, where $a(dt) = \delta_{nT}(dt)$. Hence, by definition, it is small for the continuous-time Markov process X as well.¹⁷¹⁸

Since \hat{X} is just a process with the same characteristics as X , but with zero jumps, aperiodicity also holds for \hat{X} , and we are done. \square

Remark 4.4. • *For a large class of diffusions on irreducible symmetric cones (the so-called “Bru processes with linear drift”), the existence of a strictly positive Lebesgue density for each $t > 0$ and $x \in K$ is characterized in terms of a standard controllability condition [18, Section 5.1.4]. In fact, in this case for each $t > 0$, the process X_t is non-centrally Wishart distributed, and in this case the transition density is explicitly given, is supported on K° and is strictly positive thereon. We provide in the next section as a special case the Wishart processes on the cones of positive semi-definite $d \times d$ matrices (in which the controllability condition is Theorem 5.2 (ii)).*

- *In the pure jump case, [79] establishes under a comparable controllability condition under infinite activity jumps the absolute continuity of the transition probabilities of an Ornstein-Uhlenbeck (OU) process. However, irreducibility holds under weaker conditions for the OU process, see e.g., [63] and [75] for examples of non-degenerate/absolutely continuous stationary distributions in the context of cones).*

We now turn to the pure-jump case. So $Q = 0$ in addition to the assumptions made so far. The following sufficient conditions for irreducibility and aperiodicity are inspired by similar results of [81] for a non-linear SDE in the cone of positive semi-definite matrices.

Proposition 4.5. *Let X be a pure jump affine process as described above. If*

- (i) *m or $\langle \eta, \mu(\cdot) \rangle$ for one $\eta \in K^\circ$ have an absolutely continuous component with a strictly positive density on K° ,*
- (ii) *$e^{Bt}(K^\circ) = K^\circ$ for all $t \in \mathbb{R}_+$ and $\tau(B) < 0$,*
- (iii) *$m(K^\circ) > 0$ or $b \in K^\circ$ provided m has no absolutely continuous component with a strictly positive density on K° ,*

then X is irreducible with respect to the Lebesgue measure on $B^{-1}b + K$ and aperiodic.

Proof. Since aperiodicity and irreducibility are properties of the law of the process, we can assume that X is constructed as follows on an appropriate probability space: Let $(\sigma_j)_{j \in \mathbb{N}}$ be an iid sequence of $Exp(1)$ distributed random variables. Set

$$\tau_1 = \rho_1 = \inf \left\{ z > 0 : \int_0^z \left(m(K) + \left\langle e^{Bt}x + \int_0^t e^{B(t-s)} b dt, \mu(K) \right\rangle \right) dt \geq \sigma_1 \right\}$$

$$X_t = e^{Bt}x + \int_0^t e^{B(t-s)} b ds \text{ for } t \in [0, \tau_1).$$

Condition (iii) implies that $\tau_1 < \infty$ almost surely. Next draw a K -valued random variable ΔX_{τ_1} with distribution $(m(\cdot) + \langle X_{\tau_1-}, \mu(\cdot) \rangle) / (m(K) + \langle X_{\tau_1-}, \mu(K) \rangle)$ and otherwise independent of $(\sigma_j)_{j \in \mathbb{N}}$. Set

$$X_{\tau_1} = e^{B\tau_1}x + \int_0^{\tau_1} e^{B(\tau_1-s)} b ds + \Delta X_{\tau_1}.$$

For $j = 2, 3, \dots$ apply the following procedure iteratively. Set

$$\rho_j = \inf \left\{ z > 0 : \int_{\tau_{j-1}}^{\tau_{j-1}+z} \left(m(K) + \left\langle e^{Bt}x + \int_0^t e^{B(t-s)} b dt + \sum_{k=1}^{j-1} e^{B(t-\tau_k)} \Delta X_{\tau_k}, \mu(K) \right\rangle \right) dt \geq \sigma_j \right\},$$

$$\tau_j = \tau_{j-1} + \rho_j,$$

$$X_t = e^{Bt}x + \int_0^t e^{B(t-s)} b ds + \sum_{k=1}^{j-1} e^{B(t-\tau_k)} \Delta X_{\tau_k} \text{ for } t \in (\tau_{j-1}, \tau_j).$$

Condition (iii) implies that $\rho_j < \infty$ a.s. and so $\tau_j < \infty$ a.s. Then draw a K -valued random variable ΔX_{τ_j} with distribution $(m(\cdot) + \langle X_{\tau_j-}, \mu(\cdot) \rangle) / (m(K) + \langle X_{\tau_j-}, \mu(K) \rangle)$ and otherwise independent of $\{(\sigma_j)_{j \in \mathbb{N}}, (\Delta X_{\tau_k})_{k=1, \dots, j-1}\}$. Set

$$X_{\tau_j} = e^{B\tau_j}x + \int_0^{\tau_j} e^{B(\tau_j-s)} b ds + \sum_{k=1}^{j-1} e^{B(\tau_j-\tau_k)} \Delta X_{\tau_k} + \Delta X_{\tau_j}.$$

Obviously this gives a process X with the correct semimartingale characteristics (see [51] for a comprehensive treatment).

Furthermore, by an adaption of the proof of [16, Proposition 5.3]¹⁹, the associated generalized Riccati differential equations (2.4a)–(2.4b) have a unique global solution for any initial data $u \in (K^*)^\circ$. We introduce the filtration $\mathcal{F} := (\mathcal{F}_t)_{t \geq 0}$, where $\mathcal{F}_t := \sigma(X_s \mid s \leq t)$. By Itô's formula and the generalized Riccati differential equations, we find that for any $t > 0$, the process $e^{-\Phi(t-s, u) - \langle \Psi(t-s, u), X_s \rangle}$ is a uniformly integrable martingale on $[0, t]$, hence we conclude the affine transform formula holds, that is,

$$\mathbb{E}[e^{-\langle u, X_t \rangle} \mid \mathcal{F}_s] = e^{-\Phi(t-s, u) - \langle \Psi(t-s, u), X_s \rangle}, \quad 0 \leq s < t.$$

Therefore, by a similar argument as in [34, Section 8], the Laplace transforms of the finite-dimensional distributions of X can be computed explicitly in terms of the functional characteristics Φ, Ψ .²⁰ Since the finite-dimensional distributions uniquely characterize the process in law, X is an affine process with parameters $b, B, m(d\xi)$ and $\mu(d\xi)$, as desired.

First, we establish the claimed irreducibility: Observe that for any $x \in K$,

$$\lim_{t \rightarrow \infty} \left(e^{Bt}x + \int_0^t e^{B(t-s)} b ds \right) = B^{-1}b \in K. \quad (4.7)$$

So from wherever we start we get eventually “down” towards $B^{-1}b$, as long as no jumps occur. The rest of the proof is about rigorously establishing that from there we can reach any relevant set with positive probability at any large enough time. Fix some $x \in K$ and a measurable set $A \subseteq B^{-1}b + K$ with non-vanishing Lebesgue measure. It suffices to show that there exists a $T > 0$ such that

$$\mathbb{P}^x[X_t \in A] > 0 \forall t > T.$$

The above construction of X and Condition (iii) imply that $\mathbb{P}^x[X_t \in K^\circ] > 0$ for all $x \in K$ and $t > 0$. Therefore, we can assume w.l.o.g. that $x \in K^\circ$ and thus $\mathbb{P}^x[X_t \in K^\circ \forall t \in \mathbb{R}^+] = 1$. Pick an $\eta \in K^\circ$ and an $\epsilon > 0$ such that $\lambda(\tilde{A}) > 0$ for $\tilde{A} := A \cap (\epsilon\eta + B^{-1}b + K)$. Now we can choose a $T > 0$ such that $e^{Bt}x + \int_0^t e^{B(t-s)} b ds \preceq \epsilon\eta + B^{-1}b$ for all $t \geq T$. Pick an arbitrary $t > T$.

$$\begin{aligned} \mathbb{P}^x[X_t \in A] &\geq \mathbb{P}^x[\{X_t \in \tilde{A}\} \cap \{\tau_1 \in (T, t]\} \cap \{\tau_2 > t\}] \\ &= \int_T^t \mathbb{P}^x \left[\Delta X_{\tau_1} \in e^{-B(t-s)} \left(\tilde{A} - e^{Bt}x - \int_0^t e^{B(t-u)} b du \right) \middle| \tau_2 > t, \tau_1 = s \right] \mathbb{P}^x_{\tau_1} [ds | \tau_2 > t] \\ &\quad \cdot \mathbb{P}^x[\tau_2 > t]. \end{aligned}$$

Here $\mathbb{P}^x_{\tau_1}[\cdot | \tau_2 > t]$ is the conditional law of τ_1 under \mathbb{P}^x given $\tau_2 > t$ and $\mathbb{P}^x[\cdot | \tau_2 > t, \tau_1 = s]$ a version of the regular conditional probability given $\tau_1 = s$. From the above explicit construction we now see that $\mathbb{P}^x[\tau_2 > t] > 0$ and that $\mathbb{P}^x_{\tau_1}(\cdot | \tau_2 > t)$ is equivalent to the Lebesgue measure on $(T, t]$. Observe that the measures $(\langle \eta, \mu(\cdot) \rangle)_{\eta \in K^\circ}$ are all equivalent. Hence, $\mathcal{M}^+(\eta) := \{M \in \mathcal{B}(K) : m(M) + \langle \eta, \mu(M) \rangle > 0\}$ (with $\mathcal{B}(K)$ denoting the Borel- σ -algebra over K) does not depend on $\eta \in K^\circ$. Due to Condition (i) $\mathcal{M}^+(\eta) \supseteq \{M \in \mathcal{B}(K) : \lambda(M) > 0\}$. From the explicit construction we see that $\mathbb{P}^x[\Delta X_{\tau_1} \in M | \tau_2 > t, \tau_1 = s] > 0$ holds for all $M \in \mathcal{M}^+(\eta)$ and $s \in (T, t]$. Due to our choices the set $\tilde{A} - e^{Bt}x - \int_0^t e^{B(t-u)} b du$ is in K and has non-zero Lebesgue measure. As $e^{B(t-s)}$ is a diffeomorphism the standard transformation formula (e.g. [56, Theorem 1.101]), implies that $e^{-B(t-s)} \left(\tilde{A} - e^{Bt}x - \int_0^t e^{B(t-u)} b du \right)$ has non-zero Lebesgue measure for all $s \in (T, t]$. Thus it belongs to $M \in \mathcal{M}^+(\eta)$. Hence, $\mathbb{P}^x \left[\Delta X_{\tau_1} \in e^{-B(t-s)} \left(\tilde{A} - e^{Bt}x - \int_0^t e^{B(t-u)} b du \right) \middle| \tau_2 > t, \tau_1 = s \right] > 0$ holds for all $s \in (T, t]$ and therefore $\mathbb{P}^x[X_t \in A] > 0$ for all $t > T$ concluding the proof of the irreducibility.

It remains to show the aperiodicity. Consider the set $\mathcal{C} = B_1(B^{-1}b) \cap K$. Arguing as in the proof of Proposition 4.3 we see that \mathcal{C} is small. Choose $T > 0$ such that $\int_0^t e^{B(t-s)} b ds \in B_{1/2}(B^{-1}b)$ and $\|e^{Bt}x\| < 1/2$ for all $x \in \mathcal{C}$ and $t > T$. Hence, $e^{Bt}x + \int_0^t e^{B(t-s)} b ds \in \mathcal{C}$ and therefore

$$\mathbb{P}^x[X_t \in \mathcal{C}] \geq \mathbb{P}^x[\tau_1 > t] > 0 \forall t > T, x \in \mathcal{C}.$$

Therefore the process is aperiodic. □

Remark 4.6. • *Similarly to [81, Corollary 4.7] one can relax the conditions to demanding that m or $\langle \eta, \mu(\cdot) \rangle$ for one $\eta \in K^\circ$ have an absolutely continuous component with a strictly positive density only in a neighbourhood of zero in K° . Then X is irreducible with respect to the Lebesgue measure restricted to a neighbourhood of $B^{-1}b$ and aperiodic.*

- If K is the cone of $d \times d$ symmetric positive definite matrices, one can also build the jumps by taking the outer products of jumps in \mathbb{R}^d and the above results remain true provided that they have a strictly positive density on \mathbb{R}^d (in a neighbourhood of zero). This can be shown by taking not only one jump after time T as in the above proof but at least d jumps and arguing similar to [81, Theorem 4.6].

5 Applications and extensions.

5.1 Wishart processes

Let $V = S_d$ be the symmetric $d \times d$ matrices, endowed with the inner product $\langle x, y \rangle = \text{Tr}(xy)$, where xy denotes the matrix product, and Tr is the trace of a $d \times d$ matrix. On the positive semi-definite symmetric $d \times d$ matrices $K = S_d^+ \subset V$, the Bru processes with linear drift are just the Wishart processes, which are for each initial value $x \in S_d^+$ the global unique weak solution of

$$dX_t = \sqrt{X_t} dW_t Q + Q^\top dW_t^\top \sqrt{X_t} + (X_t \beta + \beta^\top X_t + \delta Q^\top Q) dt, \quad (5.1)$$

$$X_0 = x \in S_d^+, \quad (5.2)$$

where Q, β are real $d \times d$ matrix (not necessarily symmetric), W is a standard $d \times d$ Brownian motion matrix, and $\delta > d - 1$. In this case, the parameters in the sense of Theorem 2.9 are

- $b = \delta \alpha$, where $\alpha = Q^\top Q \in S_d^+$, and $\delta > d - 1$,
- $Q : V \times V \rightarrow V$, $Q(u, u) = 2u\alpha u$, and
- $B^\top : V \rightarrow V$ is given by $B^\top(u) = \beta u + u\beta^\top$,

while $c, \gamma, m(d\xi)$ and $\mu(d\xi)$ vanish.

Under the above conditions [16] showed that the SDE (5.1) has a unique global weak solution remaining in S_d^+ . Strengthening the above assumptions to $\delta \geq d + 1$ [69] showed that a unique strong solution exists which always stays in the interior of the cone provided the initial value is strictly positive definite. Intuitively, the inward pointing drift must be sufficiently large to ensure that the Brownian term cannot move the process through the boundary.

Note that for any $x, u \in S_d^+$ by the definition of the inner product, and the cyclic property of the trace,

$$\langle B^\top(u), x \rangle = \langle \beta u + u\beta^\top, x \rangle = \langle u, x\beta \rangle + \langle u, \beta^\top x \rangle = \langle u, \beta^\top x + x\beta \rangle,$$

whence $B(u) = u\beta + \beta^\top u$.

An application of Theorem 2.7 to the self-dual cone $K = S_d^+$ is the following:

Corollary 5.1. *The following are equivalent:*

- (i) $\tau(\beta) < 0$.
- (ii) $\tau(B) < 0$.
- (iii) *There exists $v \in S_d^{++}$ such that $\beta v + v\beta^\top \in -S_d^{++}$.*

Proof. By construction, the linear map B is a qmi map on S_d^+ with respect to S_d^+ (in fact, even $-B$ is quasi monotone increasing). Therefore, Theorem 2.7 gives the equivalent of (ii) and (iii).

It remains to prove that (i) is equivalent to any of the two equivalent statements (ii) and (iii).

By [50, Theorem 4.45] we have $\sigma(B) \subseteq \sigma(\beta) + \sigma(\beta)$ (as our operator B is acting only on S_d), which settles the proof of (i) \Rightarrow (ii).

Let now $\lambda, \mu \in \mathbb{C}$ be two eigenvalues of β^T with corresponding eigenvectors $x, y \in \mathbb{C}^d$. Then $\beta^T(xy^* + yx^*) + (xy^* + yx^*)^*\beta = (\lambda + \mu)(xy^* + yx^*)$. As $xy^* + yx^*$ is Hermitian this shows that B^T has the eigenvalue $\lambda + \mu$ as an operator over the Hermitian matrices and thus over the real symmetric matrices. Hence, $\sigma(B) \supseteq \sigma(\beta) + \sigma(\beta)$. This settles the proof of (ii) \Rightarrow (i). \square

A consequence of Corollary 3.2, Proposition 4.3 and Corollary 5.1 is the geometric ergodicity of Wishart processes:

Theorem 5.2. *Let X be a Wishart process with parameters (δ, β, Q) . Assume that*

- (i) $\tau(\beta) < 0$ (or, any equivalent condition in Corollary 5.1).
- (ii) *The $d \times d^2$ matrix $Q^\top |\beta Q^\top| \dots |\beta^{d-1} Q^\top$ has maximal rank (equals d).*

Then X is geometrically ergodic, positive Harris recurrent and the stationary distribution has some finite exponential moments.

Remark 5.3. *As the distribution of Wishart processes is exactly known and the stationary distribution has to be a Wishart distribution, it is clear that the stationary distribution has some finite exponential moments. The main implication of our approach is that we get geometric ergodicity and thus exponentially fast convergence in the total variation norm (and actually in an exponential norm).*

Proof of Theorem 5.2. It suffices to show that all assumptions of Corollary 3.2 are satisfied:

By [66, Theorem 7.1], $X_t | X_0 = x$ admits a density, for each $t > 0$. By inspection of the series expansion of the density [66, eq (5.3)], we see that if for some multi-index κ , the zonal polynomial C_κ is strictly positive on S_d^{++} , then the density is strictly positive thereon. In fact, by [67, Lemma 3.1] this is the case for $\kappa = (1, \dots, 1)$, which is fairly obvious from the fact that in this case $C_\kappa = \det$, because the determinant is invariant under the action of the orthogonal group. By Proposition 4.3, X is aperiodic and irreducible and thus satisfies Condition (i) of Corollary 3.2.

By Corollary 5.1 also Condition (i) of Corollary 3.2 is satisfied, and the latter corollary concludes the proof. \square

Of course, we can extend the results of this section by adding jumps. Consider the process

$$dX_t = \sqrt{X_t} dW_t Q + Q^\top dW_t^\top \sqrt{X_t} + (X_t \beta + \beta^\top X_t + \delta Q^\top Q) dt + dJ_t, \quad (5.3)$$

$$X_0 = x \in S_d^+, \quad (5.4)$$

where Q, β are real $d \times d$ matrix (not necessarily symmetric), W is a standard $d \times d$ Brownian motion matrix, and $\delta > d - 1$. Furthermore $J_t = \sum_{s \leq t} \Delta X_s$ has absolutely continuous compensator $\nu(X_t, d\xi) dt = (m(d\xi) + \langle X_t, \mu(d\xi) \rangle) dt$, where $m(d\xi)$ is a finite positive measure on S_d^+ satisfying $m(S_d^+) = l$, and $\mu(d\xi)$ is an S_d^+ -valued measure on S_d^+ satisfying $\mu(S_d^+) = \Lambda$. In this case, the parameters in the sense of Theorem 2.9 are

- $b = \delta\alpha$, where $\alpha = Q^\top Q \in S_d^+$, and $\delta > d - 1$,
- $Q : V \times V \rightarrow V$, $Q(u, u) = 2u\alpha u$, and
- $B^\top : V \rightarrow V$ is given by $B^\top(u) = \beta u + u\beta^\top$,
- The jump measures $m(d\xi)$ and $\mu(d\xi)$ as above.

We shall refer to this class of affine jump-diffusions as *Wishart processes with jumps*.

Theorem 5.4. *Let X be a Wishart process with jumps. Assume that*

- (i) *The eigenvalues of the linear map $L : S_d \rightarrow S_d$ defined by*

$$Lu = \beta u + u\beta^\top + \int_{S_d^+} \text{Tr}(\xi u) \mu(d\xi)$$

have strictly negative real parts.

- (ii) *there exists $\eta_0 \in S_d^{++}$ such that $\int_{\{\|x\| \geq 1\}} e^{\text{Tr}(\eta_0 \xi)} (m(d\xi) + \text{Tr}(\mu(d\xi), x)) < \infty$.*
- (iii) *The $d \times d^2$ matrix $Q^\top |\beta Q^\top| \dots |\beta^{d-1} Q^\top|$ has maximal rank (equals d).*

Then X is geometrically ergodic, positive Harris recurrent and the stationary distribution has some finite exponential moments.

Proof. Due to Proposition 4.2, prior to its first jump, the Wishart process with jumps X cannot be distinguished from a Wishart process \hat{X} , killed at the exponentially affine rate $e^{-\int_0^t l + \text{Tr}(\Lambda X_t)}$. Since $\hat{X}_t | \hat{X}_0 = x$ admits a strict positive density, for each $t > 0$ (cf. proof of Theorem 5.2), the process X is aperiodic and irreducible (Proposition 4.3). Hence the main assertion follows by an application Corollary 3.2. \square

Using Theorem 3.6 we similarly get:

Theorem 5.5. *Let X be a Wishart process with jumps and $m \geq 2$ a natural number. Assume that*

- (i) *The eigenvalues of the linear map $L : S_d \rightarrow S_d$ defined by*

$$Lu = \beta u + u\beta^\top + \int_{S_d^+} \text{Tr}(\xi u) \mu(d\xi)$$

have strictly negative real parts.

- (ii) *The jump measures have finite m -th moment, that is, there exists $x_0 \in S_d^{++}$ such that*

$$\int_{S_d^+} \|\xi\|^m (m(d\xi) + \text{Tr}(x_0, \mu(d\xi))) < \infty. \quad (5.5)$$

- (iii) *The $d \times d^2$ matrix $Q^\top |\beta Q^\top| \dots |\beta^{d-1} Q^\top|$ has maximal rank (equals d).*

Then X is geometrically ergodic, positive Harris recurrent and the stationary distribution π has finite m -th moment, that is, $\int_{S_d^+} (\|\xi\|^m + 1) \pi(d\xi) < \infty$.

Remark 5.6. For dimension $d = 1$, vanishing $\mu(d\xi)$, and $m(d\xi) = cde^{-d\xi}d\xi$ (exponentially distributed jumps with mean d , intensity c) we obtain the corresponding results for the “Basic affine jump diffusion” (BAJD), [52, Theorem 5.6 and Theorem 6.3], given by the SDE

$$dX_t = a(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t + d\hat{J}_t, \quad X_0 \geq 0,$$

Here \hat{J} the compensated jump process.

Note that in this case, $\sigma > 0$ implies strictly positive transition densities (as for $J = 0$ they are non-central chi-square distributed. Alternatively, (iii) of Theorem 5.4 also implies strictly positive densities). Furthermore, since J is compensated, the standard assumption $a > 0$ implies Condition (i) of Theorem 5.4.

In this special one-dimensional case [53] could also establish ergodicity under log moments and geometric ergodicity under the finiteness of $\kappa > 0$ moments and for J being an infinite activity subordinator, i.e. only having state-independent jumps.

5.2 Affine processes on \mathbb{R}_+^m

In the following, the inner product $\langle x, y \rangle = x^\top y$ is the standard Euclidean scalar product on $V = \mathbb{R}^m$.

Affine processes with state space $K = \mathbb{R}_+^m$ are fully characterized by the paper [28]. In this case, the linear drift map $B^\top : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a $m \times m$ matrix with non-negative off-diagonal elements, hence it is easy to compute the eigenvalues of B to find a good test vector satisfying condition (ii) of Theorem 3.1. A corollary is therefore the following ergodicity result:

Theorem 5.7. Let X be an affine process on \mathbb{R}_+^m with $c = 0$, $\gamma = 0$. Assume that

- (i) X is ν -irreducible with the support of ν having non-empty interior and aperiodic,
- (ii) The eigenvalues of the linear map $B^\top + \int_{\mathbb{R}_+^m \setminus \{0\}} \langle \xi - \chi(\xi) \rangle \mu(d\xi)$ have strictly negative real parts, and for sufficiently small $w > 0$,

$$\int_{\{\|x\| \geq 1\}} e^{\|w\xi\|} (m(d\xi) + \langle \mu(d\xi), x \rangle) < \infty.$$

Then X is conservative, and there exists $\eta_0 \in \mathbb{R}_+^m$ such that for all $\eta \preceq \eta_0$, any $x \in \mathbb{R}_+^m$ and all $t \geq 0$ we have $\mathbb{E}^x[e^{\langle \eta, X_t \rangle}] < \infty$.

Furthermore, X is geometrically ergodic, positive Harris recurrent. and for any $\eta \preceq \eta_0$ the stationary distribution π has finite η -exponential moment, i.e. $\int_{K \setminus \{0\}} e^{\langle \eta_0, \xi \rangle} \pi(d\xi) < \infty$.

We refrain from stating the obvious analogue with finite m -th moments. For conditions ensuring irreducibility and aperiodicity we refer to Remarks 4.4, 4.6, Proposition 4.5 and [85] who discuss the canonical state space at length under a strong non-degeneracy assumption on the diffusive part which is stronger than the controllability assumption.

5.3 Affine term-structure models.

We consider state spaces $K \times \mathbb{R}^n$ in this section, where K is an m -dimensional generating, proper closed convex cone. There is no established theory of affine processes on these spaces, and also our theory developed in Chapter 3 does not fall into this setup, as $K \times \mathbb{R}^n$ is not a proper cone.

Due to [34, Lemma 7.1] the instantaneous diffusion of processes on $\mathbb{R}_+^m \times \mathbb{R}^n$ can always be brought into block-diagonal form. We take this statement as motivation to make the following assumption throughout. Let $Z := (X, Y)$ be an affine diffusion process on $K \times \mathbb{R}^n$ such that for each $z := (x, y) \in K \times \mathbb{R}_+^n$, X is the unique weak solution to

$$dX_t = (b + B(X_t))dt + \Sigma(X_t)dW_t^X, \quad X_0 \in K, \quad (5.6)$$

$$dY_t = (c + C(X_t) + D(Y_t))dt + \sigma(X_t)dW_t^Y, \quad Y_0 \in \mathbb{R}^n, \quad (5.7)$$

where W^X and W^Y are independent standard Brownian motions of dimensions m, n , respectively. Furthermore $b \in K$, $c \in \mathbb{R}^n$, and B is a linear map, qmi with respect to K . Moreover, $C : \mathbb{R}^m \rightarrow \mathbb{R}^n$ and $D : K \rightarrow \mathbb{R}^n$ are linear maps. The affine structure implies that $\langle u, \Sigma(x)\Sigma^\top(x)u \rangle = \langle Q(u, u), x \rangle$ as in Theorem 2.9 (iv), while $\sigma\sigma^\top(x) = E + Fx$, where $E \in S_n^+$ and $F : K \rightarrow S_n^+$ is linear.

For the associated generalized Riccati differential equations, the origin in $K \times \mathbb{R}^n$ is an asymptotically stable equilibrium, if the eigenvalues of the drift matrix

$$\Xi := \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

(or, equivalently, that of Ξ^\top) have strictly negative parts. This can be seen by linearizing the associated generalized Riccati differential equations for $\psi(t, u, v) = (\psi^X, \psi^Y)$ which are given by

$$\partial_t \psi^X = B^\top \psi^X + C^\top \psi^Y + \frac{1}{2}Q(\psi^X, \psi^X) + (\psi^Y)^\top F \psi^Y,$$

whereas

$$\partial_t \psi^Y = D^\top \psi^Y,$$

subject to $\psi(t=0, u, v) = (u, v)$.

Theorem 5.8. *Suppose Z satisfies the following:*

- (i) Z is ν -irreducible with the support of ν having non-empty interior and aperiodic,
- (ii) $\tau(\Xi) < 0$ (or, equivalently, $\tau(B) < 0$ and $\tau(D) < 0$).

Then Z is geometrically ergodic and positive Harris recurrent. Furthermore, the stationary distribution π has a finite second moment, i.e. $\int_{K \times \mathbb{R}^n} \|\xi\|^2 \pi(d\xi) < \infty$.

Remark 5.9. (i) *The inclusion of finite activity and state-dependent jump behavior is possible (similarly to the main results of this paper).*

- (ii) *For “canonical state spaces” $\mathbb{R}_+^m \times \mathbb{R}^n$, sufficient conditions for irreducibility and aperiodicity (that is (i) of Theorem 5.8) can be found in [85].*
- (iii) *The classical term-structure literature (for $K = \mathbb{R}_+^m$, see [40, Section 2, (C1)–(C4)]) uses stronger conditions on drift and diffusion coefficients of Z . In particular, the linear drift coefficients D and B must have real eigenvalues due to econometric identification issues.*
- (iv) *The ergodicity result of [4, Theorem 4.1] is the special case, where $K = \mathbb{R}_+$, $C = 0$ and the constant diffusion matrix E vanishes. They however prove directly Assumption (i). The condition (ii) clearly is just implied by strictly negative linear drift coefficients.*

Proof. Since $\tau(D) < 0$, due to [50, Theorem 2.2.1] there exists $Q \in S_n^{++}$ such that $D^\top Q + QD \in -S_n^{++}$ (and conversely, this implies $\tau(D) < 0$.)

Since B is qmi, there exists $\eta_0 \in K \setminus \{0\}$, such that $B^\top \eta_0 = \tau(B)\eta_0$ (see Lemma 2.6). We may without loss of generality assume $\inf_{x \in K, \|x\|=1} \langle \eta_0, x \rangle = 1$.

Let ϑ_0 be a positive constant satisfying

$$\vartheta_0 < \min \left(|\tau(D^\top Q + QD)|^3, \frac{(2|\tau(B)|)^{3/2}}{\|C\|^3 \|D\|^3} \right). \quad (5.8)$$

We introduce the quadratic test function $f : K \times \mathbb{R}_+^n \rightarrow [1, \infty)$ by setting

$$f(x, y) := 1 + \langle \eta_0, x \rangle^2 + \vartheta_0 \langle y, Qy \rangle, \quad x \in K, \quad y \in \mathbb{R}^n$$

and obtain for any $(x, y) \in K \times \mathbb{R}_+^n$,

$$\begin{aligned} & \langle B(x), f_x \rangle + \langle C(x) + D(y), f_y \rangle \\ &= 2\langle x, B^\top \eta_0 \rangle \langle \eta_0, x \rangle + 2\vartheta_0 \langle C(x), Qy \rangle + \vartheta_0 \langle (D^\top Q + QD)y, y \rangle \\ &\leq -2|\tau(B)| \langle \eta_0, x \rangle^2 - \vartheta_0 |\tau(D^\top Q + QD)| \|y\|^2 + 2\vartheta_0 \|C\| \|Q\| \|x\| \|y\| \\ &\leq -2|\tau(B)| \langle \eta_0, x \rangle^2 - \vartheta_0 |\tau(D^\top Q + QD)| \|y\|^2 + 2\vartheta_0 \|C\| \|Q\| \langle \eta_0, x \rangle \|y\| \\ &\leq -2|\tau(B)| \langle \eta_0, x \rangle^2 - \vartheta_0 |\tau(D^\top Q + QD)| \|y\|^2 + 2\vartheta_0^{1/3} \|C\| \|Q\| \langle \eta_0, x \rangle \vartheta_0^{2/3} \|y\| \\ &\leq -(2|\tau(B)| - \vartheta_0^{2/3} \|C\|^2 \|Q\|^2) \langle \eta_0, x \rangle^2 - \vartheta_0 (|\tau(D^\top Q + QD)| - \vartheta_0^{1/3}) \|y\|^2 \\ &\leq -\varepsilon(f(x, y) - 1) \end{aligned}$$

for some $\varepsilon > 0$, where in the last but one estimate we have used the fact that $\inf_{\|x\|=1} \langle \eta_0, x \rangle = 1$ and in the last one we have used (5.8) and the fact that $\langle Qy, y \rangle$ is non-degenerate, therefore bounded from below by a positive multiple of $\|y\|^2$.

The process Z is an affine diffusion, and thus m -polynomial for any $m \geq 2$, see [17]. Hence f lies in the domain of the extended generator.

Since the infinitesimal generator \mathcal{A} of (X, Y) is a second order differential operator with affine drift and affine diffusion coefficients, its formal application yields

$$\mathcal{A}f(x, y) \leq -\varepsilon(f(x, y) - 1) + O(\|x\| + \|y\|)$$

and since $(f - 1)^{1/2}$ is norm-like, we can find a large enough ball $\mathcal{C} = \{(x, y) \mid \|x\|^2 + \|y\|^2 \leq R^2\}$ such that

$$\mathcal{A}f(x, y) \leq -\frac{\varepsilon}{2} f(x, y) + d1_{\mathcal{C}}(x, y), \quad x \in K, \quad y \in \mathbb{R}^n.$$

We conclude that (\mathcal{A}, f) satisfies the Lyapunov -Foster drift condition.

The rest of the proof follows the lines of the proof of Theorem 3.1. \square

Notes

¹A typical statistical approach selects a parametric (sub)model and uses an estimation approach like quasi-maximum-likelihood or generalized methods of moments (see [46, 47, 64]). The parametrization needs to be sufficiently smooth and identifiable. If one assumes that the observations are following the stationary model as well as that the true parameter is included in the parameter space and if one chooses the parametric set-up additionally such

that for every parameter one has sufficiently fast strong mixing, this implies that the quasi-likelihood or the empirical moments converge to the corresponding expected values when the number of observations/the time horizon goes to infinity and that central limit theorems hold provided sufficiently many moments are finite. Combining this with the identifiability and regularity of the parametrization then typically implies consistency and asymptotic normality of the estimators. For examples of such approaches see [11, 39, 48, 78]. Strong mixing allows further to obtain an asymptotic theory for estimators of the spectral density, see [76]. Other strands of the literature like the series of papers [35–38] consider the non-parametric estimation of a drift in a one-dimensional diffusion model. Their approach needs geometric ergodicity in a uniform manner over the parameter space. We conjecture that in order to use Theorem 2.1 of [36] one could strengthen our arguments to obtain uniform Lyapunov drift conditions over suitable “compact” parameter spaces. However, establishing the additionally needed “uniform smallness (minorization)” condition seems to be intricate. Thus uniformity of geometric ergodicity results over parameter spaces are beyond the scope of the present paper and left for future research.

²The generating property, $V = K - K$ is technically necessary to identify the generalized Riccati differential equations for the exponents ϕ, ψ .

³A truncation function is continuous, bounded in norm by 1 and equal to the identity in a neighborhood of the origin.

⁴A convex cone is called symmetric, if it is self-dual and if the linear automorphism group acts transitively on it.

⁵A process is non-explosive (or conservative), if the transition function does not charge the cemetery state Δ .

⁶Our argument is longer than similar proofs (e.g., [16, 18, 68]) which overlook the fact that Lipschitz continuity does not necessarily hold in the real domain near zero, but only for the purely imaginary argument. In fact, for $R(u)$ being defined in a real neighborhood of zero, one needs the existence of exponential moments.

⁷The phrase “seemingly stronger” stems from [25]. In fact, only for $f = 1$, the two formulations coincide – then the total variation norm of a signed measure equals its f -norm. However, due to [25, Chapter 3 and Theorem 2.1], for Markov chains, f -geometric ergodicity is equivalent with weaker types of geometric ergodicity.

⁸The f -norm (of a signed measure μ) is defined for every measurable function from the state space D to $[1, \infty)$ by $\|\mu\|_f := \sup_{|g| \leq f} |\int \mu(dy)g(y)|$.

⁹In the next paragraph we avoid non-uniqueness by scaling η_0 sufficiently.

¹⁰The identification $p(t, \eta) = -\phi(t, -\eta)$ and $q(t, \eta) = -\psi(t, -\eta)$ is necessary, since we use the Laplace transform for the defining affine property, whereas [54] uses the characteristic function

¹¹The definition is independent of the choice of $x_0 \in K^\circ$, because for any $y_0 \in K^\circ$ there exists $h > 0$ such that for any $u \in K^*$ we have $h\langle y_0, u \rangle \leq \langle x_0, u \rangle \leq h^{-1}\langle y_0, u \rangle$.

¹²This follows directly from the Feller property as stated in Theorem 2.9. For a direct proof, see [18, Lemma 3.2].

¹³While the quoted material is stated within the framework of symmetric cones (Section 4), we only borrow the part that holds for general cones.

¹⁴We mean elements of $f \in C_b^2(V)$, restricted to K .

¹⁵The argument hinges on a density argument (see [16, Proof of Theorem B.3]), namely that the linear hull of $f_u(x)$, $u \in (K^*)^\circ$, is dense in the space of rapidly decreasing functions \mathcal{S}_+ on $K = S_d^+$, the cone of $d \times d$ positive semi-definite matrices. An inspection of this density argument reveals that it applies to the general cones K of this paper.

¹⁶The orders of polynomials do not depend on the choice of basis.

¹⁷Besides, any small set is petite (as it is petite by definition, for a degenerate sampling distribution $a(dt)$).

¹⁸See the paragraph after Equation (10) in [25].

¹⁹Set $Q = 0$ and replace S_d^{++} by $(K^*)^\circ$ in [16, Proposition 5.3].

²⁰Due to a density argument, it suffices to consider Laplace variables on $(K^*)^\circ$.

References

- [1] A. Alfonsi. *Affine diffusions and related processes: simulation, theory and applications*, volume 6 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan, 2015.
- [2] A. Alfonsi, A. Kebaier, and C. Rey. Maximum likelihood estimation for Wishart processes. *Stochastic Processes and their Applications*, 126.(11):3243–3282, 2016.

- [3] J. Baldeaux and E. Platen. *Functionals of multidimensional diffusions with applications to finance*, volume 5 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan, 2013.
- [4] M. Barczy, L. Döring, Z. Li, and G. Pap. Stationarity and ergodicity for an affine two-factor model. *Advances in Applied Probability*, 46(3):878–898, 2014.
- [5] O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society Series B Statistical Methodology*, 63(2):167–241, 2001.
- [6] O. E. Barndorff-Nielsen and R. Stelzer. Positive-definite matrix processes of finite variation. *Probability and Mathematical Statistics*, 27(1):3–43, 2007.
- [7] F. E. Benth and L. Vos. Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets. *Advances in Applied Probability*, 45(2):572–594, 2013.
- [8] F. Biagini, A. Gnoatto, and M. Härtel. Long-term yield in an affine HJM framework on S_d^+ . *Applied Mathematics & Optimization*, 77(3):405–441, 2018.
- [9] F. Boussama, F. Fuchs, and R. Stelzer. Stationarity and geometric ergodicity of BEKK multivariate GARCH models. *Stochastic Processes and Their Applications*, 121(10):2331–2360, 2011.
- [10] N. Branger, M. Muck, F. T. Seifried, and S. Weisheit. Optimal portfolios when variances and covariances can jump. *Journal of Economic Dynamics & Control*, 85:59–89, 2017.
- [11] P. J. Brockwell and E. Schlemm. Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations. *Journal of Multivariate Analysis*, 115: 217–251, 2013.
- [12] M.-F. Bru. Wishart Processes. *Journal of Theoretical Probability*, 4:725–751, 1991.
- [13] C. Chiarella, J. Da Fonseca, and M. Grasselli. Pricing range notes within Wishart affine models. *Insurance: Mathematics & Economics*, 58:193–203, 2014.
- [14] J. C. Cox, J. E. Ingersoll, Jr., and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.
- [15] C. Cuchiero and J. Teichmann. Path properties and regularity of affine processes on general state spaces. In *Séminaire de Probabilités XLV*, volume 2078 of *Lecture Notes in Math.*, pages 201–244. Springer, Cham, 2013.
- [16] C. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann. Affine processes on positive semidefinite matrices. *The Annals of Applied Probability*, 21(2):397–463, 2011.
- [17] C. Cuchiero, M. Keller-Ressel, and J. Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4):711–740, 2012.

- [18] C. Cuchiero, M. Keller-Ressel, E. Mayerhofer, and J. Teichmann. Affine processes on symmetric cones. *Journal of Theoretical Probability*, 29(2):359–422, 2016.
- [19] J. da Fonseca and M. Grasselli. Riding on the smiles. *Quantitative Finance*, 11:1609–1632, 2011.
- [20] Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978, 2000.
- [21] J. Dieudonné. *Foundations of Modern Analysis, enlarged and corrected printing*. Academic Press, 1969.
- [22] C. Donati-Martin. Large deviations for Wishart processes. *Probability and Mathematical Statistics*, 28(2):325–343, 2008.
- [23] C. Donati-Martin, Y. Dumerç, H. Matsumoto, and M. Yor. Some properties of the Wishart processes and a matrix extension of the Hartman-Watson laws. *Kyoto University. Research Institute for Mathematical Sciences. Publications*, 40(4):1385–1412, 2004.
- [24] P. Doukhan. *Mixing*, volume 85 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1994. ISBN 0-387-94214-9.
- [25] D. Down, S. P. Meyn, and R. L. Tweedie. Exponential and uniform ergodicity of Markov processes. *The Annals of Probability*, 23(4):1671–1691, 1995.
- [26] D. Duffie and N. Garleanu. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57(1):41–59, 2001.
- [27] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.
- [28] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Probability*, 13(3):984–1053, 2003.
- [29] L. Elsner. Quasimonotonie und Ungleichungen in halbgeordneten Räumen. *Linear Algebra and its Applications*, 8(3), 1974.
- [30] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*, volume 282. John Wiley & Sons, 2009.
- [31] K. Fan. Topological proofs for certain theorems on matrices with non-negative elements. *Monatshefte für Mathematik*, 62(3):219–237, 1958.
- [32] J. Faraut. *Analysis on symmetric cones*. Clarendon Press, Oxford, 1994.
- [33] D. Filipović. Time-inhomogeneous affine processes. *Stochastic Processes and Their Applications*, 115(4):639–659, 2005.
- [34] D. Filipović and E. Mayerhofer. Affine diffusion processes: theory and applications. In *Advanced financial modelling*, volume 8 of *Radon Ser. Comput. Appl. Math.*, pages 125–164. Walter de Gruyter, Berlin, 2009.

- [35] L. Galtchouk and S. Pergamenschchikov. Uniform concentration inequality for ergodic diffusion processes. *Stochastic Processes and their Applications*, 117:830–839, 2007.
- [36] L. Galtchouk and S. Pergamenschchikov. Geometric ergodicity for classes of homogeneous Markov chains. *Stochastic Processes and their Applications*, 124:3362–3391, 2014.
- [37] L. I. Galtchouk and S. M. Pergamenschchikov. Adaptive sequential estimation for ergodic diffusion processes in quadratic metric. *Journal of Nonparametric Statistics*, 23:255–285, 2011.
- [38] L. I. Galtchouk and S. M. Pergamenschchikov. Efficient pointwise estimation based on discrete data in ergodic nonparametric diffusions. *Bernoulli*, 2:2569–2594, 2015.
- [39] V. Genon-Catalot, T. Jeantheau, and C. Larédo. Stochastic volatility as hidden markov models and statistical applications. *Bernoulli*, 6:1051–1079, 2000.
- [40] P. Glasserman and K.-K. Kim. Moment explosions and stationary distributions in affine diffusion models. *Mathematical Finance*, 20(1):1–33, 2010.
- [41] A. Gnoatto and M. Grasselli. An affine multicurrency model with stochastic volatility and stochastic interest rates. *SIAM Journal on Financial Mathematics*, 5(1):493–531, 2014.
- [42] C. Gourieroux and R. Sufana. Derivative pricing with Wishart multivariate stochastic volatility. *Journal of Business & Economic Statistics*, 28(3):438–451, 2010.
- [43] P. Graczyk and L. Vostrikova. The moments of Wishart processes via Itô calculus. *Theory of Probability and its Applications*, 51(4):609 – 525, 2007.
- [44] A. Granelli and A. E. D. Veraart. Modeling the variance risk premium of equity indices: the role of dependence and contagion. *SIAM Journal on Financial Mathematics*, 7(1):382–417, 2016.
- [45] M. Grasselli and G. Miglietta. A flexible spot multiple-curve model. *Quantitative Finance*, 16(10):1465–1477, 2016.
- [46] A. R. Hall. *Generalized method of moments*. Advanced Texts in Econometrics. Oxford University Press, Oxford, 2005.
- [47] L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.
- [48] S. Haug, C. Klüppelberg, A. Lindner, and M. Zapp. Method of moment estimation in the COGARCH(1,1) model. *The Econometrics Journal*, 10:320–341, 2007.
- [49] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.
- [50] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
- [51] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, 2nd edition, 2003.

- [52] P. Jin, B. Rüdiger, and C. Trabelsi. Positive Harris recurrence and exponential ergodicity of the basic affine jump-diffusion. *Stochastic Analysis and Applications*, 34(1):75–95, 2016.
- [53] P. Jin, J. Kremer, and B. Rüdiger. Moments and ergodicity of the jump-diffusion CIR process. *Stochastics*, 2019. to appear, DOI: 10.1080/17442508.2019.1576686.
- [54] M. Keller-Ressel and E. Mayerhofer. Exponential moments of affine processes. *The Annals of Applied Probability*, 25(2):714–752, 2015.
- [55] M. Keller-Ressel, W. Schachermayer, and J. Teichmann. Regularity of affine processes on general state spaces. *Electronic Journal of Probability*, 18:no. 43, 17, 2013.
- [56] A. Klenke. *Probability theory*. Universitext. Springer, London, second edition, 2014.
- [57] M. G. Krein and M. A. Rutman. Linear operators leaving invariant a cone in a Banach space. *Uspekhi Matematicheskikh Nauk*, 3(1):3–95, 1948.
- [58] R. Lang. A note on the measurability of convex sets. *Archiv der Mathematik (Basel)*, 47(1):90–92, 1986.
- [59] M. Leippold and F. Trojani. Asset Pricing with Matrix Jump Diffusions. Working paper, 2010. available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1572576.
- [60] A. M. Lyapunov. The general problem of the stability of motion. *International Journal of Control*, 55(3):531–772, 1992.
- [61] D. Mahoney. *Modeling and Valuation of Energy Structures: Analytics, Econometrics, and Numerics*. Applied Quantitative Finance. Palgrave Macmillan, New York, 2016.
- [62] M. B. Marcus and J. Rosen. *Markov processes, Gaussian processes, and local times*, volume 100 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [63] H. Masuda. On multidimensional Ornstein-Uhlenbeck processes driven by a general Lévy process. *Bernoulli*, 10(1):97–120, 2004.
- [64] L. Mátyás, editor. *Generalized method of moments estimation*. Cambridge University Press, Cambridge, 1999.
- [65] E. Mayerhofer. Affine processes on positive semidefinite $d \times d$ matrices have jumps of finite variation in dimension $d > 1$. *Stochastic Processes and Their Applications*, 122(10):3445–3459, 2012.
- [66] E. Mayerhofer. Wishart processes and wishart distributions: an affine processes point of view. In *Modern methods in multivariate statistics, Travaux en Cours*. Hermann, Paris, 2014.
- [67] E. Mayerhofer. On Wishart and non-central Wishart distributions on symmetric cones. *Transactions of the American Mathematical Society*, 371(10):7093–7109, 2019.
- [68] E. Mayerhofer, J. Muhle-Karbe, and A. G. Smirnov. A characterization of the martingale property of exponentially affine processes. *Stochastic Processes and their Applications*, 121(3):568–582, 2011.

- [69] E. Mayerhofer, O. Pfaffel, and R. Stelzer. On strong solutions for positive definite jump-diffusions. *Stochastic Processes and Their Applications*, 121:2072–2086, 2011.
- [70] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes I: Criteria for discrete-time chains. *Advances in Applied Probability*, 24(3):542–574, 1992.
- [71] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes II: Continuous-time processes and sampled chains. *Advances in Applied Probability*, 25(3):487–517, 1993.
- [72] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548, 1993.
- [73] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, 2009.
- [74] J. Muhle-Karbe, O. Pfaffel, and R. Stelzer. Option pricing in multivariate stochastic volatility models of OU type. *SIAM Journal on Financial Mathematics*, 3:66–94, 2011.
- [75] C. Pigorsch and R. Stelzer. On the definition, stationary distribution and second order structure of positive semi-definite Ornstein-Uhlenbeck type processes. *Bernoulli*, 15:754–773, 2009.
- [76] M. Rosenblatt. Asymptotic normality, strong mixing and spectral density estimates. *The Annals of Probability*, 12(4):1167–1180, 1984.
- [77] K. Sato and M. Yamazato. Operator-selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Processes and Their Applications*, 17:73–100, 1984.
- [78] E. Schlemm and R. Stelzer. Quasi maximum likelihood estimation for multivariate Lévy-driven CARMA processes. *Electronic Journal of Statistics*, 6:2185–2234, 2012.
- [79] T. Simon. On the absolute continuity of multidimensional Ornstein-Uhlenbeck processes. *Probability Theory and Related Fields*, 151(1-2):173–190, 2011.
- [80] P. Spreij and E. Veerman. Affine diffusions with non-canonical state space. *Stochastic Analysis and Applications*, 30(4):605–641, 2012.
- [81] R. Stelzer and J. Vestweber. Geometric ergodicity of the multivariate COGARCH(1,1) process. Technical report, 2019. arXiv eprint 1701.07859.
- [82] P. Tuominen and R. L. Tweedie. Exponential decay and ergodicity of general Markov processes and their discrete skeletons. *Advances in Applied Probability*, 11(4):784–803, 1979.
- [83] F. Verhulst. *Nonlinear Differential Equations and Dynamical Systems*. Springer, Berlin, 1990.
- [84] P. Volkmann. Über die invarianz konvexer mengen und differentialungleichungen in einem normierten raume. *Mathematische Annalen*, 203(3):201–210, 1973.
- [85] X. Zhang and P. W. Glynn. Affine jump-diffusions: Stochastic stability and limit theorems. Technical report, Oct. 2018. ArXiv e-print 1811.00122.