The Annals of Applied Probability
2011, Vol. 21, No. 2, 397–463
DOI: 10.1214/10-AAP710

AFFINE PROCESSES ON POSITIVE SEMIDEFINITE MATRICES

BY CHRISTA CUCHIERO[1], DAMIR FILIPOVIĆ[2], EBERHARD MAYERHOFER[2] AND JOSEF TEICHMANN[1]

ETH Zürich, École Polytechnique Fédérale de Lausanne, Vienna Institute of Finance and ETH Zürich

This article provides the mathematical foundation for stochastically continuous affine processes on the cone of positive semidefinite symmetric matrices. This analysis has been motivated by a large and growing use of matrix-valued affine processes in finance, including multi-asset option pricing with stochastic volatility and correlation structures, and fixed-income models with stochastically correlated risk factors and default intensities.

CONTENTS

1. Introduction . . . 397
2. Definition and characterization of affine processes . . . 402
3. Affine processes are regular and Feller . . . 410
4. Necessary parameter restrictions . . . 413
5. Sufficient conditions for the existence and uniqueness of affine processes . . . 432
6. Proof of the main results . . . 446
Appendix A: Existence and viability of a class of jump-diffusions . . . 450
Appendix B: An approximation lemma on the cone of positive semidefinite matrices . . . 458
Acknowledgments . . . 460
References . . . 460

1. Introduction. This paper provides the mathematical foundation for stochastically continuous affine processes on the cone of positive semidefinite symmetric $d \times d$-matrices S_d^+. These matrix-valued affine processes have arisen from a large and growing range of useful applications in finance, including multi-asset option pricing with stochastic volatility and correlation structures, and fixed-income models with stochastically correlated risk factors and default intensities.

For illustration, let us consider a multi-variate stochastic volatility model consisting of a d-dimensional logarithmic price process with risk-neutral dynamics

$$dY_t = \left(r\mathbf{1} - \tfrac{1}{2}X_t^{\mathrm{diag}}\right)dt + \sqrt{X_t}\,dB_t, \qquad Y_0 = y, \tag{1.1}$$

Received October 2009; revised April 2010.
[1]Supported by FWF-Grant Y328 (START prize from the Austrian Science Fund).
[2]Supported by WWTF (Vienna Science and Technology Fund) and Swissquote.
MSC2010 subject classifications. Primary 60J25; secondary 91B70.
Key words and phrases. Affine processes, Wishart processes, stochastic volatility, stochastic invariance.

and stochastic covariation process $X = \langle Y, Y\rangle$, which is a proxy for the instantaneous covariance of the price returns. Here B denotes a standard d-dimensional Brownian motion, r the constant interest rate, $\mathbf{1}$ the vector whose entries are all equal to one and X^{diag} the vector containing the diagonal entries of X.

The necessity to specify X as a process in S_d^+ such that it qualifies as covariation process is one of the mathematically interesting and demanding aspects of such models. Beyond that, the modeling of X must allow for enough flexibility in order to reflect the stylized facts of financial data and to adequately capture the dependence structure of the different assets. If these requirements are met, the model can be used as a basis for financial decision-making in the area of portfolio optimization, pricing of multi-asset options and hedging of correlation risk.

The tractability of such a model crucially depends on the dynamics of X. A large part of the literature in the area of multivariate stochastic volatility modeling has proposed the following affine dynamics for X:

$$
\begin{aligned}
&dX_t = (b + HX_t + X_t H^\top)\,dt + \sqrt{X_t}\,dW_t\,\Sigma + \Sigma^\top\,dW_t^\top\sqrt{X_t} + dJ_t,\\
&X_0 = x \in S_d^+,
\end{aligned}
\tag{1.2}
$$

where b is some suitably chosen matrix in S_d^+, H, Σ some invertible matrices, W a standard $d \times d$-matrix of Brownian motions possibly correlated with B, and J a pure jump process whose compensator is an affine function of X.[3]

The main reason for the analytic tractability of this model is that, under some technical conditions, the following affine transform formula holds:

$$
\mathbb{E}_{x,y}\big[e^{-\operatorname{Tr}(zX_t)+v^\top Y_t}\big] = e^{\Phi(t,z,v)+\operatorname{Tr}(\Psi(t,z,v)x)+v^\top y}
$$

for appropriate arguments $z \in S_d \times iS_d$ and $v \in \mathbb{C}^d$. The functions Φ and Ψ solve a system of nonlinear ordinary differential equations (ODEs), which are determined by the model parameters. Setting $v = 0$, $\phi(t,z) = -\Phi(t,z,0)$ and $\psi(t,z) = -\Psi(t,z,0)$ and taking $z = u \in S_d^+$, we arrive at

$$
\mathbb{E}_x\big[e^{-\operatorname{Tr}(uX_t)}\big] = e^{-\phi(t,u)-\operatorname{Tr}(\psi(t,u)x)}, \qquad u \in S_d^+. \tag{1.3}
$$

In this paper, we characterize the class of all stochastically continuous time-homogeneous Markov processes with the key property (1.3)—henceforth called affine processes—on S_d^+. Our main result shows that an affine process is necessarily a Feller process whose generator has affine coefficients in the state variables. The parameters of the generator satisfy some well-determined admissibility conditions, and are in a one-to-one relation with those of the corresponding ODEs for ϕ and ψ. Conversely, and more importantly for applications, we show that for any

[3]This affine multi-variate stochastic volatility model generalizes the well-known one-dimensional models of Heston [27], for the diffusion case, or the Barndorff-Nielsen Shepard model [2], for the pure jump case.

admissible parameter set there exists a unique well-behaved affine process on S_d^+. Furthermore, we prove that any stochastically continuous infinitely decomposable Markov process on S_d^+ is affine with zero diffusion, and vice versa.

On the one hand, our findings extend the model class (1.2), since a more general drift and jumps are possible. Indeed, we allow for full generality in b, as long as $b - (d-1)\Sigma^T\Sigma \in S_d^+$, for a general linear drift part $B(x) = \sum_{ij} x_{ij}\beta^{ij}$ and for an inclusion of (infinite activity) jumps. This of course enables more flexibility in financial modeling. For example, due to the general linear drift part, the volatility of one asset can generally depend on the other ones, which is not possible for $B(x) = Hx + xH^\top$. On the other hand, we now know the exact assumptions under which affine processes on S_d^+ actually exist. Our characterization of affine processes on S_d^+ is thus exhaustive. Beyond that, the equivalence of infinitely decomposable Markov processes with state space S_d^+ and affine processes without diffusion is interesting in its own right.

This paper complements Duffie, Filipović and Schachermayer [16], who analyzed time-homogeneous affine processes on the state space $\mathbb{R}_+^m \times \mathbb{R}^n$.[4] Matrix-valued affine processes seem to have been studied systematically for the first time in the literature by Bru [5, 6], who introduced the so called Wishart processes. These are generalizations of squares of matrix Ornstein–Uhlenbeck processes, that is, of the form (1.2) for $J = 0$ and $b = k\Sigma^\top\Sigma$, for some real parameter $k > d-1$. Note that $k > d-1$ is a stronger assumption than what we require on b and $\Sigma^\top\Sigma$. Bru [6] then establishes existence and uniqueness of a local[5] S_d^+-valued solution to (1.2) under the additional assumptions that X_0 has distinct eigenvalues, $-H \in S_d^+$, and that H and Σ commute (see [6], Theorem 2″). In the more special case where $H = 0$ and $k > d-1$, Bru [6] shows global existence and uniqueness for (1.2) for any X_0 with distinct eigenvalues (see [6], Theorem 2 and last part of Section 3).[6] Bru's results concerning strong solutions have recently been extended to the case of matrix valued jump-diffusions; see [40].

Wishart processes have subsequently been introduced in the financial literature by Gourieroux and Sufana [24, 25] and Gourieroux et al. [23]. Financial applications thereof have then been taken up and carried further by various authors, including Da Fonseca et al. [9–12] and Buraschi, Cieslak and Trojani [7, 8]. Grasselli and Tebaldi [26] give some general results on the solvability of the corresponding Riccati ODEs. Barndorff-Nielsen and Stelzer [3] provide a theory for a certain class of matrix-valued Lévy driven Ornstein–Uhlenbeck processes of finite variation. Leippold and Trojani [38] introduce S_d^+-valued affine jump diffusions and

[4] For the diffusion case see also [20] or [19], Chapter 10. Time-inhomogeneous affine processes on $\mathbb{R}_+^m \times \mathbb{R}^n$ have been explored in [18].

[5] Up to the first collision time of the eigenvalues.

[6] Actually, Bru [6] establishes existence and uniqueness of solutions also for $k = 1, \dots, d-1$. But these are degenerate solutions, as they are only defined on lower-dimensional subsets of the boundary of S_d^+ (see [6], Corollary 1).

provide financial examples, including multi-variate option pricing, fixed-income models and dynamic portfolio choice. All of these models are contained in our framework.

We want to point out that the full characterization of positive semidefinite matrix-valued affine processes needs a multitude of methods. In order to prove the fundamental property of regularity of affine processes another adaption of the famous analysis of Montgommery and Zippin is necessary, which has been worked out in [34] and [35] for the state space $\mathbb{R}_+^m \times \mathbb{R}^n$. For the necessary conditions on drift, diffusion and jump parameters we need the theory of infinitely divisible distributions on S_d^+. Most interestingly, the constant drift part b must satisfy a condition depending on the magnitude of the diffusion component (see Proposition 4.18), which is in accordance with the choice of the drift in Bru's work [6] on Wishart processes, as explained above. This enigmatic additional condition on the drift b is derived by studying the process with respect to well chosen test functions, including in our case the determinant of the process. It is worth noting, as already visible in dimension one, that a naive application of classical geometric invariance conditions does not bring the correct necessary result on the drift but a stronger one. Indeed, take a one-dimensional affine diffusion process X solving

$$dX_t = b\,dt + \sqrt{X_t}\,dW_t.$$

Then a back-of-the-envelope calculation would yield the Stratonovich drift at the boundary point $x = 0$ of value $b - \frac{1}{4}$, leading to the necessary parameter restriction $b \geq \frac{1}{4}$, which is indeed too strong. It is well known that the correct parameter restriction is $b \geq 0$. We see two reasons why geometric conditions on the drift cannot be applied: first, precisely at the boundary of our state spaces the diffusion coefficients are not Lipschitz continuous anymore, and, second, the boundary of the cone of positive semi-definite matrices is not a smooth submanifold but a more complicated object.

For the sufficient direction refined methods from stochastic invariance theory are applied. Having established viability of a particular class of jump-diffusions on S_d^+, existence of affine processes on S_d^+—under the necessary parameter conditions—is shown through the solution of a martingale problem. Uniqueness follows by semigroup methods which need the theory of multi-dimensional Riccati equations.

Summing up, we face two major problems in the analysis of positive matrix valued affine processes. First, the candidate stochastic differential equations necessarily lead to volatility terms which are not Lipschitz continuous at the boundary of the state space. This makes every existence, uniqueness and invariance question delicate. Second, the jump behavior transversal to the boundary is of finite total variation.

1.1. *Program of the article.* For affine processes on S_d^+, results and proofs deviate in essential points from the theory on state spaces of the form $\mathbb{R}_+^m \times \mathbb{R}^n$ given in [16, 34], which is a consequence of the more involved geometry of this nonpolyhedral cone. The program of the paper as outlined below therefore includes a comparison with the approach in [16].

Section 2 contains the main definition and a summary of the results of this article. In Section 3, we then derive two main properties, namely the regularity of the process and the Feller property of the associated semigroup. The Feller property, in turn, is a simple consequence of an important positivity result of the characteristic exponents ϕ, ψ, which is proved in Lemma 3.3. This lemma is further employed as a tool for the treatment of the generalized Riccati differential equations in Section 5.1 (see proof of Proposition 5.3). The global existence and uniqueness of these equations is then used to show uniqueness of the martingale problem for affine processes (see proof of Proposition 5.9).

In Section 4, we define a set of admissible parameters specifying the infinitesimal generator of affine semigroups and prove the necessity of the parameter restrictions (see Proposition 4.9).

The sufficient direction is then treated in Section 5. It is known that, for $d \geq 2$, there exist continuous affine processes on S_d^+ which are—in contrast to those on the state space $\mathbb{R}_+^m \times \mathbb{R}^n$—*not* infinitely divisible (see Example 2.8). The analysis of this paper reveals the failure of infinite divisibility as a consequence of the drift condition (see proof of Theorem 2.9). This has substantial influence on the approach chosen here to prove existence of affine processes associated with a given parameter set: Being in general hindered to recognize the solutions of the generalized Riccati differential equations as cumulant generating functions of substochastic measures, as done in [16], Section 7, we solve the martingale problem for the associated Lévy type generator on S_d^+, as exposed in Section 5 and Appendix A. In Section 5.3, however, we deliver a variant of the existence proof of [16] for pure jump processes, which is possible in this case due to the absence of a diffusion component.

Finally, Section 6 contains the proofs of the main results which build on the propositions of the previous sections.

1.2. *Notation.* For the stochastic background and notation, we refer to standard text books such as [31] and [43]. We write $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_{++} = (0, \infty)$. Moreover:

- S_d denotes the space of symmetric $d \times d$-matrices equipped with the scalar product $\langle x, y \rangle = \mathrm{Tr}(xy)$. Note that S_d is isomorphic, but not isometric, to the standard Euclidean space $\mathbb{R}^{d(d+1)/2}$. We denote by $\{c^{ij}, i \leq j\}$ the standard basis of S_d, that is, the (kl)th component of c^{ij} is given by $c_{kl}^{ij} = \delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}(1 - \delta_{ij})$, where δ_{ij} denotes the Kronecker delta. Additionally, we sometimes consider the following basis elements $\{e^{ij}, i \le j\}$ which are positive semidefinite and form a basis of S_d:

$$e^{ij} = \begin{cases} c^{ii}, & \text{if } i = j, \\ c^{ii} + c^{ij} + c^{jj}, & \text{if } i \neq j. \end{cases}$$

- S_d^+ stands for the cone of symmetric $d \times d$-positive semidefinite matrices, S_d^{++} for its interior in S_d, the cone of strictly positive definite matrices. The boundary is denoted by $\partial S_d^+ = S_d^+ \setminus S_d^{++}$, the complement is denoted by $(S_d^+)^c$, and $S_d^+ \cup \{\Delta\}$ denotes the one-point compactification. Recall that S_d^+ is self-dual [w.r.t. the scalar product $\langle x, y\rangle = \mathrm{Tr}(xy)$], that is,

$$S_d^+ = \{x \in S_d \mid \langle x, y\rangle \ge 0, \forall y \in S_d^+\}.$$

  Both cones, S_d^+ and S_d^{++}, induce a partial and strict order relation on S_d, respectively: we write $x \preceq y$ if $y - x \in S_d^+$, and $x \prec y$ if $y - x \in S_d^{++}$.
- M_d is the space of $d \times d$-matrices and $O(d)$ the orthogonal group of dimension d over $\mathbb{R}$.
- I_d denotes the $d \times d$-identity matrix.

Throughout this paper, a function $f : S_d \to \mathbb{R}$ is understood as the restriction $f = g|_{S_d}$ of a function $g : M_d \to \mathbb{R}$ which satisfies $g(x) = g(x^\top)$ for all $x \in M_d$. Without loss of generality $g(x) = f((x + x^\top)/2)$. We avoid using the *vech* operator, that is, to identify $x \in S_d$ with a vector in $\mathbb{R}^{d(d+1)/2}$ by stringing the columns of x together, while only taking the entries x_{ij} with $i \le j$.

Throughout this article, we shall consider the following function spaces for measurable $U \subseteq S_d$. We write $\mathcal{B}(U)$ for the Borel σ-algebra on U. bU corresponds to the Banach space of bounded real-valued Borel measurable functions f on U with norm $\|f\|_\infty = \sup_{x \in U} |f(x)|$. We write $C(U)$ for the space of real-valued continuous functions f on U, $C_b(U)$ for $C(U) \cap bU$, $C_c(U)$ for the space of functions $f \in C(U)$ with compact support and $C_0(U)$ for the Banach space of functions $f \in C(U)$ with $\lim_{x \to \Delta} f(x) = 0$ and norm $\|f\|_\infty = \sup_{x \in U} |f(x)|$. Furthermore, $C^k(U)$ is the space of k times differentiable functions f on U°, the interior of U, such that all partial derivatives of f up to order k belong to $C(U)$. As usual, we set $C^\infty(U) = \bigcap_{k \ge 1} C^k(U)$, and we write $C_c^k(U) = C_c(U) \cap C^k(U)$ and $C_b^k(U) = C_b(U) \cap C^k(U)$, for $k \le \infty$.

2. Definition and characterization of affine processes. We consider a time-homogeneous Markov process X with state space S_d^+ and semigroup $(P_t)_{t \ge 0}$ acting on functions $f \in bS_d^+$,

$$P_t f(x) = \int_{S_d^+} f(\xi) p_t(x, d\xi), \qquad x \in S_d^+.$$

We note that X may not be conservative. Then there is a standard extension of the transition probabilities to the one-point compactification $S_d^+ \cup \{\Delta\}$ of S_d^+ by defining

$$p_t(x, \{\Delta\}) = 1 - p_t(x, S_d^+), \qquad p_t(\Delta, \{\Delta\}) = 1$$

for all t and $x \in S_d^+$, with the convention that $f(\Delta) = 0$ for any function f on S_d^+. Thus X becomes conservative on $S_d^+ \cup \{\Delta\}$.

DEFINITION 2.1. The Markov process X is called *affine* if:

(i) it is stochastically continuous, that is, $\lim_{s\to t} p_s(x, \cdot) = p_t(x, \cdot)$ weakly on S_d^+ for every t and $x \in S_d^+$, and

(ii) its Laplace transform has exponential-affine dependence on the initial state

$$P_t e^{-\langle u,x\rangle} = \int_{S_d^+} e^{-\langle u,\xi\rangle} p_t(x, d\xi) = e^{-\phi(t,u)-\langle\psi(t,u),x\rangle}, \tag{2.1}$$

for all t and $u, x \in S_d^+$, for some functions $\phi : \mathbb{R}_+ \times S_d^+ \to \mathbb{R}_+$ and $\psi : \mathbb{R}_+ \times S_d^+ \to S_d^+$.

Note that stochastic continuity of X implies that $\phi(t, u)$ and $\psi(t, u)$ are jointly continuous in (t, u); see Lemma 3.2(iii) below. Moreover, due to the Markov property, this also means that $p_t(x, \{\Delta\}) < 1$ for all $x \in S_d^+$ and $t \geq 0$. In contrast to [16], we take stochastic continuity as part of the definition of affine processes, and consider the Laplace transform instead of the characteristic function. The latter is justified by the nonnegativity of X, the former is by convenience since, as we will see in Proposition 3.4 below, it automatically implies regularity in the following sense.

DEFINITION 2.2. The affine process X is called *regular* if the derivatives

$$F(u) = \left.\frac{\partial\phi(t, u)}{\partial t}\right|_{t=0+}, \qquad R(u) = \left.\frac{\partial\psi(t, u)}{\partial t}\right|_{t=0+} \tag{2.2}$$

exist and are continuous at $u = 0$.

We remark that there are simple examples of Markov processes which satisfy Definition 2.1(ii) but are not stochastically continuous; see [16], Remark 2.11. However, such processes are of limited interest for applications and will not be considered.

In the following, we shall provide an equivalent characterization of the affine property in terms of the generator of X. As we shall see in (2.12), the diffusion, drift, jump and killing characteristics of X depend in an affine way on the underlying state. We denote by $\chi : S_d \to S_d$ some bounded continuous truncation function with $\chi(\xi) = \xi$ in a neighborhood of 0. Then the involved parameters are admissible in the following sense.

DEFINITION 2.3. An *admissible parameter set* $(\alpha, b, \beta^{ij}, c, \gamma, m, \mu)$ associated with χ consists of:

- a linear diffusion coefficient

$$\alpha \in S_d^+, \tag{2.3}$$

- a constant drift term

$$b \succeq (d-1)\alpha, \tag{2.4}$$

- a constant killing rate term

$$c \in \mathbb{R}^+, \tag{2.5}$$

- a linear killing rate coefficient

$$\gamma \in S_d^+, \tag{2.6}$$

- a constant jump term: a Borel measure m on $S_d^+ \setminus \{0\}$ satisfying

$$\int_{S_d^+ \setminus \{0\}} (\|\xi\| \wedge 1) m(d\xi) < \infty, \tag{2.7}$$

- a linear jump coefficient: a $d \times d$-matrix $\mu = (\mu_{ij})$ of finite signed measures on $S_d^+ \setminus \{0\}$ such that $\mu(E) \in S_d^+$ for all $E \in \mathcal{B}(S_d^+ \setminus \{0\})$ and the kernel

$$M(x, d\xi) := \frac{\langle x, \mu(d\xi)\rangle}{\|\xi\|^2 \wedge 1} \tag{2.8}$$

  satisfies

$$\int_{S_d^+ \setminus \{0\}} \langle \chi(\xi), u\rangle M(x, d\xi) < \infty \qquad \text{for all } x, u \in S_d^+ \text{ with } \langle x, u\rangle = 0, \tag{2.9}$$

- a linear drift coefficient: a family $\beta^{ij} = \beta^{ji} \in S_d$ such that the linear map $B : S_d \to S_d$ of the form

$$B(x) = \sum_{i,j} \beta^{ij} x_{ij} \tag{2.10}$$

  satisfies

$$\begin{aligned} &\langle B(x), u\rangle - \int_{S_d^+ \setminus \{0\}} \langle \chi(\xi), u\rangle M(x, d\xi) \geq 0 \\ &\qquad\qquad \text{for all } x, u \in S_d^+ \text{ with } \langle x, u\rangle = 0. \end{aligned} \tag{2.11}$$

We shall comment more on the admissibility conditions in Section 2.1 below. The following three theorems contain the main results of this article. Their proofs are given in Section 6. First, we provide a characterization of affine processes on S_d^+ in terms of the admissible parameter set introduced in Definition 2.3. As for the domain of the generator, we consider the space $\mathcal{S}_+$ of rapidly decreasing C^∞-functions on S_d^+, defined in (B.1) below. It is shown in Appendix B that $e^{-\langle u, \cdot\rangle} \in \mathcal{S}_+$, for $u \in S_d^{++}$, as well as $C_c^\infty(S_d^+) \subset \mathcal{S}_+$.

THEOREM 2.4. *Suppose X is an affine process on S_d^+. Then X is regular and has the Feller property. Let $\mathcal{A}$ be its infinitesimal generator on $C_0(S_d^+)$. Then $\mathcal{S}_+ \subset D(\mathcal{A})$ and there exists an admissible parameter set* $(\alpha, b, \beta^{ij}, c, \gamma, m, \mu)$

such that, for $f \in \mathcal{S}_+$,

$$
\begin{aligned}
\mathcal{A}f(x) &= \frac{1}{2}\sum_{i,j,k,l} A_{ijkl}(x)\,\frac{\partial^2 f(x)}{\partial x_{ij}\,\partial x_{kl}} \\
&\quad + \sum_{i,j}\big(b_{ij} + B_{ij}(x)\big)\,\frac{\partial f(x)}{\partial x_{ij}} - (c + \langle \gamma, x\rangle) f(x) \\
&\quad + \int_{S_d^+\setminus\{0\}} \big(f(x+\xi) - f(x)\big) m(d\xi) \\
&\quad + \int_{S_d^+\setminus\{0\}} \big(f(x+\xi) - f(x) - \langle \chi(\xi), \nabla f(x)\rangle\big) M(x, d\xi),
\end{aligned}
\tag{2.12}
$$

where $B(x)$ *is defined by* (2.10), $M(x, d\xi)$ *by* (2.8) *and*

$$
A_{ijkl}(x) = x_{ik}\alpha_{jl} + x_{il}\alpha_{jk} + x_{jk}\alpha_{il} + x_{jl}\alpha_{ik}. \tag{2.13}
$$

Moreover, $\phi(t,u)$ *and* $\psi(t,u)$ *in* (2.1) *solve the generalized Riccati differential equations, for* $u \in S_d^+$,

$$
\frac{\partial \phi(t,u)}{\partial t} = F(\psi(t,u)), \qquad \phi(0,u) = 0, \tag{2.14}
$$

$$
\frac{\partial \psi(t,u)}{\partial t} = R(\psi(t,u)), \qquad \psi(0,u) = u, \tag{2.15}
$$

with

$$
F(u) = \langle b, u\rangle + c - \int_{S_d^+\setminus\{0\}} \big(e^{-\langle u,\xi\rangle} - 1\big) m(d\xi), \tag{2.16}
$$

$$
\begin{aligned}
R(u) &= -2u\alpha u + B^\top(u) + \gamma \\
&\quad - \int_{S_d^+\setminus\{0\}} \left(\frac{e^{-\langle u,\xi\rangle} - 1 + \langle \chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1}\right) \mu(d\xi),
\end{aligned}
\tag{2.17}
$$

where $B_{ij}^\top(u) = \langle \beta^{ij}, u\rangle$.

Conversely, let $(\alpha, b, \beta^{ij}, c, \gamma, m, \mu)$ *be an admissible parameter set. Then there exists a unique affine process on* S_d^+ *with infinitesimal generator* (2.12) *and* (2.1) *holds for all* $(t,u) \in \mathbb{R}_+ \times S_d^+$, *where* $\phi(t,u)$ *and* $\psi(t,u)$ *are given by* (2.14) *and* (2.15).

REMARK 2.5. It can be proved as in [39] that X is conservative if and only if $c = 0$ and $\psi(t,0) \equiv 0$ is the only S_d^+-valued local solution of (2.15) for $u = 0$. The latter condition clearly requires that $\gamma = 0$.

Hence, a sufficient condition for X to be conservative is $c = 0$ and $\gamma = 0$ and

$$
\int_{S_d^+ \cap \{\|\xi\| \ge 1\}} \|\xi\| \big(\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)\big) < \infty \qquad \text{for all } 1 \le i \le j \le d,
$$

where $\mu_{ij} = \mu_{ij}^+ - \mu_{ij}^-$ denotes the Jordan decomposition of μ_{ij}. Indeed, it can be shown similarly as in [16], Section 9, that the latter property implies Lipschitz continuity of $R(u)$ on S_d.

Due to the Feller property, as established in Theorem 2.4, any affine process X on S_d^+ admits a càdlàg modification, still denoted by X (see, e.g., [43], Chapter III.2). It can and will thus be realized on the space $\Omega = \mathbb{D}(S_d^+ \cup \{\Delta\})$ of càdlàg paths $\omega : \mathbb{R}_+ \to S_d^+ \cup \{\Delta\}$ with $\omega(s) = \Delta$ for $s > t$ whenever $\omega(t-) = \Delta$ or $\omega(t) = \Delta$. For every $x \in S_d^+$, we denote by $\mathbb{P}_x$ the law of X given $X_0 = x$ and by $(\mathcal{F}_t^X)$ the natural filtration generated by X_t. We also consider the usual augmentation

$$\widetilde{\mathcal{F}}_t := \bigcap_{x \in S_d^+} \mathcal{F}_t^{(x)} \tag{2.18}$$

of $(\mathcal{F}_t^X)$, where $(\mathcal{F}_t^{(x)})$ is the augmentation of $(\mathcal{F}_t^X)$ with respect to $\mathbb{P}_x$. Then $(\widetilde{\mathcal{F}}_t)$ is right continuous and X is still a Markov process under $(\widetilde{\mathcal{F}}_t)$. We shall now relate conservative affine processes to semimartingales, where semimartingales are understood with respect to the stochastic basis $(\Omega, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}})_t, \mathbb{P}_x)$ for every x.

THEOREM 2.6. *Let X be a conservative affine process on S_d^+ and let $(\alpha, b, \beta^{ij}, c = 0, \gamma = 0, m, \mu)$ be the related admissible parameter set associated with the truncation function χ. Then X is a semimartingale whose characteristics (B, A, ν) with respect to χ are given by*

$$A_{t,ijkl} = \int_0^t A_{ijkl}(X_s)\, ds, \tag{2.19}$$

$$B_t = \int_0^t \left(b + \int_{S_d^+ \setminus \{0\}} \chi(\xi) m(d\xi) + B(X_s) \right) ds, \tag{2.20}$$

$$\nu(dt, d\xi) = \big(m(d\xi) + M(X_t, d\xi)\big)\, dt, \tag{2.21}$$

where $B(x)$ is given by (2.10), $A_{ijkl}(x)$ by (2.13) and $M(x, d\xi)$ by (2.8). Furthermore, there exists, possibly on an enlargement of the probability space, a $d \times d$-matrix of standard Brownian motions W such that X admits the following representation:

$$\begin{aligned} X_t = x + B_t &+ \int_0^t \left(\sqrt{X_s}\, dW_s\, \Sigma + \Sigma^\top dW_s \sqrt{X_s}\right) \\ &+ \int_0^t \int_{S_d^+ \setminus \{0\}} \chi(\xi)\big(\mu^X(ds, d\xi) - \nu(ds, d\xi)\big) \\ &+ \int_0^t \int_{S_d^+ \setminus \{0\}} (\xi - \chi(\xi))\mu^X(ds, d\xi), \end{aligned} \tag{2.22}$$

where $\Sigma \in M_d$ *satisfies* $\Sigma^\top \Sigma = \alpha$ *and* μ^X *denotes the random measure associated with the jumps of* X.

Hence, X *is continuous if and only if* m *and* μ *vanish.*

Let $\mathcal{P}$ be the set of all families of probability measures $(\mathbb{P}_x)_{x \in S_d^+}$ on the canonical probability space $(\Omega, \mathcal{F}^X)$ such that $(X, (\mathbb{P}_x)_{x \in S_d^+})$ is a stochastically continuous Markov processes on S_d^+ with $\mathbb{P}_x[X_0 = x] = 1$, for all $x \in S_d^+$. Note that in contrast to [16], there is no need to impose regularity of X. For two probability measures $\mathbb{P}, \mathbb{Q}$ on $(\Omega, \mathcal{F}^X)$, the convolution $\mathbb{P} * \mathbb{Q}$ is defined as the push-forward of $\mathbb{P} \times \mathbb{Q}$ under the map $(\omega, \omega') \mapsto \omega + \omega' : (\Omega \times \Omega, \mathcal{F}^X \otimes \mathcal{F}^X) \to (\Omega, \mathcal{F}^X)$.

DEFINITION 2.7. An element $(\mathbb{P}_x)_{x \in S_d^+} \in \mathcal{P}$ is called:

(i) infinitely decomposable, if for each $k \geq 1$, there exists $(\mathbb{P}_x^{(k)})_{x \in S_d^+} \in \mathcal{P}$ such that

$$\mathbb{P}_{x^{(1)} + \cdots + x^{(k)}} = \mathbb{P}^{(k)}_{x^{(1)}} * \cdots * \mathbb{P}^{(k)}_{x^{(k)}};$$

(ii) infinitely divisible, if the one-dimensional marginal distributions $\mathbb{P}_x \circ X_t^{-1}$ are infinitely divisible, for all $(t, x) \in \mathbb{R}_+ \times S_d^+$.

In [16] it was shown that regular affine processes on $\mathbb{R}_+^m \times \mathbb{R}^n$ are infinitely decomposable Markov processes, and vice versa. In fact, this property was the core for the existence proof of affine processes in [16]. On S_d^+ the situation is different. The following counterexample reveals that not all affine processes on S_d^+ are infinitely divisible.

EXAMPLE 2.8. The affine process X on S_d^+ corresponding to the parameter set $(\alpha = I_d, b = \delta I_d, 0, 0, 0, 0, 0)$, where $\delta \in [d-1, \infty)$, is the diffusion process initially studied by Bru [6]. By [6], Theorem 3, the Laplace-transforms

$$\mathbb{E}_x\big[e^{-\langle X_t, u\rangle}\big] = \big(\det(I + 2tu)\big)^{-\delta/2} e^{-\langle (I+2tu)^{-1}u, x\rangle}$$

are those of noncentral Wishart distributions $\mathrm{WIS}(\delta, d, x)$. By a well-known result due to Paul Lévy, these Wishart distributions are not infinitely divisible if $d \geq 2$ (see [15], Section 2.C).

Here, is our main result on infinite divisibility of affine processes on S_d^+.

THEOREM 2.9. *Let* $d \geq 2$ *and* $(\mathbb{P}_x)_{x \in S_d^+} \in \mathcal{P}$. *The following assertions are equivalent*:

(i) $(\mathbb{P}_x)_{x \in S_d^+}$ *is infinitely decomposable.*
(ii) $(X, (\mathbb{P}_x)_{x \in S_d^+})$ *is affine with vanishing diffusion parameter* $\alpha = 0$.
(iii) $(X, (\mathbb{P}_x)_{x \in S_d^+})$ *is affine and infinitely divisible.*

2.1. *Discussion of the parameters.* We discuss and highlight some properties of the admissible parameter set $(\alpha, b, \beta^{ij}, c, \gamma, m, \mu)$ of an affine process.

Let us therefore define the normal cone

$$N_{S_d^+}(x) = \{u \in S_d^+ \mid \langle u, x\rangle = 0\}, \tag{2.23}$$

containing the inward pointing normal vectors, to S_d^+ at $x \in S_d^+$.[7] It will be shown in Lemma 4.1 below that $N_{S_d^+}(x) \neq \{0\}$ only for boundary elements $x \in \partial S_d^+$.

2.1.1. *Diffusion.* The diffusion term does not admit a constant part, and its linear part is of the very specific form

$$\langle u, A(x)u\rangle = 4\langle x, u\alpha u\rangle.$$

This property of $A(x)$ has also been stated in the setting of symmetric cones in [26]. We could thus write the second order differential operator in (2.12) as

$$\frac{1}{2}\sum_{i,j,k,l} A_{ijkl}(x)\frac{\partial^2 f(x)}{\partial x_{ij}\,\partial x_{kl}} = 2\langle x, \nabla\alpha\nabla f(x)\rangle.$$

The reason why we introduce and use the symmetrization (2.13) of $A(x)$ is that it corresponds to the quadratic characteristic (2.19) of the semimartingale X.

2.1.2. *Drift.* The remarkable drift condition (2.4) has been assumed in many previous papers. Here is the first time where necessity and sufficiency of (2.4) are proved in the full generality in the presence of jumps. Note that in dimension $d = 1$, the drift condition simply reduces to nonnegativity $b \geq 0$. But for dimension $d \geq 2$, the boundary of the state space S_d^+ becomes curved and kinked, implying a nontrivial trade-off between diffusion α and b.

Concerning the form of B, let us note the following: condition (2.11) implies in particular

$$\beta^{ii}_{\setminus\{i\}} - \int_{S_d^+\setminus\{0\}} \frac{\chi(\xi)_{\setminus\{i\}}}{\|\xi\|^2 \wedge 1}\mu_{ii}(d\xi) \in S_{d-1}^+ \qquad \text{for all } 1 \leq i \leq d, \tag{2.24}$$

where for any matrix $u \in S_d$, $u_{\setminus\{i\}}$ denotes the matrix where the ith row and column are deleted. Indeed, inserting $x = c^{ii}$ in condition (2.11) yields

$$\langle B(c^{ii}), u\rangle - \int_{S_d^+\setminus\{0\}} \frac{\langle\chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1}\mu_{ii}(d\xi) \geq 0$$

[7] Indeed, we obtain (2.23) from the general definition in (A.8) below by choosing $y = 0$ and $y = 2x$, and using the self-duality of S_d^+: $\langle u, y\rangle \geq 0$ for all $y, u \in S_d^+$.

for all $u \in S_d^+$ with $\langle c^{ii}, u\rangle = 0$. Since the ith column and row of such an element $u \in S_d^+$ is zero, it follows that

$$\begin{aligned}&\langle B(c^{ii}), u\rangle - \int_{S_d^+\setminus\{0\}} \frac{\langle \chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1}\mu_{ii}(d\xi)\\ &\quad = \langle \beta^{ii}_{\setminus\{i\}}, u_{\setminus\{i\}}\rangle - \int_{S_d^+\setminus\{0\}} \frac{\langle \chi(\xi)_{\setminus\{i\}}, u_{\setminus\{i\}}\rangle}{\|\xi\|^2 \wedge 1}\mu_{ii}(d\xi) \geq 0.\end{aligned} \tag{2.25}$$

By choosing appropriate elements $u_{\setminus\{i\}} \in S_{d-1}^+$, we can further derive the integrability of $\chi(\xi)_{kl}$ for all $k \neq i, l \neq i$, which implies

$$\int_{S_d^+\setminus\{0\}} \frac{\langle \chi(\xi)_{\setminus\{i\}}, u_{\setminus\{i\}}\rangle}{\|\xi\|^2 \wedge 1}\mu_{ii}(d\xi) = \left\langle \int_{S_d^+\setminus\{0\}} \frac{\chi(\xi)_{\setminus\{i\}}}{\|\xi\|^2 \wedge 1}\mu_{ii}(d\xi), u_{\setminus\{i\}}\right\rangle. \tag{2.26}$$

As (2.25) and (2.26) must hold true for all $u_{\setminus\{i\}} \in S_{d-1}^+$, assertion (2.24) is proved.

Note that the (ij)th component of the adjoint operator $B^\top$ is given by

$$B^\top_{ij}(u) = \langle \beta^{ij}, u\rangle, \tag{2.27}$$

since $\langle B(x), u\rangle = \langle \sum_{i,j} \beta^{ij} x_{ij}, u\rangle = \sum_{i,j}\langle \beta^{ij}, u\rangle x_{ij} = \langle B^\top(u), x\rangle$.

In most previous papers, $B(x)$ is of the form

$$B(x) = Hx + xH^\top. \tag{2.28}$$

In this case,

$$\langle B(x), u\rangle = \langle Hx + xH^\top, u\rangle = 0 \qquad \text{for all } x, u \in S_d^+ \text{ with } \langle x, u\rangle = 0, \tag{2.29}$$

and hence (2.11) is equivalent to

$$\int_{S_d^+\setminus\{0\}} \langle \chi(\xi), u\rangle M(x, d\xi) = 0,$$

for all $x, u \in S_d^+$ with $\langle x, u\rangle = 0$.

If $B(x)$ is of the form

$$B(x) = Hx + xH^\top + \Gamma(x), \tag{2.30}$$

where $H \in M_d$ and $\Gamma : S_d \to S_d$ linear satisfying $\Gamma(S_d^+) \subseteq S_d^+$, then, in view of (2.29), condition (2.11) holds true as long as

$$\langle \Gamma(x), u\rangle - \int_{S_d^+\setminus\{0\}} \langle \chi(\xi), u\rangle M(x, d\xi) \geq 0$$

for all $x, u \in S_d^+$ with $\langle x, u\rangle = 0$. As a bold conjecture, we claim that any $B(x)$ satisfying (2.11) is of form (2.30).

Here is a simple example where $B(x)$ is of the form (2.30) but not of the usual form (2.28): let $d = 2$ and

$$B(x) = \begin{pmatrix} x_{22} & x_{12} \\ x_{12} & x_{11} \end{pmatrix}.$$

It can be easily checked that (2.11) is satisfied, while $B(x)$ cannot be brought into the form (2.28). If x_{ii} models the (squared) volatility of the ith stock price, as in (1.1), then this drift specification admits level impacts of the volatility of stock 1 on the volatility of stock 2, and vice versa.

2.1.3. *Killing*. See Remark 2.5.

2.1.4. *Jumps*. Condition (2.7) means that jumps described by m, which can for instance appear at $x = 0$, should be of finite variation entering the cone S_d^+, since infinite variation transversal to the boundary would let the process leave the state space. Similarly, condition (2.9) asserts finite variation for the inward pointing directions, while we could a priori have a general jump behavior (supported by S_d^+ due to the affine structure) parallel to the boundary. Note that in the case $d = 1$, which corresponds to $\mathbb{R}_+$, the linear jump part can have infinite total variation (see [16], equation (2.11)). However, due to the geometry of the cone S_d^+, we conjecture that in higher dimensions $d \geq 2$ such a behavior is no longer possible and that all jumps are in fact of finite total variation. In any case, for $d \geq 2$, affine positive matrix valued diffusion processes cannot be approximated (in law) by pure jump processes, since this would yield a contradiction to condition (2.4). See also Remark 5.12 below.

3. Affine processes are regular and Feller. Suppose X is an affine process on S_d^+. The main result of this section is that X is regular in the sense of Definition 2.1. In addition, we shall prove that P_t is a Feller semigroup on $C_0(S_d^+)$. In order to show both properties, we shall mainly rely on Lemma 3.3 below. The Feller property is then a simple consequence of this statement and regularity is obtained by arguing as in Keller-Ressel, Schachermayer and Teichmann [35], who obtained the corresponding statements for affine processes on the state space $\mathbb{R}_+^m \times \mathbb{R}^n$; see [35], Theorem 4.3, and also the Ph.D. thesis of Keller-Ressel [34]. We observe that most arguments of [35] translate to our setting without major changes. It is only required to tailor some technicalities to the cone S_d^+. We start with the following elementary observations.

LEMMA 3.1. *If $u \in \partial S_d^+$ and $S_d^+ \ni v \preceq u$, then $v \in \partial S_d^+$.*

PROOF. Let $x \in S_d^+ \setminus \{0\}$ such that $\langle x, u\rangle = 0$. Then, $S_d^+ \ni v \preceq u$ implies $0 \leq \langle v, x\rangle \leq \langle u, x\rangle = 0$. Hence, $v \in \partial S_d^+$. □

We now derive some first properties of the functions ϕ and ψ in (2.1).

LEMMA 3.2. *Let X be an affine process on S_d^+. Then, we have*:

(i) *The functions* ϕ *and* ψ *satisfy*

$$\phi(t+s,u)=\phi(t,u)+\phi(s,\psi(t,u)), \tag{3.1}$$

$$\psi(t+s,u)=\psi(s,\psi(t,u)) \tag{3.2}$$

for all $t,s\in\mathbb{R}_+$.

(ii) *For all* $u,v\in S_d^+$ *with* $v\preceq u$ *and for all* $t\geq 0$, *the order relations*

$$\phi(t,v)\leq\phi(t,u)\quad\textit{and}\quad\psi(t,v)\preceq\psi(t,u) \tag{3.3}$$

hold true.

(iii) *The functions* ϕ *and* ψ *are jointly continuous in* $\mathbb{R}_+\times S_d^+$. *Furthermore*, $u\mapsto\phi(t,u)$ *and* $u\mapsto\psi(t,u)$ *are analytic on* S_d^{++}.

PROOF. Assertion (i) follows directly from the Chapman–Kolmogorov equation,

$$\begin{aligned}e^{-\phi(t+s,u)-\langle\psi(t+s,u),x\rangle}&=\int_{S_d^+}p_s(x,d\xi)\int_{S_d^+}e^{-\langle u,\widetilde{\xi}\rangle}p_t(\xi,d\widetilde{\xi})\\&=e^{-\phi(t,u)}\int_{S_d^+}e^{-\langle\psi(t,u),\xi\rangle}p_s(x,d\xi)\\&=e^{-\phi(t,u)-\phi(s,\psi(t,u))-\langle\psi(s,\psi(t,u)),x\rangle}.\end{aligned}$$

For the proof of (ii), note that $v\preceq u$ is equivalent to $\langle v,x\rangle\leq\langle u,x\rangle$ for all $x\in S_d^+$. By the monotonicity of the exponential function, we have for all $x\in S_d^+$ and for all $t\geq 0$,

$$\begin{aligned}e^{-\phi(t,v)-\langle\psi(t,v),x\rangle}&=\int_{S_d^+}e^{-\langle v,\xi\rangle}p_t(x,d\xi)\geq\int_{S_d^+}e^{-\langle u,\xi\rangle}p_t(x,d\xi)\\&=e^{-\phi(t,u)-\langle\psi(t,u),x\rangle},\end{aligned}$$

and the assertion follows by taking logarithms.

Concerning statement (iii), note that stochastic continuity of X implies joint continuity of $P_te^{-\langle u,x\rangle}$ in $(t,u)\in\mathbb{R}_+\times S_d^+$ (this follows, e.g., from [4], Lemma 23.7), for all $x\in S_d^+$. This in turn yields continuity of the functions $(t,u)\mapsto\phi(t,u)$ and $(t,u)\mapsto\psi(t,u)$. The second assertion follows from analyticity properties of the Laplace transform. □

The following property of ψ is crucial.

LEMMA 3.3. *Let* $\psi:\mathbb{R}_+\times S_d^+\to S_d^+$ *be any map satisfying* $\psi(0,u)=u$ *and the properties* (i)–(iii) *of Lemma* 3.2 (*regarding the function* ψ). *Then* $\psi(t,u)\in S_d^{++}$ *for all* $(t,u)\in\mathbb{R}_+\times S_d^{++}$.

PROOF. We adapt the proof of [34], Proposition 1.10, to our setting. Assume by contradiction that there exists some $(t,u) \in \mathbb{R}_+ \times S_d^{++}$ such that $\psi(t,u) \in \partial S_d^+$. Let us consider the interval $(0, \lambda_{\min}(u)] \neq \varnothing$, where $\lambda_{\min}(u) > 0$ denotes the smallest eigenvalue of u. Then for all $v \in (0, \lambda_{\min}(u)]$ we have $vI_d \preceq u$. Since $\psi(t,u)$ admits property (ii) of Lemma 3.2, we obtain

$$S_d^+ \ni \psi(t, vI_d) \preceq \psi(t,u) \in \partial S_d^+$$

for all $v \in (0, \lambda_{\min}(u)]$. Consequently, Lemma 3.1 yields that $\psi(t, vI_d) \in \partial S_d^+$. Hence,

$$\det(\psi(t, vI_d)) = 0$$

for all $v \in (0, \lambda_{\min}(u)]$. The analyticity of $u \mapsto \psi(t,u)$ on S_d^{++} carries over to $u \mapsto \det(\psi(t,u))$ and implies that $\det(\psi(t, vI_d)) = 0$ for all $v \in \mathbb{R}_{++}$. Indeed, the set of zeros of $\det(\psi(t, vI_d))$ has an accumulation point in $\mathbb{R}_{++}$, which implies that $\det(\psi(t, vI_d))$ vanishes entirely on $\mathbb{R}_{++}$. The same statement holds true for t replaced by $\frac{t}{2}$. Indeed, if $\psi(\frac{t}{2}, u) \in \partial S_d^+$, then the assertion is shown by the same arguments as above. Otherwise, if $\psi(\frac{t}{2}, u) \in S_d^{++}$, we have for all $v \in \mathbb{R}_{++}$ with $vI_d \preceq \psi(\frac{t}{2}, u)$, that is, for all $v \in (0, \lambda_{\min}(\psi(\frac{t}{2}, u)]$

$$S_d^+ \ni \psi\left(\frac{t}{2}, vI_d\right) \preceq \psi\left(\frac{t}{2}, \psi\left(\frac{t}{2}, u\right)\right) = \psi(t,u) \in \partial S_d^+,$$

which yields again $\psi(\frac{t}{2}, vI_d) \in \partial S_d^+$ and $\det(\psi(\frac{t}{2}, vI_d)) = 0$ for all $v \in (0, \lambda_{\min}(\psi(\frac{t}{2}, u)]$. The same reasoning as before then leads to $\det(\psi(\frac{t}{2}, vI_d)) = 0$ for all $v \in \mathbb{R}_{++}$. By reapplying this argument, we finally get for every $n \in \mathbb{N}$ and for all $v \in \mathbb{R}_{++}$

$$\det\left(\psi\left(\frac{t}{2^n}, vI_d\right)\right) = 0.$$

From the continuity of the function $t \mapsto \psi(t,u)$ and of the determinant, we deduce that for any $v \in \mathbb{R}_{++}$,

$$0 = \lim_{n\to\infty} \det\left(\psi\left(\frac{t}{2^n}, vI_d\right)\right) = \det(\psi(0, vI_d)) = \det(vI_d) = v^d > 0,$$

a contradiction, and the assertion is proved. □

We may now formulate the main result of this section.

PROPOSITION 3.4. *Let X be an affine process with state space S_d^+. Then, we have*:

(i) *X is a Feller process.*
(ii) *X is regular.*

PROOF. In order to prove (i), it suffices to show that for all $f \in C_0(S_d^+)$

$$\lim_{t\to 0+} P_t f(x) = f(x) \qquad \text{for all } x \in S_d^+, \tag{3.4}$$

$$P_t f \in C_0(S_d^+) \qquad \text{for all } t \in \mathbb{R}_+, \tag{3.5}$$

(see, e.g., [43], Propostion III.2.4). Property (3.4) is a consequence of stochastic continuity, which implies for all $f \in C_0(S_d^+)$ and $x \in S_d^+$

$$\lim_{t\to 0^+} P_t f(x) = f(x).$$

Concerning (3.5), it suffices to verify this property for a dense subset of $C_0(S_d^+)$. By a locally compact version of Stone–Weierstrass' theorem (see, e.g., [48]), the linear span of the set $\{e^{-\langle u,x\rangle} \mid u \in S_d^{++}\}$ is dense in $C_0(S_d^+)$. Indeed, it is a subalgebra of $C_0(S_d^+)$, separates points and vanishes nowhere, as all elements are strictly positive functions on S_d^+. From Lemma 3.3, we can deduce that $P_t e^{-\langle u,x\rangle} \in C_0(S_d^+)$ if $u \in S_d^{++}$, since $\psi(t,u) \in S_d^{++}$ and $\langle \psi(t,u), x\rangle > 0$ for $x \neq 0$ implying that

$$P_t e^{-\langle u,x\rangle} = e^{-\phi(t,u)-\langle \psi(t,u),x\rangle}$$

goes to 0 as $x \to \Delta$. Hence, statement (i) is proved.

The proof of (ii) follows precisely the lines of [35], proof of Theorem 4.3. Using Lemma 3.3, one may mimic the proof of [35], Theorem 4.3, to obtain that differentiability of $\psi(t,u)$ in $u \in S_d^{++}$, which follows from Lemma 3.2(iii), implies differentiability of $\psi(t,u)$ in t for $t = 0$ and for all $u \in S_d^+$. $\square$

By the regularity of X, we are now allowed to differentiate the equations (3.1) and (3.2) with respect to t and evaluate them at $t = 0$. As a consequence, ϕ and ψ satisfy the system of differential equations

$$\begin{aligned}
\frac{\partial \phi(t,u)}{\partial t} &= F(\psi(t,u)), \qquad \phi(0,u) = 0,\\
\frac{\partial \psi(t,u)}{\partial t} &= R(\psi(t,u)), \qquad \psi(0,u) = u \in S_d^+,
\end{aligned}$$

where F and R are defined as in (2.2). The analysis of these (generalized Riccati) differential equations is subject of Section 5.1, whereas the specific form of F and R is elaborated in the following.

4. Necessary parameter restrictions. In this section, we derive *necessary* parametric restrictions, that is, given an affine process on S_d^+, we determine necessary implications on a set of parameters which only ensue from Definition 2.1. These conditions are precisely the conditions on the admissible parameter set as of Definition 2.3. The form of the functions F and R as defined by (2.2) is then characterized by means of this parameter set, which is stated in Proposition 4.9 below. For its proof, we first provide a number of technical prerequisites.

LEMMA 4.1. *Let $x, u \in S_d^+$ and*

$$x = O \Lambda O^\top = O \operatorname{diag}(\lambda_1 > 0, \dots, \lambda_{d-r} > 0, 0, \dots, 0) O^\top \tag{4.1}$$

be the diagonalization of x with $r \geq 0$ and $O \in O(d)$. Then the following assertions are equivalent:

(i) $ux = xu = 0$,
(ii) $\langle x, u \rangle = 0$,
(iii) *u is of form*

$$u = O \begin{pmatrix} 0 & 0 \\ 0 & w \end{pmatrix} O^\top \tag{4.2}$$

with $w \in S_r^+$.

PROOF. The direction (i) $\Rightarrow$ (ii) is obvious. In order to prove the implication (ii) $\Rightarrow$ (iii), define v as $v = O^\top u O$. Then we have

$$0 = \langle x, u \rangle = \langle \Lambda, O^\top u O \rangle = \sum_{i \leq d-r} \lambda_i v_{ii},$$

which implies $v_{ii} = 0$ for all $i \leq d - r$ and by the positive definiteness of v it must then be of form

$$v = \begin{pmatrix} 0 & 0 \\ 0 & w \end{pmatrix}$$

with $w \in S_r^+$. Thus, u is given by (4.2). This then implies that $ux = xu = 0$, which proves the direction (iii) $\Rightarrow$ (i). $\square$

LEMMA 4.2. *Let p be an orthogonal projector, that is, $p \in S_d^+$ and $p^2 = p$* (*see*, *e.g.*, *Kato* [33], *Section* I.6.7), *and define $q = I_d - p$. Then q is an orthogonal projector and the orthogonal complement of p in S_d^+ equals*

$$\{v \in S_d^+ \mid \langle p, v \rangle = 0\} = \{quq \mid u \in S_d^+\}.$$

PROOF. That q is an orthogonal projector follows by inspection. The diagonalization of p is of the form $p = O \Lambda O^\top$ with $\Lambda = \operatorname{diag}(1, \dots, 1, 0, \dots, 0)$, and thus $q = O(I_d - \Lambda) O^\top$. In view of Lemma 4.1, we conclude that $v \in S_d^+$ is orthogonal to p if and only if $v = qvq$. This proves the assertion. $\square$

LEMMA 4.3. *Let u be in S_d and $x \in \partial S_d^+$ such that $ux = xu = 0$. Then, the linear map T_u defined by*

$$T_u : S_d \to S_d, \qquad v \mapsto T_u v := uvu$$

has the following properties:

(i) *T_u is self-adjoint and* $T_u(S_d^+ - \mathbb{R}_+ x) \subseteq S_d^+$.
(ii) *There exists an element* $v \in S_d$ *such that* $T_u v = u$.

PROOF. The assertion (i) is obvious, since for every $k \in \mathbb{R}_+$, $T_u kx = kuxu = 0$ and $T_u v = uvu \in S_d^+$ if $v \in S_d^+$. For proving part (ii), we use the fact that x is of form (4.1) and that all zero divisors u in S_d of x can be represented by (4.2) with $w \in S_r$. Thus, setting

$$v = O \begin{pmatrix} 0 & 0 \\ 0 & w^+ \end{pmatrix} O^\top,$$

where w^+ satisfies $ww^+w = w$, yields $T_u v = u$. □

LEMMA 4.4. *Let V denote a vector space*[8] *over $\mathbb{R}$. Let $L : S_d^+ \to V$ be an additive (resp., homogeneous additive) map, that is, for all $x, y \in S_d^+$ and $\lambda = 1$ (resp., for all $\lambda \in \mathbb{R}_+$) we have*

$$L(x + \lambda y) = L(x) + \lambda L(y).$$

Then $L(x)$ is the restriction of an additive (resp., $\mathbb{R}$-linear) map on S_d.

PROOF. We define the map $\widetilde{L} : S_d \to V$ as

$$\widetilde{L}(x - y) := L(x) - L(y), \qquad x, y \in S_d^+.$$

$\widetilde{L}$ is well defined, as for $u, v, x, y \in S_d^+$ such that $u - v = x - y$ we have

$$L(u) - L(v) = \widetilde{L}(u - v) = \widetilde{L}(x - y) = L(x) - L(y).$$

Since $S_d^+ - S_d^+ = S_d$, the domain of $\widetilde{L}$ is all of S_d. Also, $L(0) = 0$ by the additivity of L. Hence, L is the restriction of $\widetilde{L}$ to S_d^+. Homogeneity of $\widetilde{L}$ holds, as for $\lambda > 0$, $z = x - y \in S_d$ we have by definition

$$\widetilde{L}(\lambda z) = L(\lambda x) - L(\lambda y) = \lambda L(x) - \lambda L(y) = \lambda \widetilde{L}(z).$$

Finally, we show additivity of $\widetilde{L}$. Choose $w, z \in S_d$ such that $z = x - y$, $w = u - v$, hence $w + z = (x + u) - (y + v)$. By the definition of $\widetilde{L}$, we have

$$\widetilde{L}(z) = L(x) - L(y), \qquad \widetilde{L}(w) = L(u) - L(v),$$

and by the additivity of L we obtain

$$\begin{aligned} \widetilde{L}(w + z) &= L(x + u) - L(y + v) = L(x) + L(u) - L(y) - L(v) \\ &= \widetilde{L}(z) + \widetilde{L}(w). \end{aligned}$$ □

We now provide a convergence result for Laplace transforms (in fact Laplace–Fourier transforms), which is most relevant for the analysis of affine processes.

[8] In the proof of Proposition 4.9 below, V corresponds to S_d, the vector space of linear maps $S_d \to S_d$, or the vector space of finite signed measures on S_d.

LEMMA 4.5. *Let ν_n be a sequence of measures on S_d with*

$$L_n(u) = \int_{S_d} e^{-\langle u,\xi\rangle} \nu_n(d\xi) < \infty \quad \text{and} \quad \lim_{n\to\infty} L_n(u) = L(u), \qquad u \in S_d^+,$$

pointwise, for some finite function L on S_d^+ continuous at $u = 0$. Then ν_n converges weakly to some finite measure ν on S_d and the Fourier–Laplace transforms converge for $u \in S_d^{++} \cup \{0\}$ and $v \in S_d$ to the Fourier–Laplace transforms of ν, that is,

$$\lim_{n\to\infty} \int_{S_d} e^{-\langle u+iv,\xi\rangle} \nu_n(d\xi) = \int_{S_d} e^{-\langle u+iv,\xi\rangle} \nu(d\xi).$$

In particular, $\nu(S_d) = \lim_{n\to\infty} \nu_n(S_d)$ and

$$L(u) = \int_{S_d} e^{-\langle u,\xi\rangle} \nu(d\xi),$$

for all $u \in S_d^{++} \cup \{0\}$.

REMARK 4.6. Instead of $u = 0$ we could take any set K of points at the boundary $K \subset \partial S_d^+$: if we assume continuity of L at points in K, then we obtain the equality of L with the Laplace transform of ν for all points in K. Additionally, continuity is too strong an assumption, since we only need right continuity of L along the segment $u + \varepsilon I_d$ for $\varepsilon = 0$ at the points from the boundary under consideration.

PROOF OF LEMMA 4.5. Since $\nu_n(S_d) = L_n(0)$ is bounded, we know by general theory that ν_n has a vague accumulation point ν, which is a finite measure on S_d.

Since $L_n(u) < \infty$ on S_d^+, it follows by well-known regularity properties of Laplace transforms (see, e.g., [19], Lemma 10.8) that the functions L_n admit an analytic extension on the strip $S_d^{++} + iS_d$, still denoted by L_n:

$$(u + iv) \mapsto L_n(u + iv) = \int_{S_d} e^{-\langle u+iv,\xi\rangle} \nu_n(d\xi).$$

Moreover, pointwise convergence of the finite convex functions L_n to L on S_d^+ implies that this convergence is in fact uniform on compact subsets of S_d^{++} (see, e.g., Rockafellar [44], Theorem 10.8). Hence, the functions L_n are uniformly bounded on compact subsets of S_d^{++} and since $|L_n(u + iv)| \le L_n(u)$, also on compact subsets of $S_d^{++} + iS_d$. Therefore, and since S_d^{++} is a set of uniqueness in $S_d^{++} + iS_d$, it follows by Vitali's theorem ([41], Chapter 1, Proposition 7) that the analytic functions L_n converge uniformly on compact subsets of $S_d^{++} + iS_d$ to an analytic limit thereon. By Lévy's continuity theorem, we therefore know that for any $u \in S_d^{++}$ the finite measures $\exp(-\langle u, \xi\rangle)\nu_n(d\xi)$ converge weakly to a limit, which by uniqueness of the weak limit has to equal $\exp(-\langle u, \xi\rangle)\nu(d\xi)$. Whence the only vague accumulation point of ν_n is ν. Vague convergence implies weak convergence if mass is conserved. Continuity of $L(u)$ at $u = 0$ implies this mass conservation: indeed, by weak convergence of $e^{-\langle \varepsilon I_d, \xi \rangle} \nu_n$ we arrive at

$$\begin{aligned} L(\varepsilon I_d) &= \lim_{n \to \infty} \int_{S_d} e^{-\langle \varepsilon I_d, \xi \rangle} \nu_n(d\xi) = \int_{S_d} e^{-\langle \varepsilon I_d, \xi \rangle} \nu(d\xi) \\ &= \int_{S_d} e^{-\langle \varepsilon I_d, \xi \rangle} 1_{\{\langle I_d, \xi \rangle \le 0\}} \nu(d\xi) + \int_{S_d} e^{-\langle \varepsilon I_d, \xi \rangle} 1_{\{\langle I_d, \xi \rangle > 0\}} \nu(d\xi) \end{aligned}$$

and therefore—by dominated convergence—we obtain that the limit $\varepsilon \to 0$ yields

$$L(0) = \int_{S_d} \nu(d\xi),$$

which is the desired mass conservation, hence weak convergence, which means in turn convergence of the Fourier–Laplace transform at $u = 0$. $\square$

Finally, let us state a general comparison result for ODEs and hereto introduce the notion of *quasi-monotonicity*, which we shall need several times throughout this article, in particular in the proofs of Propositions 4.9 and 5.3 below.

DEFINITION 4.7. Let $U \subset S_d$ be an open set. A function $f : U \to S_d$ is called *quasi-monotone increasing* if for all elements $x, y \in U$, $u \in S_d^+$ which satisfy $x \preceq y$ and $\langle x, u \rangle = \langle y, u \rangle$,

$$\langle f(x), u \rangle \le \langle f(y), u \rangle$$

holds true. Accordingly, we call f *quasi-constant* if both f and $-f$ are quasi-monotone increasing.

The following comparison result can be deduced from a more general theorem proved by Volkmann [52].

THEOREM 4.8. *Let $U \subset S_d$ be an open set. Let $f : [0, T) \times U \to S_d$ be a continuous locally Lipschitz map such that $f(t, \cdot)$ is quasi-monotone increasing on U for all $t \in [0, T)$. Let $0 < t_0 \le T$ and $x, y : [0, t_0) \to U$ be differentiable maps such that $x(0) \preceq y(0)$ and*

$$\dot{x}(t) - f(t, x(t)) \preceq \dot{y}(t) - f(t, y(t)), \qquad 0 \le t < t_0.$$

Then we have $x(t) \preceq y(t)$ for all $t \in [0, t_0)$.

4.1. *The functions F and R.* The main result of this section characterizes the form of the functions F and R as defined by (2.2).

PROPOSITION 4.9. *Let X be an affine process with state space S_d^+. Then there exist parameters $(\alpha, b, \beta^{ij}, c, \gamma, m, \mu)$, where $\alpha, \beta^{ij}, c, \gamma, m, \mu$ satisfy the admissibility conditions of Definition 2.3 and $b \in S_d^+$, such that the functions F and R are of the form (2.16) and (2.17).*

REMARK 4.10. Note that for the moment we only obtain $b \in S_d^+$, and not (2.4).

PROOF OF PROPOSITION 4.9. As the proof of Proposition 4.9 is rather long, we divide it into several steps:

Step 1. *Necessary admissibility conditions for b, c, γ, m.* In order to derive the particular form of F and R with the above parameter restrictions, we follow the approach of Keller-Ressel [34], Theorem 2.6. Note that the t-derivative of $P_t e^{-\langle u,x\rangle}$ at $t = 0$ exists for all $x, u \in S_d^+$, since

$$\begin{aligned}\lim_{t\to 0^+} \frac{P_t e^{-\langle u,x\rangle} - e^{-\langle u,x\rangle}}{t} &= \lim_{t\to 0^+} \frac{e^{-\phi(t,u)-\langle\psi(t,u),x\rangle} - e^{-\langle u,x\rangle}}{t} \\ &= \bigl(-F(u) - \langle R(u), x\rangle\bigr)e^{-\langle u,x\rangle}\end{aligned} \tag{4.3}$$

is well defined by Proposition 3.4. Moreover, we can also write

$$\begin{aligned}&-F(u) - \langle R(u), x\rangle \\ &\quad= \lim_{t\to 0^+} \frac{P_t e^{-\langle u,x\rangle} - e^{-\langle u,x\rangle}}{t e^{-\langle u,x\rangle}} \\ &\quad= \lim_{t\to 0^+} \frac{1}{t}\biggl(\int_{S_d^+\setminus\{0\}} e^{-\langle u,\xi - x\rangle} p_t(x, d\xi) - 1\biggr) \\ &\quad= \lim_{t\to 0^+}\biggl(\frac{1}{t}\int_{S_d^+ - x} \bigl(e^{-\langle u,\xi\rangle} - 1\bigr) p_t(x, d\xi + x) + \frac{p_t(x, S_d^+) - 1}{t}\biggr).\end{aligned}$$

By the above equalities and the fact that $p_t(x, S_d^+) \le 1$, we then obtain for $u = 0$

$$0 \ge \lim_{t\to 0^+} \frac{p_t(x, S_d^+) - 1}{t} = -F(0) - \langle R(0), x\rangle.$$

Setting $F(0) = c$ and $R(0) = \gamma$ yields $c \in \mathbb{R}^+$ as in (2.5) and $\gamma \in S_d^+$ as in (2.6). We thus obtain

$$\begin{aligned}&-\bigl(F(u) - c\bigr) - \langle R(u) - \gamma, x\rangle \\ &\qquad= \lim_{t\to 0^+} \frac{1}{t}\int_{S_d^+ - x} \bigl(e^{-\langle u,\xi\rangle} - 1\bigr) p_t(x, d\xi + x).\end{aligned} \tag{4.4}$$

For every fixed $t > 0$, the right-hand side of (4.4) is the logarithm of the Laplace transform of a compound Poisson distribution supported on $S_d^+ - \mathbb{R}_+ x$ with intensity $p_t(x, S_d^+)/t$ and compounding distribution $p_t(x, d\xi + x)/p_t(x, S_d^+)$. Concerning the support, note that the compounding distribution is concentrated on $S_d^+ - x$, which implies that the compound Poisson distribution has support on the convex cone $S_d^+ - \mathbb{R}_+ x$. By Lemma 4.5, the pointwise convergence of (4.4) for $t \to 0$ to some function being continuous at 0, implies weak convergence of the compound Poisson distributions to some infinitely divisible probability distribution $K(x, dy)$ supported on $S_d^+ - \mathbb{R}_+ x$. Indeed, this follows from the fact that any compound Poisson distribution is infinitely divisible and the class of infinitely divisible distributions is closed under weak convergence ([47], Lemma 7.8). Again, by Lemma 4.5 the Laplace transform of $K(x, dy)$ is then given as exponential of the left-hand side of (4.4).

In particular, for $x = 0$, $K(0, dy)$ is an infinitely divisible distribution with support on the cone S_d^+. By the Lévy–Khintchine formula on proper cones (see [49], Theorem 3.21), its Laplace transform is therefore of the form

$$\exp\biggl(-\langle b, u\rangle + \int_{S_d^+ \setminus \{0\}} (e^{-\langle u, \xi\rangle} - 1) m(d\xi)\biggr),$$

where $b \in S_d^+$ and m is a Borel measure supported on S_d^+ such that

$$\int_{S_d^+ \setminus \{0\}} (\|\xi\| \wedge 1) m(d\xi) < \infty,$$

yielding (2.7). Therefore,

$$F(u) = \langle b, u\rangle + c - \int_{S_d^+ \setminus \{0\}} \bigl(e^{-\langle u, \xi\rangle} - 1\bigr) m(d\xi).$$

Step 2. *Necessary admissibility conditions for* β^{ij}, μ. We next obtain the particular form of R. Observe that for each $x \in S_d^+$ and $k \in \mathbb{N}$,

$$\exp\bigl(-\bigl(F(u) - c\bigr)/k - \langle R(u) - \gamma, x\rangle\bigr)$$

is the Laplace transform of the infinitely divisible distribution $K(kx, dy)^{*1/k}$, where $*\frac{1}{k}$ denotes the $\frac{1}{k}$ convolution power. For $k \to \infty$, these Laplace transforms obviously converge to $\exp(-\langle R(u) - \gamma, x\rangle)$ pointwise in u. Using again the same arguments as before [an application of Lemma 4.5 as below (4.4)], we can deduce that $K(kx, dy)^{*1/k}$ converges weakly to some infinitely divisible distribution $L(x, dy)$ on $S_d^+ - \mathbb{R}_+ x$ with Laplace transform $\exp(-\langle R(u) - \gamma, x\rangle)$ for $u \in S_d^+$.

By the Lévy–Khintchine formula on S_d ([47], Theorem 8.1, indeed on $\mathbb{R}^{(d(d+1)/2)}$ by modifying the scalar product appropriately), the characteristic function of $L(x, dy)$ has the form

$$\begin{aligned}\widehat{L}(x, u) = \exp\biggl(&\frac{1}{2}\langle u, A(x)u\rangle + \langle B(x), u\rangle \\ &+ \int_{S_d \setminus \{0\}} \bigl(e^{-\langle u, \xi\rangle} - 1 - \langle \chi(\xi), u\rangle\bigr) M(x, d\xi)\biggr),\end{aligned}$$

for $u \in i S_d$, where $A(x)$ is a symmetric positive semidefinite linear operator on S_d, $B(x) \in S_d$, $M(x, \cdot)$ a measure on $S_d \setminus \{0\}$ satisfying

$$\int_{S_d \setminus \{0\}} (\|\xi\|^2 \wedge 1) M(x, d\xi) < \infty,$$

and χ some appropriate truncation function. Furthermore, by [47], Theorem 8.7,

$$\begin{aligned}
&\int_{S_d \setminus \{0\}} f(\xi) \frac{1}{t} p_t(x, d\xi + x) \\
&\qquad \xrightarrow{t \to 0} \int_{S_d \setminus \{0\}} f(\xi) m(d\xi) + \int_{S_d \setminus \{0\}} f(\xi) M(x, d\xi)
\end{aligned} \tag{4.5}$$

holds true for all $f : S_d \to \mathbb{R}$ which are bounded, continuous and vanishing on a neighborhood of 0. We conclude that $M(x, d\xi)$ has support in $S_d^+ - x$. Therefore, the characteristic function $\widehat{L}(x, u)$ admits an analytic extension to $S_d^+ \times i S_d$, which then has to coincide with the Laplace transform for $u \in S_d^+$. We conclude that, for all $x \in S_d^+$,

$$\begin{aligned}
&-\langle R(u) - \gamma, x \rangle \\
&\quad = \frac{1}{2} \langle u, A(x) u \rangle - \langle B(x), u \rangle \\
&\qquad + \int_{S_d \setminus \{0\}} \bigl(e^{-\langle u, \xi \rangle} - 1 + \langle \chi(\xi), u \rangle\bigr) M(x, d\xi), \qquad u \in S_d^+.
\end{aligned} \tag{4.6}$$

As the left-hand side of (4.6) is linear in the components of x, it follows that $x \mapsto A(x)$, $x \mapsto B(x)$ as well as $x \mapsto \int_E (\|\xi\|^2 \wedge 1) M(x, d\xi)$ for every $E \in \mathcal{B}(S_d \setminus \{0\})$ are homogeneous additive maps on S_d^+ in the sense of Lemma 4.4. This then implies that they are restrictions of linear maps on S_d, such that we can write

$$A(x) = \sum_{i,j} a_{ij} x_{ij}, \qquad B(x) = \sum_{i,j} \beta^{ij} x_{ij},$$

$$\int_E (\|\xi\|^2 \wedge 1) M(x, d\xi) = \langle x, \mu(E) \rangle = \sum_{i,j} \mu_{ij}(E) x_{ij},$$

where (recall that c^{ij} denotes the standard basis of S_d defined in Section 1.2):

$$\begin{aligned}
a_{ij} &= a_{ji} = (1 + \delta_{ij}) \frac{A(c^{ij})}{2} : S_d \to S_d \qquad \text{linear}, \\
\beta^{ij} &= \beta^{ji} = (1 + \delta_{ij}) \frac{B(c^{ij})}{2} \in S_d
\end{aligned}$$

and

$$E \mapsto \mu_{ij}(E) = \mu_{ji}(E) = (1 + \delta_{ij}) \frac{\int_E (\|\xi\|^2 \wedge 1) M(c^{ij}, d\xi)}{2}$$

are finite signed measures on $S_d \setminus \{0\}$. The fact that $M(x, \cdot)$ is a nonnegative measure for each $x \in S_d^+$ implies immediately that $\mu(E)$ is a positive semidefinite matrix.

In (4.5), take now $x = \frac{1}{n}e^{ij}$ and nonnegative functions $f = f_n \in C_b(S_d)$ with $f_n = 0$ on $S_d^+ - \frac{1}{n}e^{ij}$. Then for each n the left-hand side of (4.5) is zero since the $p_t(\frac{1}{n}e^{ij}, d\xi + \frac{1}{n}e^{ij})$ is concentrated on $S_d^+ - \frac{1}{n}e^{ij}$. As $\operatorname{supp}(m) \subseteq S_d^+$, the first integral on the right vanishes as well. Hence,

$$\begin{aligned} 0 &= \int_{S_d \setminus \{0\}} f_n(\xi) M\left(\frac{1}{n}e^{ij}, d\xi\right) = \int_{S_d \setminus \{0\}} \frac{f_n(\xi)}{\|\xi\|^2 \wedge 1} \left\langle \frac{1}{n}e^{ij}, \mu(d\xi)\right\rangle \\ &= \frac{1}{n} \int_{S_d \setminus \{0\}} \frac{f_n(\xi)}{\|\xi\|^2 \wedge 1} \big(\mu_{ii}(d\xi) + (1-\delta_{ij})\big(\mu_{jj}(d\xi) + 2\mu_{ij}(d\xi)\big)\big) \end{aligned}$$

for any nonnegative function $f_n \in C_b(S_d)$ with $f_n = 0$ on $S_d^+ - \frac{1}{n}e^{ij}$ implies that $\operatorname{supp}(\mu_{ij}) \subseteq S_d^+ - \frac{1}{n}e^{ij}$ for each n. Thus, we can conclude that $\operatorname{supp} \mu_{ij} \subseteq S_d^+$ for all $1 \leq i, j \leq d$.

Now let $T : S_d \to S_d$ be any linear map with the property $T(S_d^+ - \mathbb{R}_+ x) \subseteq S_d^+$. Then $T(\operatorname{supp}(L(x, dy))) \subseteq S_d^+$. This implies that the pushforward $T_* L(x, \cdot)$ of $L(x, dy)$ under T is an infinitely divisible distribution supported on S_d^+. By the Lévy–Khintchine formula on proper cones (see [49], Theorem 3.21, and by [47], Proposition 11.10) this implies that for all $x \in S_d^+$

$$TA(x)T^\top = 0, \tag{4.7}$$

$$TB(x) + \int_{S_d^+ \setminus \{0\}} \big(\widetilde{\chi}(T\xi) - T(\chi(\xi))\big) M(x, d\xi) \in S_d^+, \tag{4.8}$$

$$\int_{S_d^+ \setminus \{0\}} (\|\xi\| \wedge 1) T_* M(x, d\xi) < \infty, \tag{4.9}$$

where $\widetilde{\chi}$ denotes some truncation function associated with $T_* L(x, \cdot)$ and $T_* M$ the pushforward of M under T. Due to (4.9), we can set $\widetilde{\chi} = 0$. Thus, (4.8) becomes

$$TB(x) - \int T(\chi(\xi)) M(x, d\xi) \in S_d^+. \tag{4.10}$$

Moreover, equations (4.7), (4.10) and (4.9) are equivalent to

$$\begin{aligned} \langle T^\top v, A(x) T^\top v \rangle &= 0 \qquad \text{for all } v \in S_d, \\ \langle B(x), T^\top v \rangle - \int_{S_d^+ \setminus \{0\}} \langle (\chi(\xi)), T^\top v \rangle M(x, d\xi) &\geq 0 \qquad \text{for all } v \in S_d^+, \end{aligned}$$

$$\int_{S_d^+ \setminus \{0\}} (\|T\xi\| \wedge 1) M(x, d\xi) < \infty. \tag{4.11}$$

In particular, we claim that

$$\begin{aligned} &\langle u, A(x) u \rangle = 0 \\ &\qquad \text{for all } u \in S_d \text{ s.t. } ux = xu = 0, \end{aligned} \tag{4.12}$$

$$\langle B(x), u\rangle - \int_{S_d^+\setminus\{0\}} \langle \chi(\xi), u\rangle M(x, d\xi) \geq 0 \quad \text{for all } u \in S_d^+ \text{ s.t. } ux = xu = 0, \tag{4.13}$$

$$\int_{S_d^+\setminus\{0\}} \langle \chi(\xi), u\rangle M(x, d\xi) < \infty \quad \text{for all } u \in S_d^+ \text{ s.t. } ux = xu = 0. \tag{4.14}$$

Indeed, if x is invertible then $ux = 0$ is equivalent to $u = 0$ and the assertions are obvious. Otherwise, if x is in ∂S_d^+, the linear map T_u defined in Lemma 4.3 is self-adjoint and satisfies $T_u(S_d^+ - \mathbb{R}_+ x) \subseteq S_d^+$. Furthermore, by Lemma 4.3(ii), there exists an element $v \in S_d$ such that $T_u v = u$. Hence,

$$\langle u, A(x)u\rangle = \langle T_u^\top v, A(x) T_u^\top v\rangle = 0.$$

It follows from the proof of Lemma 4.3 that for $u \in S_d^+$, v is an element of S_d^+ as well and we have $\langle B(x), u\rangle = \langle B(x), T_u^\top v\rangle$ and $\langle \chi(\xi), u\rangle = \langle (\chi(\xi)), T^\top v\rangle$. Equation (4.14) is obtained by choosing $T = T_{\sqrt{u}}$ in (4.11). Indeed,

$$\begin{aligned}
\int_{S_d^+\cap\{\|\xi\|\leq 1\}} \langle \xi, u\rangle M(x, d\xi) &= \int_{S_d^+\cap\{\|\xi\|\leq 1\}} \langle I_d, \xi u\rangle M(x, d\xi) \\
&\leq \|I_d\| \int_{S_d^+\cap\{\|\xi\|\leq 1\}} \|\xi u\| M(x, d\xi) \\
&= \|I_d\| \int_{S_d^+\cap\{\|\xi\|\leq 1\}} \|T_{\sqrt{u}}\xi\| M(x, d\xi) < \infty.
\end{aligned}$$

From these arguments and Lemma 4.1, properties (2.11) and (2.9) can be derived so far. Thus, only (2.3) remains to be shown.

Step 3. *Necessary admissibility condition for* α. Due to the linearity of $A(x)$, $\langle u, A(x)u\rangle$ can be written as $4\langle x, \vartheta(u)\rangle$, where the (ij)th component of $\vartheta(u) \in S_d$ is defined by $\vartheta_{ij}(u) = 1/4\langle u, a_{ij} u\rangle$. Note that ϑ is defined on all of S_d. Given that for all $x \in S_d^+$, $A(x)$ is a positive semidefinite operator on S_d, $\langle u, A(x)u\rangle \geq 0$ for all $u \in S_d$ and therefore, by the self duality of S_d^+, $\vartheta(u) \in S_d^+$. By (4.12), we have for all u such that $ux = xu = 0$

$$0 = \langle u, A(x)u\rangle = 4\langle x, \vartheta(u)\rangle. \tag{4.15}$$

Next, we show that ϑ is quasi-constant, that is, $\langle x, \vartheta(u + w) - \vartheta(u)\rangle = 0$ for all $x, u, w \in S_d^+$ with $\langle x, w\rangle = 0$ (see Definition 4.7). Indeed, pick $x, u, w \in S_d^+$ with $\langle x, w\rangle = 0$. According to our assumptions, $A(x)w = 0$, due to (4.15) and the positivity of A. Hence,

$$\begin{aligned}
4\langle x, \vartheta(u + w) - \vartheta(u)\rangle &= \langle u + w, A(x)(u + w)\rangle - \langle u, A(x)u\rangle \\
&= \langle u, A(x)w\rangle + \langle A(x)w, u\rangle = 0,
\end{aligned}$$

where the second last equality holds in view of the symmetry of $A(x)$.

We now claim that there exists some $\alpha \in S_d^+$ such that $\vartheta(u) = u\alpha u$, for each $u \in S_d$. It is sufficient to show that this statements holds for all orthogonal projectors $p \in S_d^+$, that is, there exists some $\alpha \in S_d^+$ such that $\vartheta(p) = p\alpha p$ for all orthogonal projectors p. Indeed, if this is the case, we can derive the general statement in the following way: take $u \in S_d^+$, then—by spectral decomposition—there are numbers $\lambda_i \geq 0$ and orthogonal projectors p_i, which are mutually orthogonal, such that $u = \sum_{i=1}^d \lambda_i p_i$ (see, e.g., Kato [33], Section I.6.9). Since the assertion holds for all orthogonal projectors, we have that

$$2\vartheta(u) = \sum_{i,j=1}^d \lambda_i \lambda_j \big(\vartheta(p_i + p_j) - \vartheta(p_i) - \vartheta(p_j)\big)$$

by the property that ϑ is quadratic. Since $p_i + p_j$ is again an orthogonal projector, we obtain the result.

We prove the assertion on orthogonal projectors by quasi-constancy. Take an arbitrary orthogonal projector p and define $q = I_d - p$. Additionally, we define $\alpha = \vartheta(I_d)$. By quasi-constancy, we obtain

$$\langle x, \vartheta(p+q) - \vartheta(q)\rangle = \langle y, \vartheta(p+q) - \vartheta(p)\rangle = 0$$

and

$$\langle x, \vartheta(p)\rangle = \langle y, \vartheta(q)\rangle = 0,$$

for all $x, y \in S_d^+$ with $\langle x, p\rangle = 0$ and $\langle y, q\rangle = 0$. Therefore, $\alpha - \vartheta(q)$ and $\vartheta(p)$ are orthogonal to the orthogonal complement of p in S_d^+ (i.e., the positive symmetric matrices of the form quq by Lemma 4.2), and $\alpha - \vartheta(p)$ and $\vartheta(q)$ are orthogonal to the orthogonal complement of q in S_d^+ (the positive symmetric matrices of the form pup by Lemma 4.2). This means that we can write

$$\alpha = \vartheta(p) + \vartheta(q) + \beta,$$

where the symmetric matrix β is orthogonal to all elements which are orthogonal to p and q (in S_d^+), that is, β is orthogonal to the linear span of matrices of the form pup and quq. However, such a decomposition is unique, since all vectors in the sum are mutually orthogonal, and the decomposition is given by

$$\alpha = (p+q)\alpha(p+q) = p\alpha p + q\alpha q + (p\alpha q + q\alpha p).$$

Therefore, we can conclude the assertion $\vartheta(p) = p\alpha p$. Since p was arbitrary the assertion is proved.

Finally, all the derived restrictions on the parameters together with (4.6) then yield (2.17). $\square$

REMARK 4.11. An alternative proof for the special form of the diffusion matrix $A(x)$ can also be established by Stokes' theorem [50] on Riccati ODEs.

4.2. *Infinitesimal generator.* The aim of this section is to prove the form of the infinitesimal generator as stated in (2.12).

PROPOSITION 4.12. *The infinitesimal generator $\mathcal{A}$ of an affine process on S_d^+ satisfies $\mathcal{S}_+ \subset \mathcal{D}(\mathcal{A})$ and is of the form* (2.12) *for all $f \in \mathcal{S}_+$ and $x \in S_d^+$.*

PROOF. As already mentioned in the proof of Proposition 4.9, the t-derivative of $P_t e^{-\langle u,x\rangle}$ at $t=0$ exists pointwise for all $x, u \in S_d^+$ and is given by (4.3). Furthermore, $x \mapsto (-F(u) - \langle R(u), x\rangle)e^{-\langle u,x\rangle} \in C_0(S_d^+)$, for $u \in S_d^{++}$. As (P_t) is a Feller semigroup on $C_0(S_d^+)$, it follows from [47], Lemma 31.7, that $\{e^{-\langle u,x\rangle} \mid u \in S_d^{++}\} \in D(\mathcal{A})$ and

$$\mathcal{A}e^{-\langle u,x\rangle} = \bigl(-F(u) - \langle R(u), x\rangle\bigr)e^{-\langle u,x\rangle}.$$

Combined with Proposition 4.9, we thus obtain

$$\begin{aligned}
\mathcal{A}e^{-\langle u,x\rangle} &= \biggl(-\langle b, u\rangle - c + \int_{S_d^+\setminus\{0\}} (e^{-\langle u,\xi\rangle} - 1) m(d\xi) \\
&\qquad + \biggl\langle 2u\alpha u - B^\top(u) - \gamma \\
&\qquad\qquad + \int_{S_d^+\setminus\{0\}} \biggl(\frac{e^{-\langle u,\xi\rangle} - 1 + \langle \chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1}\biggr)\mu(d\xi), x\biggr\rangle\biggr) e^{-\langle u,x\rangle} \\
&= \frac{1}{2}\sum_{i,j,k,l} A_{ijkl}(x) u_{ij} u_{kl} e^{-\langle u,x\rangle} + \bigl\langle b + B(x), \nabla e^{-\langle u,x\rangle}\bigr\rangle \\
&\quad - (c + \langle \gamma, x\rangle) e^{-\langle u,x\rangle} \\
&\quad + \int_{S_d^+\setminus\{0\}} (e^{-\langle u,x+\xi\rangle} - e^{-\langle u,x\rangle}) m(d\xi) \\
&\quad + \int_{S_d^+\setminus\{0\}} \bigl(e^{-\langle u,x+\xi\rangle} - e^{-\langle u,x\rangle} + \bigl\langle \chi(\xi), \nabla e^{-\langle u,x\rangle}\bigr\rangle\bigr) M(x, d\xi).
\end{aligned} \tag{4.16}$$

Indeed, in order to obtain the form of the diffusion part, observe that we have by symmetrization

$$\begin{aligned}
2\langle u\alpha u, x\rangle &= 2\sum_{i,j,k,l} \alpha_{jk} x_{il} u_{ij} u_{kl} \\
&= \frac{1}{2}\sum_{i,j,k,l} (x_{ik}\alpha_{jl} + x_{il}\alpha_{jk} + x_{jk}\alpha_{il} + x_{jl}\alpha_{ik}) u_{ij} u_{kl} \\
&= \frac{1}{2}\sum_{i,j,k,l} A_{ijkl}(x) u_{ij} u_{kl};
\end{aligned}$$

see (2.13).

According to Theorem B.3, the linear hull $\mathcal{M}$ of $\{e^{-\langle u,\cdot\rangle} \mid u \in S_d^{++}\}$ is dense in $\mathcal{S}_+$ with respect to the family of seminorms $p_{k,+}$ defined in (B.2). Denoting the right-hand side of (2.12) by $\mathcal{A}^\sharp$, we now claim that for every $f \in \mathcal{S}_+$

$$\lim_{n\to\infty} \|\mathcal{A}^\sharp f_n - \mathcal{A}^\sharp f\|_\infty = 0, \tag{4.17}$$

where $f_n \in \mathcal{M}$ such that $\lim_{n\to\infty} p_{k,+}(f - f_n) = 0$ for every k. Indeed, this is obvious for the differential operator part of $\mathcal{A}^\sharp$. By choosing $\chi(\xi) = 1_{\{\|\xi\|\le 1\}}\xi$ and by denoting $g(x) := f_n(x) - f(x)$, we obtain the following estimate for the integral part:

$$\begin{aligned}
&\left\| \int_{S_d^+\setminus\{0\}} \left(\frac{g(x+\xi) - g(x) - \langle 1_{\{\|\xi\|\le 1\}}\xi, \nabla g(x)\rangle}{\|\xi\|^2 \wedge 1} \right) x_{ij}\mu_{ij}(d\xi) \right\|_\infty \\
&\quad \le \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|\le 1\}} \left\| \left(\sum_{k,l,m,n} \left(\int_0^1 \frac{\partial^2 g(x+s\xi)}{\partial x_{kl}\,\partial x_{mn}} (1-s)\,ds \right) \frac{\xi_{kl}\xi_{mn}}{\|\xi\|^2} \right) x_{ij} \right\|_\infty \\
&\qquad\qquad \times \big(\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)\big) \\
&\qquad + \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|>1\}} \big\|\big(g(x+\xi) - g(x)\big)x_{ij}\big\|_\infty \big(\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)\big) \\
&\quad \le C_1 p_{3,+}(g) \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|\le 1\}} \frac{\|\xi\|^2}{\|\xi\|^2} \big(\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)\big) \\
&\qquad + \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|>1\}} \big(\big\|g(x+\xi)(\|x+\xi\|)\big\|_\infty + \|g(x)x_{ij}\|_\infty\big) \\
&\qquad\qquad \times \big(\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)\big) \\
&\quad \le C_2 p_{3,+}(g)\big(\mu_{ij}^+(S_d^+) + \mu_{ij}^-(S_d^+)\big) \le C_3, p_{3,+}(g),
\end{aligned}$$

where C_1, C_2 and C_3 denote some constants and μ_{ij}^+, μ_{ij}^- correspond to the Jordan decomposition $\mu_{ij} = \mu_{ij}^+ - \mu_{ij}^-$. In the second last inequality, we use the estimate $x_{ij} \le \|x + \xi\|$. The same as above can be shown for the measure $m(d\xi)$, whence (4.17) holds true. As by the first part of the proof, we have $\mathcal{A}^\sharp = \mathcal{A}$ for all elements of $\mathcal{M}$, (4.17) implies

$$\lim_{n\to\infty} \|\mathcal{A} f_n - \mathcal{A}^\sharp f\|_\infty = 0.$$

Since the infinitesimal generator of every Feller process is a closed operator, it follows that $\mathcal{S}_+ \subset \mathcal{D}(\mathcal{A})$ and $\mathcal{A} = \mathcal{A}^\sharp$ on $\mathcal{S}_+$. $\square$

4.3. *Linear transformations and canonical representation.* In this subsection, we shall deal with linear transformations of affine processes. The proposition below states how the parameters of an affine process on S_d^+ change under such linear maps, which allows us to establish a canonical representation of an affine process.

PROPOSITION 4.13. *Suppose X is an affine process on S_d^+ with parameters $\alpha, \beta^{ij}, c, \gamma, m, \mu$ as specified in Definition 2.3 and $b \in S_d^+$. Furthermore, let $G : S_d^+ \to S_d^+, x \mapsto gxg^\top$ be an automorphism, where $g \in M_d$ is invertible. Then, $Y := gXg^\top$ is an affine process on S_d^+, whose parameters, denoted by $\widetilde{\cdot}$, are given as follows with respect to the truncation function $\widetilde{\chi} = g\chi(g^{-1}\xi(g^\top)^{-1})g^\top$:*

$$
\begin{aligned}
\widetilde{b} &= gbg^\top, \\
\widetilde{c} &= c, \\
\widetilde{m}(d\xi) &= G_* m(d\xi), \\
\widetilde{\alpha} &= g\alpha g^\top, \\
\widetilde{\gamma} &= (g^\top)^{-1}\gamma g^{-1}, \\
\widetilde{\mu}(d\xi) &= \left(\frac{\|\xi\|^2 \wedge 1}{\|g^{-1}\xi(g^\top)^{-1}\|^2 \wedge 1}\right)(g^\top)^{-1} G_*\mu(d\xi) g^{-1}, \\
\widetilde{B}^\top(u) &= (g^\top)^{-1} B^\top(g^\top u g) g^{-1},
\end{aligned}
$$

where $G_* m$ $(G_*\mu)$ is the pushforward of the measure m (μ, resp.).

PROOF. Let us consider the process

$$
Y_t^y = gX_t^{g^{-1}y(g^\top)^{-1}} g^\top,
$$

for which we have

$$
\begin{aligned}
&\mathbb{E}[\exp(-\langle u, Y_t^y\rangle)] \\
&\quad = \mathbb{E}\big[\exp\big(-\big\langle u, gX_t^{g^{-1}y(g^\top)^{-1}} g^\top\big\rangle\big)\big] \\
&\quad = \mathbb{E}\big[\exp\big(-\big\langle g^\top u g, X_t^{g^{-1}y(g^\top)^{-1}}\big\rangle\big)\big] \\
&\quad = \exp\big(-\phi(t, g^\top u g) - \langle \psi(t, g^\top u g), g^{-1}y(g^\top)^{-1}\rangle\big) \\
&\quad = \exp\big(-\phi(t, g^\top u g) - \langle (g^\top)^{-1}\psi(t, g^\top u g) g^{-1}, y\rangle\big).
\end{aligned}
$$

Define now $\widetilde{\phi}$ and $\widetilde{\psi}$ by

$$
\widetilde{\phi}(t,u) = \phi(t, g^\top u g) \quad \text{and} \quad \widetilde{\psi}(t,u) = (g^\top)^{-1}\psi(t, g^\top u g) g^{-1},
$$

to see that Y is an affine process on S_d^+. Using (2.14) and (2.16), we consequently obtain

$$
\begin{aligned}
\frac{\partial \widetilde{\phi}(t,u)}{\partial t} &= \frac{\partial \phi(t, g^\top u g)}{\partial t} \\
&= F(\psi(t, g^\top u g))
\end{aligned}
$$

$$
\begin{aligned}
&= \langle b, \psi(t, g^\top u g)\rangle + c - \int_{S_d^+\setminus\{0\}} \bigl(e^{-\langle \psi(t,g^\top ug),\xi\rangle} - 1\bigr) m(d\xi) \\
&= \langle gbg^\top, \widetilde{\psi}(t,u)\rangle + c - \int_{S_d^+\setminus\{0\}} \bigl(e^{-\langle \widetilde{\psi}(t,u), g\xi g^\top\rangle} - 1\bigr) m(d\xi) \\
&= \langle \widetilde{b}, \widetilde{\psi}(t,u)\rangle + c - \int_{S_d^+\setminus\{0\}} \bigl(e^{-\langle \widetilde{\psi}(t,u),\xi\rangle} - 1\bigr) G_* m(d\xi).
\end{aligned}
$$

Due to the uniqueness of the Lévy–Khintchine decomposition, this implies that b transforms to $\widetilde{b} = gbg^\top$, c remains constant and m becomes $\widetilde{m}(d\xi) = G_* m(d\xi)$. For $\widetilde{\psi}$ we proceed similarly, that is, we have

$$
\begin{aligned}
\frac{\partial \widetilde{\psi}(t,u)}{\partial t} &= (g^\top)^{-1} \frac{\partial \psi(t, g^\top u g)}{\partial t} g^{-1} = (g^\top)^{-1} R(\psi(t, g^\top ug)) g^{-1} \\
&= (g^\top)^{-1}\biggl(-2\psi(t, g^\top ug)\alpha\psi(t, g^\top ug) + B^\top(\psi(t, g^\top ug)) + \gamma \\
&\qquad - \int_{S_d^+\setminus\{0\}} \biggl(\frac{e^{-\langle \psi(t,g^\top ug),\xi\rangle} - 1 + \langle \chi(\xi), \psi(t, g^\top ug)\rangle}{\|\xi\|^2 \wedge 1}\biggr)\mu(d\xi)\biggr) g^{-1},
\end{aligned}
$$

from which it can be seen that α transforms to $\widetilde{\alpha} = g\alpha g^\top$, γ becomes $\widetilde{\gamma} = (g^\top)^{-1}\gamma g^{-1}$, and μ changes to

$$
\widetilde{\mu}(E) = (g^\top)^{-1}\biggl(\int_E \biggl(\frac{\|\xi\|^2 \wedge 1}{\|g^{-1}\xi(g^\top)^{-1}\|^2 \wedge 1}\biggr) G_*\mu(d\xi)\biggr) g^{-1}
$$

for every $E \in \mathcal{B}(S_d^+ \setminus \{0\})$. Moreover, since $\widetilde{\chi} = g\chi(g^{-1}\xi(g^\top)^{-1})g^\top$

$$
\widetilde{B}^\top(u) = (g^\top)^{-1} B^\top(g^\top ug) g^{-1}. \tag{4.18}
$$

$\square$

By means of Proposition 4.13, we can derive a canonical representation for affine processes.

PROPOSITION 4.14. *Let X be an affine process on S_d^+ with parameters $\alpha, \beta^{ij}, c, \gamma, m, \mu$ as specified in Definition 2.3 and $b \in S_d^+$. Then there exists an automorphism $G : S_d^+ \to S_d^+, x \mapsto gxg^\top$ such that the parameters of the affine process $Y = gXg^\top$, denoted by $\widetilde{\cdot}$, are as in Proposition 4.13 with*

$$
\widetilde{b} = \theta = \operatorname{diag}(\theta_{11}, \dots, \theta_{dd}), \qquad \widetilde{\alpha} = I_r^d,
$$

where we define

$$
I_r^d = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.
$$

PROOF. By Proposition 4.13, the parameters of $Y = gXg^\top$ transform as

$$\widetilde{\alpha} = g\alpha g^\top, \qquad \widetilde{b} = gbg^\top.$$

Since α and $b \in S_d^+$, they are jointly diagonalizable through an automorphism on S_d^+. More precisely, there exists an invertible matrix $g \in M_d$ such that

$$g\alpha g^\top = I_r^d \qquad \text{with } r = \mathrm{rk}(\alpha)$$

and

$$gbg^\top = \mathrm{diag}(\theta_{11}, \ldots, \theta_{dd}) =: \theta,$$

where rk denotes the rank of a matrix. For the proof of this fact, we refer to [22], Theorem 8.7.1. □

4.4. *Condition on the constant drift*. This subsection is devoted to show that condition (2.4) holds true for any affine process X on S_d^+. Since the automorphism $G : S_d^+ \to S_d^+$ in Proposition 4.14 is order preserving, it suffices to consider affine processes of the canonical form as specified in Proposition 4.14. The following result is a consequence of the Lévy–Khintchine formula on $\mathbb{R}_+$.

LEMMA 4.15. *Let Y be an affine process of canonical form as specified in Proposition 4.14 with parameters denoted by $\widetilde{\cdot}$. Then, for any $y \in \partial S_d^+$, we have*

$$\nabla \det(y) \in N_{S_d^+}(y), \qquad \int_{S_d^+ \setminus \{0\}} \langle \widetilde{\chi}(\xi), \nabla \det(y) \rangle \widetilde{M}(y, d\xi) < \infty \tag{4.19}$$

and

$$\begin{aligned} &\langle \theta, \nabla \det(y) \rangle + \langle \widetilde{B}(y), \nabla \det(y) \rangle - \int_{S_d^+ \setminus \{0\}} \langle \widetilde{\chi}(\xi), \nabla \det(y) \rangle \widetilde{M}(y, d\xi) \\ &\qquad + \frac{1}{2} \sum_{i,j,k,l} \widetilde{A}_{ijkl}(y)\, \partial_{ij}\, \partial_{kl} \det(y) \geq 0. \end{aligned} \tag{4.20}$$

PROOF. Let $y \in \partial S_d^+$ and let $f \in C_c^\infty(S_d^+)$ be a function with $f \geq 0$ and $f(x) = \det(x)$ for all x in a neighborhood of y. Then, for any $v \in \mathbb{R}_+$, the function $x \mapsto e^{-vf(x)} - 1$ lies in $C_c^\infty(S_d^+)$ and thus in $\mathcal{D}(\widetilde{\mathcal{A}})$, where $\widetilde{\mathcal{A}}$ denotes the infinitesimal generator of Y. Note that $f(y) = 0$. Hence, the limit

$$\begin{aligned} \widetilde{\mathcal{A}}(e^{-vf(y)} - 1) &= \lim_{t \to 0^+} \frac{1}{t} \int_{S_d^+} (e^{-vf(\xi)} - 1) \widetilde{p}_t(y, d\xi) \\ &= \lim_{t \to 0^+} \frac{1}{t} \int_{\mathbb{R}_+} (e^{-vz} - 1) p_t^f(y, dz), \end{aligned}$$

exists for any $v \in \mathbb{R}_+$, where $\widetilde{p}_t(y, d\xi)$ denotes the transition function of Y, and $p_t^f(y, dz) = f_* \widetilde{p}_t(y, dz)$ is the pushforward of $\widetilde{p}_t(y, \cdot)$ under f, which is a probability measure supported on $\mathbb{R}_+$.

Using the same arguments as in Proposition 4.9 [i.e., applying Lemma 4.5 as done below (4.4)], and noting that $f(y) = 0$, we conclude that

$$
\begin{aligned}
v \mapsto\ & \widetilde{\mathcal{A}}(e^{-vf(y)} - 1) \\
&= \frac{1}{2} \sum_{i,j,k,l} \widetilde{A}_{ijkl}(y) \big(v^2 \partial_{ij} f(y) \partial_{kl} f(y) - v \partial_{ij} \partial_{kl} f(y)\big) \\
&\quad - v \langle \theta + \widetilde{B}(y), \nabla f(y) \rangle \\
&\quad + \int_{S_d^+ \setminus \{0\}} (e^{-vf(y+\xi)} - 1) \widetilde{m}(d\xi) \\
&\quad + \int_{S_d^+ \setminus \{0\}} \big(e^{-vf(y+\xi)} - 1 + v \langle \widetilde{\chi}(\xi), \nabla f(y) \rangle\big) \widetilde{M}(y, d\xi)
\end{aligned} \tag{4.21}
$$

is the logarithm of the Laplace transform of an infinitely divisible distribution on $\mathbb{R}_+$. Note that

$$
\langle \nabla \det(y), x \rangle = \frac{d}{dt} \det(y + tx) \Big|_{t=0} \begin{cases} \geq 0, & x \in S_d^+, \\ = 0, & x = y. \end{cases}
$$

Hence, $\nabla \det(y) \in N_{S_d^+}(y)$ and the admissibility condition (2.9) implies (4.19). By the Lévy–Khintchine formula on $\mathbb{R}_+$ (see [49], Theorem 3.21), the linear coefficient in v in (4.21) has to be nonpositive. But this is now just (4.20), whence the lemma is proved. $\square$

It now remains to show that (2.4) follows from (4.20). For this purpose, it suffices to evaluate (4.20) at diagonal elements $y \in \partial S_d^+$. Thus, we state the following lemma.

LEMMA 4.16. *Let $y \in S_d^+$ be diagonal, and let $f \in C_c^\infty(S_d^+)$. Then we have*

$$
\begin{aligned}
&\frac{1}{2} \sum_{i,j,k,l=1}^{d} \big(y_{ik}(I_r^d)_{jl} + y_{il}(I_r^d)_{jk} + y_{jk}(I_r^d)_{il} + y_{jl}(I_r^d)_{ik}\big) \frac{\partial^2 f(x)}{\partial x_{ij} \partial x_{kl}} \Big|_{x=y} \\
&\qquad = \frac{1}{2} \sum_{i,j=1}^{d} \big(y_{ii} 1_{\{j \leq r\}} + y_{jj} 1_{\{i \leq r\}}\big) \left(\frac{\partial^2 f(x)}{\partial x_{ij}^2} \Big|_{x=y} + \frac{\partial^2 f(x)}{\partial x_{ij} x_{ji}} \Big|_{x=y} \right).
\end{aligned}
$$

PROOF. Obvious. $\square$

Next, we calculate the partial derivatives of the determinant.

LEMMA 4.17. *Let $y \in S_d^+$ be diagonal, $y = \operatorname{diag}(y_{11}, y_{22}, \dots, y_{dd})$. Then we have*

$$\frac{\partial \det(x)}{\partial x_{ij}}\bigg|_{x=y} = \begin{cases} \prod_{k\neq i} y_{kk}, & \text{if } i = j, \\ 0, & \text{else}, \end{cases}$$

and

$$\frac{\partial^2 \det(x)}{\partial x_{ij} x_{ji}}\bigg|_{x=y} = -\prod_{k=1,k\neq i,k\neq j}^{d} y_{kk} \qquad \text{for } 1 \le i < j \le d,$$

$$\frac{\partial^2 \det(x)}{\partial x_{ij}^2}\bigg|_{x=y} = 0 \qquad \text{for } 1 \le i \le j \le d,$$

where the empty product is defined to be 1.

PROOF. In dimension $d = 2$, the assertion is easily checked, as $\det(y) = y_{11}y_{22} - y_{12}y_{21}$. Therefore, we have

$$\partial_{11} \det(y) = y_{22}, \qquad \partial_{22} \det(y) = y_{11}, \qquad \partial_{12} \det(y) = \partial_{21} \det(y) = 0$$

as well as

$$\partial_{11}^2 \det(y) = \partial_{22}^2 \det(y) = \partial_{12}^2 \det(y) = \partial_{21}^2 \det(y) = 0,$$
$$\partial_{12}\,\partial_{21} \det(y) = \partial_{21}\,\partial_{12} \det(y) = -1.$$

For dimension strictly larger than 2, we employ a combinatorial argument. Recall Leibniz's definition of the determinant,

$$\det(x) = \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \prod_{k=1}^{d} x_{k\sigma(k)}, \tag{4.22}$$

where σ is an element of the permutation group Σ on the set $\{1, 2, \dots, d\}$ and sgn denotes the signum function on Σ, that is, $\operatorname{sgn} = 1$ if σ is an even permutation and $\operatorname{sgn} = -1$ if it is odd. Differentiation of (4.22) with respect to x_{ij} yields

$$\frac{\partial \det(x)}{\partial x_{ij}}\bigg|_{x=y} = \left(\sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) 1_{\{\sigma(i)=j\}} \prod_{k\neq i} x_{k\sigma(k)}\right)\bigg|_{x=y} = \begin{cases} \prod_{k\neq i} y_{kk}, & \text{if } i = j, \\ 0, & \text{else}. \end{cases}$$

Thus, for the second derivative we have

$$\frac{\partial^2 \det(x)}{\partial x_{ij}\,\partial x_{ji}}\bigg|_{x=y} = \left(\sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) 1_{\{\sigma(i)=j\}} 1_{\{\sigma(j)=i\}} \prod_{k\neq i\neq j} x_{k\sigma(k)}\right)\bigg|_{x=y} = -\prod_{k\neq i\neq j} y_{kk},$$

where the last equality holds since y is diagonal. For $\partial_{ij}^2 \det(x)$, the statement is obvious. □

We are prepared to prove the admissibility condition on the constant drift.

PROPOSITION 4.18. *Let X be an affine process on S_d^+, then* (2.4) *holds, that is*,

$$b \succeq (d-1)\alpha.$$

PROOF. Since the automorphism $G : S_d^+ \to S_d^+$ in Proposition 4.14 is order preserving, it suffices to show that (4.20) in Lemma 4.15 implies

$$\theta \succeq (d-1) I_r^d. \tag{4.23}$$

We show that $\theta_{mm} \ge d-1$, if $r \ge m$. To this end, take again some diagonal $y \in \partial S_d^+$ of form $y = \operatorname{diag}(y_{11} > 0, \dots, y_{mm} = 0, \dots, y_{dd} > 0)$. By Lemmas 4.16 and 4.17, we obtain

$$\begin{aligned}
&\sum_{i=1}^d \theta_{ii}\, \partial_{ii} \det(y) + \sum_{i,j} (\widetilde{B}(y))_{ij}\, \partial_{ij} \det(y) \\
&\quad - \int_{S_d^+ \setminus \{0\}} \Big(\sum_{i,j} (\widetilde{\chi}(\xi))_{ij}\, \partial_{ij} \det(y) \Big) \widetilde{M}(y, d\xi) \\
&\quad + \frac{1}{2} \sum_{i,j=1}^d \big((y_{ii} 1_{\{j \le r\}} + y_{jj} 1_{\{i \le r\}}) (\partial_{ij}^2 \det(y) + \partial_{ij}\, \partial_{ji} \det(y)) \big) \\
&\quad = \sum_{i=1}^d \Big(\theta_{ii} \prod_{k \ne i} y_{kk} \Big) + \sum_{l \ne m} \Big(\widetilde{\beta}_{mm}^{ll} y_{ll} \prod_{k \ne m} y_{kk} \Big) \\
&\qquad - \sum_{l \ne m} \int_{S_d^+ \setminus \{0\}} \frac{(\widetilde{\chi}(\xi))_{mm} y_{ll} \prod_{k \ne m} y_{kk}}{\|\xi\|^2 \wedge 1} \widetilde{\mu}_{ll}(d\xi) \\
&\qquad - \frac{1}{2} \sum_{i \ne j} \Big(\prod_{k \ne j} y_{kk} 1_{\{j \le r\}} + \prod_{k \ne i} y_{kk} 1_{\{i \le r\}} \Big) \\
&\quad = \theta_{mm} \prod_{k \ne m} y_{kk} + \prod_{k \ne m} y_{kk} \Big(\sum_{l \ne m} \Big(\widetilde{\beta}_{mm}^{ll} y_{ll} - y_{ll} \int_{S_d^+ \setminus \{0\}} \frac{(\widetilde{\chi}(\xi))_{mm}}{\|\xi\|^2 \wedge 1} \widetilde{\mu}_{ll}(d\xi) \Big) \Big) \\
&\qquad - (d-1) \prod_{k \ne m} y_{kk} 1_{\{m \le r\}} \ge 0.
\end{aligned}$$

As $\prod_{k \ne m} y_{kk} > 0$ and by (2.11) also

$$\Big(\widetilde{\beta}_{mm}^{ll} y_{ll} - y_{ll} \int_{S_d^+ \setminus \{0\}} \frac{(\widetilde{\chi}(\xi))_{mm}}{\|\xi\|^2 \wedge 1} \widetilde{\mu}_{ll}(d\xi) \Big) \ge 0$$

for $l \ne m$, letting $y_{ll} \to 0, l \ne m$ yields $\theta_{mm} \ge d-1$ for $r \ge m$. Relabeling of indices then proves (4.23). $\square$

5. Sufficient conditions for the existence and uniqueness of affine processes. In this section, we prove that for a given admissible parameter set α, b, β^{ij}, c, γ, m, μ satisfying the conditions of Definition 2.3, there exists a unique affine process on S_d^+, whose infinitesimal generator $\mathcal{A}$ is of form (2.12). Our approach to derive this result is to consider the martingale problem for the operator $\mathcal{A}$. In order to prove uniqueness for this martingale problem, we shall need the following existence and uniqueness result for the generalized Riccati differential equations (2.14) and (2.15).

5.1. *Generalized Riccati differential equations*. We first derive some properties of the function R given in (2.17).

LEMMA 5.1. *R is analytic on S_d^{++} and quasi-monotone increasing on S_d^+.*

PROOF. That R is analytic on S_d^{++} follows by dominated convergence (see, e.g., [16], Lemma A.2).

Now let $\delta > 0$, and define

$$\begin{aligned}
R^\delta(u) &= -2u\alpha u + B^\top(u) + \gamma - \int_{\{\|\xi\|\geq\delta\}} \biggl(\frac{e^{-\langle u,\xi\rangle} - 1 + \langle \chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1}\biggr)\mu(d\xi) \\
&= -2u\alpha u + \gamma + \biggl(B^\top(u) - \int_{\{\|\xi\|\geq\delta\}} \frac{\langle \chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1}\mu(d\xi)\biggr) \\
&\quad + \int_{\{\|\xi\|\geq\delta\}} \biggl(\frac{1 - e^{-\langle u,\xi\rangle}}{\|\xi\|^2 \wedge 1}\biggr)\mu(d\xi).
\end{aligned}$$

Now, the map $u \mapsto -2u\alpha u + \gamma$ is quasi-monotone increasing, as it is shown in Step 3 of the proof of Proposition 4.9. Furthermore, it follows from the admissibility condition (2.11) that

$$u \mapsto B^\top(u) - \int_{\{\|\xi\|\geq\delta\}} \frac{\langle \chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1}\mu(d\xi)$$

is a quasi-monotone increasing linear map on S_d^+. Finally, the quasi-monotonicity of

$$u \mapsto \int_{\{\|\xi\|\geq\delta\}} \biggl(\frac{1 - e^{-\langle u,\xi\rangle}}{\|\xi\|^2 \wedge 1}\biggr)\mu(d\xi)$$

is a consequence of the monotonicity of the exponential and that $\operatorname{supp}(\mu) \subseteq S_d^+$.

By dominated convergence, we have $\lim_{\delta\to 0} R^\delta(u) = R(u)$ pointwise for each $u \in S_d^+$. Hence, the quasi-monotonicity carries over to R. Indeed, choose $x, u, v \in S_d^+$ such that $u \preceq v$ and $\langle v - u, x\rangle = 0$. Then we have for all δ, $\langle R^\delta(v) - R^\delta(u), x\rangle \geq 0$. Thus,

$$\langle R^\delta(v) - R^\delta(u), x\rangle \to \langle R(v) - R(u), x\rangle \geq 0,$$

as $\delta \to 0$, which proves that R is quasi-monotone increasing. $\square$

LEMMA 5.2. *There exists a constant K such that*

$$\langle u, R(u)\rangle \leq \frac{K}{2}(\|u\|^2+1), \qquad u \in S_d^+. \tag{5.1}$$

PROOF. We may assume, without loss of generality, that the truncation function in Definition 2.3 takes the form $\chi(\xi) = 1_{\{\|\xi\|\leq 1\}}\xi$ [otherwise adjust $B(u)$ accordingly]. Then, for all $u \in S_d^+$ we have

$$\begin{aligned}
R(u) &= -2u\alpha u + B^\top(u) + \gamma - \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|\leq 1\}} \underbrace{\left(\frac{e^{-\langle u,\xi\rangle} - 1 + \langle \xi, u\rangle}{\|\xi\|^2}\right)}_{\geq 0} \mu(d\xi) \\
&\quad - \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|>1\}} (e^{-\langle u,\xi\rangle} - 1)\mu(d\xi) \\
&\preceq -2u\alpha u + B^\top(u) + \gamma + \mu(S_d^+ \cap \{\|\xi\| > 1\}) \\
&\preceq B^\top(u) + \gamma + \mu(S_d^+ \cap \{\|\xi\| > 1\}),
\end{aligned} \tag{5.2}$$

where we use that

$$-\int_{S_d^+\setminus\{0\}\cap\{\|\xi\|>1\}} (e^{-\langle u,\xi\rangle} - 1)\mu(d\xi) \preceq \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|>1\}} \mu(d\xi).$$

Set now

$$\overline{\gamma} := \gamma + \mu(S_d^+ \cap \{\|\xi\| > 1\}) \in S_d^+.$$

By (5.2), we obtain, for $u \in S_d^+$, that

$$\langle u, R(u)\rangle \leq \langle u, B^\top(u)\rangle + \langle u, \overline{\gamma}\rangle,$$

from which we derive the existence of a positive constant K such equation (5.1) holds. □

Here is our main existence and uniqueness result for the generalized Riccati differential equations (2.14) and (2.15).

PROPOSITION 5.3. *For every $u \in S_d^{++}$, there exists a unique global $\mathbb{R}_+ \times S_d^{++}$-valued solution (ϕ, ψ) of* (2.14) *and* (2.15). *Moreover, $\phi(t,u)$ and $\psi(t,u)$ are analytic in $(t,u) \in \mathbb{R}_+ \times S_d^{++}$.*

PROOF. We only have to show that, for every $u \in S_d^{++}$, there exists a unique global S_d^{++}-valued solution ψ of (2.15), as then ϕ is uniquely determined by integrating (2.14) and has the desired properties by admissibility of the parameter set.

Let $u \in S_d^{++}$. Since R is analytic on S_d^{++}, standard ODE results (e.g., [14], Theorem 10.4.5) yield there exists a unique local S_d^{++}-valued solution $\psi(t,u)$ of (2.15) for $t \in [0, t_+(u))$, where

$$t_+(u) = \lim_{n\to\infty} \inf\{t \geq 0 \mid \|\psi(t,u)\| \geq n \text{ or } \psi(t,u) \in \partial S_d^+\} \leq \infty.$$

It thus remains to show that $t_+(u) = \infty$. That $\psi(t,u)$, and hence $\phi(t,u)$, is analytic in $(t,u) \in \mathbb{R}_+ \times S_d^{++}$ then follows from [14], Theorem 10.8.2.

Since R may not be Lipschitz continuous at ∂S_d^+ (see Remark 5.4 below), we first have to regularize it. We thus define

$$\widetilde{R}(u) = -2u\alpha u + B^\top(u) + \gamma - \int_{S_d^+\setminus\{0\}\cap\{\|\xi\|\leq 1\}} \left(\frac{e^{-\langle u,\xi\rangle} - 1 + \langle \xi, u\rangle}{\|\xi\|^2}\right)\mu(d\xi).$$

It then follows as in Lemmas 5.1 and 5.2 that $\widetilde{R}$ is quasi-monotone increasing on S_d^+ and that (5.1) holds for some constant $\widetilde{K}$. Moreover, $\widetilde{R}$ is analytic on S_d. Hence, for all $u \in S_d$, there exists a unique local S_d-valued solution $\widetilde{\psi}$ of

$$\frac{\partial \widetilde{\psi}(t,u)}{\partial t} = \widetilde{R}(\widetilde{\psi}(t,u)), \qquad \widetilde{\psi}(0,u) = u,$$

for all $t \in [0, \widetilde{t}_+(u))$ with maximal lifetime

$$\widetilde{t}_+(u) = \lim_{n\to\infty} \inf\{t \geq 0 \mid \|\widetilde{\psi}(t,u)\| \geq n\} \leq \infty.$$

From (5.1), we infer that for all $u \in S_d^+$ and $t < \widetilde{t}_+(u)$,

$$\partial_t \|\widetilde{\psi}(t,u)\|^2 = 2\langle \widetilde{\psi}(t,u), \partial_t \widetilde{\psi}(t,u)\rangle \leq \widetilde{K}\big(\|\widetilde{\psi}(t,u)\|^2 + 1\big).$$

Gronwall's inequality (e.g., [14], (10.5.1.3)) implies

$$\|\widetilde{\psi}(t,u)\|^2 \leq e^{\widetilde{K}t}(\|u\|^2 + 1), \qquad t < \widetilde{t}_+(u). \tag{5.3}$$

Hence, $\widetilde{t}_+(u) = \infty$ for $u \in S_d^+$. As $\widetilde{R}$ is quasi-monotone increasing on S_d^+, Volkmann's comparison Theorem 4.8 now yields

$$0 \preceq \widetilde{\psi}(t,u) \preceq \widetilde{\psi}(t,v), \qquad t \geq 0, \text{ for all } 0 \preceq u \preceq v.$$

Therefore and since $\widetilde{\psi}(t,u)$ is also analytic in u, Lemma 3.3 implies that $\widetilde{\psi}(t,u) \in S_d^{++}$ for all $(t,u) \in \mathbb{R}_+ \times S_d^{++}$.

We now carry this over to $\psi(t,u)$ and assume without loss of generality, as in the proof of Lemma 5.2, that the truncation function in Definition 2.3 takes the form $\chi(\xi) = 1_{\{\|\xi\|\leq 1\}}\xi$. Then

$$R(u) - \widetilde{R}(u) = -\int_{S_d^+\setminus\{0\}\cap\{\|\xi\|>1\}} \big(e^{-\langle u,\xi\rangle} - 1\big)\mu(d\xi) \succeq 0, \qquad u \in S_d^+.$$

Hence, for $u \in S_d^{++}$ and $t < t_+(u)$, we have

$$\frac{\partial \widetilde{\psi}(t,u)}{\partial t} - \widetilde{R}(\widetilde{\psi}(t,u)) = \frac{\partial \psi(t,u)}{\partial t} - R(\psi(t,u)) \preceq \frac{\partial \psi(t,u)}{\partial t} - \widetilde{R}(\psi(t,u)).$$

Theorem 4.8 thus implies

$$\psi(t,u) \succeq \widetilde{\psi}(t,u) \in S_d^{++}, \qquad t \in [0, t_+(u)).$$

Hence, $t_+(u) = \liminf_{n\to\infty}\{t \geq 0 \mid \|\psi(t,u)\| \geq n\}$. Using (5.1) again, we now can show as for $\widetilde{\psi}$ that

$$\|\psi(t,u)\|^2 \leq e^{Kt}(\|u\|^2 + 1), \qquad t < t_+(u).$$

Hence $t^+(u) = \infty$, as desired. $\square$

REMARK 5.4. Lemma 5.1 states that the admissibility of the parameters $\alpha, \beta^{ij}, \gamma, \mu$ implies quasi-monotonicity of R on S_d^+.[9] Moreover, quasi-monotonicity just means that R is "inward pointing" close to the boundary S_d^+. Indeed, let $u, x \in S_d^+$ with $\langle u, x\rangle = 0$. Then $\langle R(u), x\rangle \geq \langle \gamma, x\rangle \geq 0$. Hence, if R were Lipschitz continuous on S_d^+, a deterministic variant of Theorem A.5 would imply the invariance of S_d^+ with respect to (2.15) right away. However, the map R might fail to be Lipschitz at ∂S_d^+ (see the one-dimensional counterexample [16], Example 9.3), even though it is analytic on the interior S_d^{++}. Here, quasi-monotonicity plays the decisive role. It leads to the phenomenon that $\psi(t,u)$ stays away from the boundary ∂S_d^+ for $u \in S_d^{++}$, which is of crucial importance in our analysis.

5.2. *The martingale problem for $\mathcal{A}$.* We are now prepared to study the martingale problem for the operator $\mathcal{A}$ given by (2.12). For the notion of martingale problems, we refer to [17], Chapter 4. We shall proceed in four steps. First, we approximate $\mathcal{A}$ by regular operators $\mathcal{A}^{\varepsilon,\delta,n}$ on the space $\mathcal{S}_+$ of rapidly decreasing C^∞-functions on S_d^+, defined in (B.1). Second, using Theorem A.5 below, we show that there exists an S_d^+-valued càdlàg solution of the martingale problem for $\mathcal{A}^{\varepsilon,\delta,n}$. Third, a subsequence of these solutions is shown to converge to an $S_d^+ \cup \{\Delta\}$-valued càdlàg solution of the martingale problem for $\mathcal{A}$. Finally, we show that this solution is unique, Markov and affine, as desired.

Note that we cannot employ Stroock's [51] seminal existence and uniqueness results for martingale problems, since those are solved on $\mathbb{R}^n$ and require uniform elliptic diffusion parts. Neither of these is satisfied in our case.

Now let $(\alpha, b, \beta^{ij}, c = 0, \gamma = 0, m, \mu)$ be some admissible parameter set. Fix some $\varepsilon, \delta > 0$ and $n \in \mathbb{N}$. In order to bound the coefficients and cut off the small jumps, we let

$$\varphi_n \in C_b^\infty(S_d), \qquad 0 \leq \varphi_n \leq 1, \qquad \varphi_n(x) = \begin{cases} 1, & \|x\| \leq n, \\ \dfrac{n}{\|x\|}, & \|x\| \geq n+1. \end{cases} \tag{5.4}$$

[9]We conjecture that the converse also holds: R is quasi-monotone on S_d^+ and $\mathrm{supp}(\mu) \subseteq S_d^+$ if and only if the parameters $\alpha, \beta^{ij}, \gamma, \mu$ are admissible.

We then define the bounded and smooth parameters

$$B^n(x) = B(\varphi_n(x)x),$$
$$m^\delta(d\xi) = m(d\xi)1_{\{\|\xi\|>\delta\}},$$
$$M^{\delta,n}(x, d\xi) = \left\langle \varphi_n(x)x, \frac{\mu(d\xi)}{\|\xi\|^2 \wedge 1} 1_{\{\|\xi\|>\delta\}} \right\rangle.$$

Concerning the diffusion function $A_{ijkl}(x)$ given by (2.13), we first find an appropriate factorization which will allow us to write the continuous martingale part of X as a stochastic integral. Thereto observe that any S_d^+-valued solution, presumed that it exists, of the following symmetric matrix-valued diffusion SDE:

$$dZ_t = \sqrt{Z_t}\, dW_t\, \Sigma + \Sigma^\top dW_t^\top \sqrt{Z_t}, \tag{5.5}$$

where W is a standard $d \times d$-matrix Brownian motion and $\Sigma \in M_d$ with $\Sigma^\top \Sigma = \alpha$, has quadratic variation $d\langle Z_{ij}, Z_{kl}\rangle_t = A_{ijkl}(Z_t)$. Define now $\sigma^{kl}(x) \in S_d$ by

$$\sigma^{kl}(x) = \sqrt{x} M^{kl} \Sigma + \Sigma^\top M^{lk} \sqrt{x}, \tag{5.6}$$

where $M_{ij}^{kl} = \delta_{ik}\delta_{jl}$. Then (5.5) can be written as

$$dZ_t = \sum_{k,l=1}^{d} \sigma^{kl}(Z_t)\, dW_{t,kl}$$

and $A_{ijkl}(x) = \sum_{m,n=1}^{d} \sigma_{ij}^{mn}(x)\sigma_{kl}^{mn}(x)$.

Since $\sigma^{kl}(x)$ involves the matrix square root, which is neither Lipschitz continuous nor bounded nor globally defined, we need to introduce some approximating regularization in order to meet the assumptions of Theorem A.5 below. Thereto fix some truncation function

$$\eta_\varepsilon \in C_b^\infty(S_d), \qquad \eta_\varepsilon(x) = \begin{cases} 1, & x \in S_d^+, \\ 0, & x \notin S_d^+ - \varepsilon I_d, \end{cases}$$

and define

$$s_{\varepsilon,n}(x) = \begin{cases} \eta_\varepsilon(\varphi_n(x)x)\big(\sqrt{\varphi_n(x)x + \varepsilon I_d} - \sqrt{\varepsilon I_d}\big), & \text{if } x \in S_d^+ - \varepsilon I_d, \\ 0, & \text{otherwise.} \end{cases} \tag{5.7}$$

Note that $s_{\varepsilon,n}$ satisfies:

- $s_{\varepsilon,n} \in C_b^\infty(S_d, S_d)$,
- $s_{\varepsilon,n}(x) = (\sqrt{\varphi_n(x)x + \varepsilon I_d} - \sqrt{\varepsilon I_d})$ on S_d^+,
- $\lim_{\varepsilon\to 0^+} s_{\varepsilon,n}(x) = \sqrt{\varphi_n(x)x}$.

With this, we can now define the regularization of σ^{kl} by

$$\sigma_{\varepsilon,n}^{kl}(x) = s_{\varepsilon,n}(x) M^{kl} \Sigma + \Sigma^\top M^{lk} s_{\varepsilon,n}(x), \tag{5.8}$$

which then satisfies the smoothness condition of Theorem A.5. Finally, we set

$$
\begin{aligned}
A^{\varepsilon,n}_{ijkl}(x) &= \sum_{m,n}^{d} (\sigma^{mn}_{\varepsilon,n}(x))_{ij} (\sigma^{mn}_{\varepsilon,n}(x))_{kl} \\
&= (s^2_{\varepsilon,n}(x))_{ik}\alpha_{jl} + (s^2_{\varepsilon,n}(x))_{il}\alpha_{jk} \\
&\quad + (s^2_{\varepsilon,n}(x))_{jk}\alpha_{il} + (s^2_{\varepsilon,n}(x))_{jl}\alpha_{ik},
\end{aligned}
\tag{5.9}
$$

and define the corresponding regularized operator on $C_0(S_d)$

$$
\begin{aligned}
\mathcal{A}^{\varepsilon,\delta,n} f(x) &= \frac{1}{2} \sum_{i,j,k,l} A^{\varepsilon,n}_{ijkl}(x) \frac{\partial^2 f(x)}{\partial x_{ij}\,\partial x_{kl}} \\
&\quad + \sum_{i,j} \bigl(b_{ij} + B^n_{ij}(x)\bigr) \frac{\partial f(x)}{\partial x_{ij}} \\
&\quad + \int_{S_d^+\setminus\{0\}} \bigl(f(x+\xi) - f(x)\bigr) m^\delta(d\xi) \\
&\quad + \int_{S_d^+\setminus\{0\}} \bigl(f(x+\xi) - f(x) - \langle \chi(\xi), \nabla f(x)\rangle\bigr) M^{\delta,n}(x, d\xi).
\end{aligned}
\tag{5.10}
$$

We now show that $\mathcal{A}^{\varepsilon,\delta,n}$ approximates $\mathcal{A}$. We let $\mathcal{S} = \mathcal{S}(S_d)$ and $\mathcal{S}_+$ denote the locally convex spaces of rapidly decreasing C^∞-functions on S_d and S_d^+ defined in (B.1) below, respectively.

LEMMA 5.5. *$\mathcal{S} \subset \mathcal{D}(\mathcal{A}^{\varepsilon,\delta,n})$ and, for every $f \in \mathcal{S}_+$,*

$$
\lim_{\varepsilon,\delta,n} \|\mathcal{A}^{\varepsilon,\delta,n} f - \mathcal{A} f\|_\infty = 0. \tag{5.11}
$$

PROOF. Since φ_n as defined in (5.4) converges uniformly on compact sets to 1, this is clear for the differential operator part. Concerning the integral part, we have

$$
\begin{aligned}
&\left\| \int_{S_d^+\setminus\{0\}} \bigl(f(x+\xi) - f(x) - \langle 1_{\{\|\xi\|\le 1\}}\xi, \nabla f(x)\rangle\bigr)\bigl(M^{\delta,n}(x, d\xi) - M(x, d\xi)\bigr) \right\| \\
&\qquad \le \left\| \sum_{i,j} \int_{S_d^+\setminus\{0\}} \left(\frac{f(x+\xi) - f(x) - \langle 1_{\{\|\xi\|\le 1\}}\xi, \nabla f(x)\rangle}{\|\xi\|^2 \wedge 1} \right) \right. \\
&\qquad\qquad\qquad \left. \times x_{ij}\bigl(\varphi_n(x) - 1\bigr)\mu^\delta_{ij}(d\xi) \right\| \\
&\qquad\quad + \left\| \sum_{i,j} \int_{S_d^+\setminus\{0\}} \left(\frac{f(x+\xi) - f(x) - \langle 1_{\{\|\xi\|\le 1\}}\xi, \nabla f(x)\rangle}{\|\xi\|^2 \wedge 1} \right) \right. \\
&\qquad\qquad\qquad \left. \times x_{ij}\bigl(1_{\{\|\xi\|>\delta\}} - 1\bigr)\mu_{ij}(d\xi) \right\|.
\end{aligned}
$$

By dominated convergence the second term goes uniformly in x to 0, thus we only have to consider the first one. By splitting the first integral into $\int_{\{\|\xi\|\le 1\}} + \int_{\{\|\xi\|>1\}}$, we note that $\| \int_{\{\|\xi\|\le 1\}} \|$ converges uniformly in x to 0. Hence, it remains to analyze

$$\left\| \sum_{i,j} \int_{\{\|\xi\|>1\}} (f(x+\xi) - f(x)) x_{ij} (\varphi_n(x) - 1) \mu_{ij}(d\xi) \right\|,$$

which can be estimated by

$$\begin{aligned}\sum_{i,j} \bigg(\int_{\{\|\xi\|>1\}} &\| f(x+\xi) x_{ij} (\varphi_n(x) - 1) \| (\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)) \\ &+ \int_{\{\|\xi\|>1\}} \| f(x) x_{ij} (\varphi_n(x) - 1) \| (\mu_{ij}^+(d\xi) + \mu_{ij}^-(d\xi)) \bigg),\end{aligned}$$

where μ_{ij}^+, μ_{ij}^- correspond to the Jordan decomposition of $\mu_{ij} = \mu_{ij}^+ - \mu_{ij}^-$. As f lies in $\mathcal{S}_+$, the second term converges uniformly to 0. For the first one, observe that for every n

$$\| f(x+\xi) x_{ij} (\varphi_n(x) - 1) \| \le \| f(x+\xi) x_{ij} \| \le \| f(x+\xi) \| \| x + \xi \|,$$

such that we can apply dominated convergence. Again, since f lies in $\mathcal{S}_+$, the first integral converges uniformly in x to 0 as well. Hence (5.11) holds true, and $\mathcal{S} \subset \mathcal{D}(\mathcal{A}^{\varepsilon,\delta,n})$ follows similarly. $\square$

We now establish existence for the martingale problem for $\mathcal{A}^{\varepsilon,\delta,n}$.

LEMMA 5.6. *For every $x \in S_d^+$ there exists an S_d^+-valued càdlàg solution X to the martingale problem for $\mathcal{A}^{\varepsilon,\delta,n}$ with $X_0 = x$. That is,*

$$f(X_t) - \int_0^t \mathcal{A}^{\varepsilon,\delta,n} f(X_s)\, ds$$

is a martingale, for all $f \in \mathcal{S}$.

PROOF. Consider the following SDE of type (A.1):

$$\begin{aligned} X_t^{\varepsilon,\delta,n} &= x + \int_0^t \left(b + B^n(X_s^{\varepsilon,\delta,n}) - \int_{S_d^+ \setminus \{0\}} \chi(\xi) M^{\delta,n}(X_s^{\varepsilon,\delta,n}, d\xi) \right) ds \\ &\quad + \sum_{k,l}^{d} \int_0^t \sigma_{\varepsilon,n}^{kl}(X_s^{\varepsilon,\delta,n})\, dW_{s,kl} + J_t, \end{aligned} \tag{5.12}$$

where W is a $d \times d$-matrix of standard Brownian motions and J a finite activity jump process with compensator $m^\delta(d\xi) + M^{\delta,n}(X_t^{\varepsilon,\delta,n}, d\xi)$. Note that the quadratic variation of the continuous martingale part of $X_t^{\varepsilon,\delta,n}$ is given by $A_{ijkl}^{\varepsilon,n}(x)$

as defined in (5.9). It thus follows by inspection that any càdlàg solution $X^{\varepsilon,\delta,n}$ of (5.12) solves the martingale problem for $\mathcal{A}^{\varepsilon,\delta,n}$.

Hence, it remains to show that there exists an S_d^+-valued càdlàg solution of (5.12). Let us recall the normal cone (2.23) to S_d^+. As $b + B^n(x) - \int_{S_d^+\setminus\{0\}} \chi(\xi) \times M^{\delta,n}(x, d\xi)$, $\sigma_{\varepsilon,n}^{kl}(x)$ and $m^\delta(d\xi) + M^{\delta,n}(x, d\xi)$ are designed to satisfy the assumptions of Theorem A.5 and since $x + \operatorname{supp}(m^\delta(\cdot) + M^{\delta,n}(x,\cdot)) \subseteq S_d^+$ for all $x \in S_d^+$, we only have to show that for all $x \in \partial S_d^+$ and $u \in N_{S_d^+}(x)$

$$\langle \sigma_{\varepsilon,n}^{kl}(x), u\rangle = 0, \tag{5.13}$$

$$\begin{aligned}\Big\langle b + B^n(x) - \int_{S_d^+\setminus\{0\}} \chi(\xi) M^{\delta,n}(x, d\xi)\\ - \frac{1}{2}\sum_{k,l=1}^d D\sigma_{\varepsilon,n}^{kl}(x)\sigma_{\varepsilon,n}^{kl}(x), u\Big\rangle \geq 0.\end{aligned} \tag{5.14}$$

Due to the definition of $\sigma_{\varepsilon,n}^{kl}(x)$, respectively, the definition of $s_{\varepsilon,n}(x)$ given in (5.7), condition (5.13) is satisfied. Concerning (5.14), we have by (2.11)

$$\Big\langle B^n(x) - \int_{S_d^+\setminus\{0\}} \chi(\xi) M^{\delta,n}(x, d\xi), u\Big\rangle \geq 0.$$

Moreover, it is shown in Lemma 5.7 below that

$$\Big\langle b - \frac{1}{2}\sum_{k,l=1}^d D\sigma_{\varepsilon,n}^{kl}(x)\sigma_{\varepsilon,n}^{kl}(x), u\Big\rangle \geq 0.$$

The lemma now follows from Theorem A.5. $\square$

LEMMA 5.7. *Let $x = O\Lambda O^\top \in S_d^+$ where $\Lambda = \operatorname{diag}(\lambda_1,\dots,\lambda_d)$ contains the eigenvalues in decreasing order and let $\sigma_{\varepsilon,n}^{kl}$ be defined by* (5.8). *Then, for all $x \in S_d^+$*,

$$\begin{aligned}\frac{1}{2}\sum_{k,l=1}^d D\sigma_{\varepsilon,n}^{kl}(x)\sigma_{\varepsilon,n}^{kl}(x) &= \frac{1}{2}\sum_{i=1}^d \frac{\varphi_n(x)(\sqrt{\lambda_i\varphi_n(x)+\varepsilon} - \sqrt{\varepsilon})}{\sqrt{\lambda_i\varphi_n(x)+\varepsilon}} U^i\\ &\quad + \frac{1}{2}\sum_{i\neq j} \frac{\varphi_n(x)(\sqrt{\lambda_j\varphi_n(x)+\varepsilon} - \sqrt{\varepsilon})}{\sqrt{\lambda_i\varphi_n(x)+\varepsilon} + \sqrt{\lambda_j\varphi_n(x)+\varepsilon}} U^i\\ &\quad + \frac{1}{2}\sum_{i,k,l} \frac{\lambda_i}{2\sqrt{\lambda_i\varphi_n(x)+\varepsilon}} \langle\nabla\varphi_n(x), \sigma_{\varepsilon,n}^{kl}\rangle Z^{ikl},\end{aligned} \tag{5.15}$$

where $U_{mn}^i = ((\Sigma^\top\Sigma)O)_{mi}O_{ni} + ((\Sigma^\top\Sigma)O)_{ni}O_{mi}$ and $Z_{mn}^{ikl} = O_{mi}O_{ki}\Sigma_{ln} + O_{ni}O_{ki}\Sigma_{lm}$.

Furthermore, *if*

$$b \succeq (d-1)\Sigma^\top \Sigma, \tag{5.16}$$

then

$$\left\langle b - \frac{1}{2}\sum_{k,l=1}^{d} D\sigma_{\varepsilon,n}^{kl}(x)\sigma_{\varepsilon,n}^{kl}(x), u\right\rangle \geq 0 \tag{5.17}$$

for all $x \in \partial S_d^+$ *and for all* $u \in N_{S_d^+}(x)$.

PROOF. Let us denote

$$C^{\varepsilon,n}(x) = \frac{1}{2}\sum_{k,l=1}^{d} D\sigma_{\varepsilon,n}^{kl}(x)\sigma_{\varepsilon,n}^{kl}(x),$$

and notice that

$$\begin{aligned} C^{\varepsilon,n}(x) = \frac{1}{2}\sum_{k,l}\biggl(&\frac{d}{dt}s_{\varepsilon,n}\bigl(x + t\sigma_{\varepsilon,n}^{kl}(x)\bigr)\big|_{t=0} M^{kl}\Sigma \\ &+ \Sigma^\top (M^{kl})^\top \frac{d}{dt}s_{\varepsilon,n}\bigl(x + t\sigma_{\varepsilon,n}^{kl}(x)\bigr)\big|_{t=0}\biggr). \end{aligned}$$

We now use the following formula from [29], Theorem 6.6.30:

$$\frac{d}{dt}f(V(t)) = O(t)\biggl(\sum_{i,j}\Delta f(\lambda_i(t), \lambda_j(t))M^{ii}[O(t)^\top V'(t)O(t)]M^{jj}\biggr)O(t)^\top,$$

where $V(t) = O(t)\operatorname{diag}(\lambda_1(t), \dots, \lambda_d(t))O(t)^\top$ is a family of symmetric matrices and $\Delta f(u, v) = \frac{(f(u)-f(v))}{(u-v)}$ for $u \neq v$ and $\Delta f(u, u) = f'(u)$. This holds true if $V(\cdot)$ is continuously differentiable for $t \in (a, b)$ and $f(\cdot)$ is continuously differentiable on an open real interval which contains all eigenvalues of $V(t)$ for all $t \in (a, b)$.

We now apply this formula to our case, where $f(t) = \sqrt{t}$ and

$$V(t) = \varphi_n\bigl(x + t\sigma_{\varepsilon,n}^{kl}(x)\bigr)\bigl(x + t\sigma_{\varepsilon,n}^{kl}(x)\bigr) + \varepsilon I_d.$$

Since we take the derivative at $t = 0$, we only have to consider

$$V(0) = O\bigl(\varphi_n(x)\Lambda + \varepsilon I_d\bigr)O^\top,$$

where O is the orthogonal matrix diagonalizing x and

$$V'(0) = \langle \nabla\varphi_n(x), \sigma_{\varepsilon,n}^{kl}(x)\rangle x + \varphi_n(x)\sigma_{\varepsilon,n}^{kl}(x).$$

Note that we do not have an explicit contribution of η_ε which is part of the definition of $s_{\varepsilon,n}$, since $\eta_\varepsilon(S_d^+) = 1$ and $\nabla\eta_\varepsilon(S_d^+) = 0$. Some lines of calculations then yield (5.15).

Let us now verify (5.17). Take an arbitrary $x = O\Lambda O^\top \in \partial S_d^+$ and assume first that it has rank $d-1$, that is, $\lambda_d = 0$ and all other eigenvalues are strictly positive. By Lemma 4.1 and (2.23), the elements of $N_{S_d^+}(x)$ can then be written as $u = OKO^\top$, where $K = \operatorname{diag}(0,\dots,0,k)$ with $k \geq 0$. Thus, (5.17) now reads

$$\langle b - C^{\varepsilon,n}(x), OKO^\top\rangle = k[O^\top bO - O^\top C^{\varepsilon,n}(x)O]_{dd}.$$

As $[O^\top U^i O]_{dd} = 2\delta_{id}(O^\top\Sigma^\top\Sigma O)_{id}$ and $O^\top Z^{ikl}O = 2\delta_{id}O_{ki}(\Sigma O)_{ld}$, we have

$$[O^\top C^{\varepsilon,n}(x)O]_{dd} = \sum_{j\neq d}\frac{\varphi_n(x)(\sqrt{\lambda_j\varphi_n(x)+\varepsilon}-\sqrt{\varepsilon})}{\sqrt{\lambda_j\varphi_n(x)+\varepsilon}+\sqrt{\varepsilon}}[O^\top\Sigma^\top\Sigma O]_{dd}.$$

Since $\sum_{j\neq d}\frac{\varphi_n(x)(\sqrt{\lambda_j\varphi_n(x)+\varepsilon}-\sqrt{\varepsilon})}{\sqrt{\lambda_j\varphi_n(x)+\varepsilon}+\sqrt{\varepsilon}} \leq d-1$, we obtain by condition (5.16)

$$[O^\top bO - O^\top C^{\varepsilon,n}(x)O]_{dd} \geq \bigl[O^\top\bigl(b-(d-1)\Sigma^\top\Sigma\bigr)O\bigr]_{dd} \geq 0,$$

which proves (5.17) for $x \in \partial S_d^+$ with $\operatorname{rk} = d-1$. In the general case, we can proceed similarly. For $x \in \partial S_d^+$ with $\operatorname{rk} = r \leq d-1$, the elements of $N_{S_d^+}(x)$ are given by $u = OKO^\top$, where

$$K = \begin{pmatrix} 0 & 0 \\ 0 & k \end{pmatrix}$$

with $k \in S_{d-r}^+$. This follows again from Lemma 4.1 and (2.23). Now, (5.17) can be written as

$$\begin{aligned}
&\langle b - C^{\varepsilon,n}(x), OKO^\top\rangle \\
&\quad = \biggl\langle O^\top\biggl(b - \sum_{j\leq r}\frac{\varphi_n(x)(\sqrt{\lambda_j\varphi_n(x)+\varepsilon}-\sqrt{\varepsilon})}{\sqrt{\lambda_j\varphi_n(x)+\varepsilon}+\sqrt{\varepsilon}}\Sigma^\top\Sigma\biggr)O, K\biggr\rangle \\
&\quad \geq \langle O^\top(b - r\Sigma^\top\Sigma)O, K\rangle \geq 0,
\end{aligned}$$

which proves the assertion. $\square$

Combining Lemmas 5.5 and 5.6, we obtain the announced existence result for the martingale problem for $\mathcal{A}$.

LEMMA 5.8. *For every $x \in S_d^+$, there exists an $S_d^+ \cup \{\Delta\}$-valued càdlàg solution X to the martingale problem for $\mathcal{A}$ with $X_0 = x$. That is,*

$$f(X_t) - \int_0^t \mathcal{A}f(X_s)\,ds$$

is a martingale, for all $f \in \mathcal{S}_+$.

PROOF. By Lemma 5.6, there exists a solution $X^{\varepsilon,\delta,n}$ to the martingale problem for $\mathcal{A}^{\varepsilon,\delta,n}$ with sample paths in $\mathbb{D}(S_d^+)$ (the space of S_d^+-valued càdlàg paths), and hence also in $\mathbb{D}(S_d^+ \cup \{\Delta\})$. We now claim that $(X^{\varepsilon,\delta,n})$ is relatively compact considered as a sequence of processes with sample paths in $\mathbb{D}(S_d^+ \cup \{\Delta\})$.[10] For the proof of this assertion, we shall make use of Theorems 9.1 and 9.4 in Chapter 3 of [17]. In order to meet the assumption of [17], Chapter 3, Theorem 9.4, we take $C_c^\infty(S_d^+)$ as subalgebra of $C_b(S_d^+)$. Then, for every $T > 0$ and $f \in C_c^\infty(S_d^+)$, we have

$$\sup_{\varepsilon,\delta,n} \mathbb{E}_x\Big[\operatorname*{essup}_{t\in[0,T]} |\mathcal{A}^{\varepsilon,\delta,n} f(X_t^{\varepsilon,\delta,n})|\Big] < \infty,$$

since there exists a constant C such that $\|\mathcal{A}^{\varepsilon,\delta,n} f\|_\infty \le C p_{3,+}(f) < \infty$ for all n, ε, δ, where $p_{k,+}$ are the semi-norms as defined in (B.2) (see also the proof of Proposition 4.12). Thus, the requirements of [17], Chapter 3, Theorem 9.4, are satisfied. Note that Y in the notation of [17], Chapter 3, Theorem 9.4, corresponds in our case to $f(X)$ such that [17], Chapter 3, Condition (9.17), is automatically fulfilled. It then follows by the conclusion of [17], Chapter 3, Theorem 9.4, that $(f(X_t^{\varepsilon,\delta,n}))$ is relatively compact [as family of processes with sample paths in $\mathbb{D}(\mathbb{R})$] for each $f \in C_c^\infty(S_d^+)$. Furthermore, since we consider $S_d^+ \cup \{\Delta\}$, the compact containment condition is always satisfied, that is, for every $\eta > 0$ and $T > 0$, there exists a compact set $\Gamma_{\eta,T} \subset (S_d^+ \cup \{\Delta\})$ for which

$$\inf_{\varepsilon,\delta,n} \mathbb{P}_x\big[X_t^{\varepsilon,\delta,n} \in \Gamma_{\varepsilon,T} \text{ for } t \in [0,T]\big] \ge 1-\eta$$

holds true. By [17], Chapter 3, Theorem 9.1, and the fact that $\{1, C_c^\infty(S_d^+)\}$ is dense in $C(S_d^+ \cup \{\Delta\})$, we therefore obtain that $(X^{\varepsilon,\delta,n})$ is relatively compact in $\mathbb{D}(S_d^+ \cup \{\Delta\})$. Thus, there exists a subsequence $(\mathbb{P}^{\varepsilon_k,\delta_k,n_k})$ of the probability distributions associated to $(X^{\varepsilon,\delta,n})$ which converges in the Prohorov metric to some limit probability distribution. By [17], Chapter 3, Theorem 3.1, this implies weak convergence of $(\mathbb{P}^{\varepsilon_k,\delta_k,n_k})$ and hence the subsequence $(X^{\varepsilon_k,\delta_k,n_k})$ converges in distribution to some limit process X in $\mathbb{D}(S_d^+ \cup \{\Delta\})$.

Combining this with Lemma 5.5 and [17], Chapter 4, Lemma 5.1, we conclude that X is a solution to the martingale problem for $\mathcal{A}$. Hence, the lemma is proved. □

We can now prove the existence and uniqueness of an affine process for any admissible parameter set.

[10]This means that the family of probability distributions associated to $(X^{\varepsilon,\delta,n})$ is relatively compact, that is, the closure of $(\mathbb{P}^{\varepsilon,\delta,n})$ in $\mathcal{P}(\mathbb{D}(S_d^+ \cup \{\Delta\}))$ is compact. Here, $\mathcal{P}(\mathbb{D}(S_d^+ \cup \{\Delta\}))$ denotes the family of probability distributions on $\mathbb{D}(S_d^+ \cup \{\Delta\})$ and $\mathbb{P}^{\varepsilon,\delta,n}$ the distribution of $X^{\varepsilon,\delta,n}$.

PROPOSITION 5.9. *Let* $(\alpha, b, \beta^{ij}, c, \gamma, m, \mu)$ *be an admissible parameter set. Then there exists a unique affine process on* S_d^+ *with infinitesimal generator* (2.12), *and* (2.1) *holds for all* $(t, u) \in \mathbb{R}_+ \times S_d^+$, *where* $\phi(t, u)$ *and* $\psi(t, u)$ *are given by* (2.14) *and* (2.15).

PROOF. Suppose first that $c = 0$ and $\gamma = 0$. Let $x \in S_d^+$. Then Lemma 5.8 implies the existence of an $S_d^+ \cup \{\Delta\}$-valued càdlàg solution X of the martingale problem for $\mathcal{A}$ with $X_0 = x$. We now show that X is unique in distribution.

Thereto, note that by [17], Chapter 4, Theorem 7.1,

$$f(t, X_t) - \int_0^t \big(\mathcal{A} f(s, X_s) + \partial_s f(s, X_s)\big)\, ds \tag{5.18}$$

is a martingale for all rapidly decreasing functions $f \in \mathcal{S}(\mathbb{R}_+ \times S_d^+)$, similarly defined as $\mathcal{S}_+$ in (B.1). Now let ϕ and ψ be the unique solutions of the generalized Riccati differential equations (2.14) and (2.15), given by Proposition 5.3. Fix $t > 0$, $u \in S_d^{++}$, and some $f \in \mathcal{S}(\mathbb{R}_+ \times S_d^+)$ such that

$$f(s, x) = e^{-\phi(t-s,u) - \langle \psi(t-s,u), x\rangle}, \qquad 0 \le s \le t, x \in S_d^+.$$

Then

$$\mathcal{A} f(s, x) + \partial_s f(s, x) = 0, \qquad 0 \le s \le t, x \in S_d^+.$$

In view of (5.18), the Laplace transform of X_t at u is thus given by

$$\mathbb{E}_x\big[e^{-\langle u, X_t\rangle}\big] = \mathbb{E}_x[f(t, X_t)] = f(0, x) - 0 = e^{-\phi(t,u) - \langle \psi(t,u), x\rangle}. \tag{5.19}$$

Since $u \in S_d^{++}$ was arbitrary, we conclude that the distribution of X_t is uniquely determined for all $t > 0$. From [17], Chapter 4, Theorem 4.2, we infer that X is a Markov process with generator $\mathcal{A}$ on $\mathcal{S}_+$ and thus unique in law as solution of the martingale problem for $\mathcal{A}$. Moreover, by (5.19), X is stochastically continuous and affine. Thus, the proposition is proved under the premise that $c = 0$ and $\gamma = 0$.

For general parameters c and γ, we employ a Feynman–Kac argument. Denote by $\mathcal{B}$ and (Q_t) the affine generator and corresponding Feller semigroup associated with $(\alpha, b, \beta^{ij}, c = 0, \gamma = 0, m, \mu)$ from the first part of the proof, respectively. Since $x \mapsto c + \langle \gamma, x\rangle$ is nonnegative on S_d^+, it follows along the lines of [16], Proposition 11.1, that

$$P_t f(x) = \mathbb{E}_x\big[e^{-\int_0^t c + \langle \gamma, X_s\rangle\, ds} f(X_t)\big]$$

defines a Feller semigroup (P_t) on $C_0(S_d^+)$ with infinitesimal generator $\mathcal{A} f(x) = \mathcal{B} f(x) - (c + \langle \gamma, x\rangle) f(x)$ for $f \in \mathcal{S}_+$, which is the desired solution. $\square$

5.3. *An alternative existence proof for jump processes.* For affine processes without diffusion component (i.e., the admissible parameter α vanishes), the existence question can be handled entirely as in the case of affine processes on $\mathbb{R}_+^m \times \mathbb{R}^n$ [16], Section 7. In this section, we elaborate an alternative existence proof in this specific case, by following the lines of [16]. Note that the OU-type processes driven by matrix Lévy subordinators [3] are contained in the class of pure jump processes of this section.

We call a function $f : S_d^+ \to \mathbb{R}$ of Lévy–Khintchine form on S_d^+, if

$$f(u) = \langle b_0, u\rangle - \int_{S_d^+\setminus\{0\}} (e^{-\langle u,\xi\rangle} - 1) m_0(d\xi),$$

where $b_0 \in S_d^+$ and m_0 is a Borel measure supported on S_d^+ such that

$$\int_{S_d^+\setminus\{0\}} (\|\xi\| \wedge 1) m_0(d\xi) < \infty.$$

Once again, we recall that a distribution on S_d^+ is infinitely divisible if and only if its Laplace transform takes the form $e^{-f(u)}$, where f is of the above form (see also Step 1 in the proof of Proposition 4.9).

Similarly to [16], we introduce the sets

$$\mathcal{C} := \{f + c \mid f : S_d^+ \to \mathbb{R} \text{ is of Lévy–Khintchine form on } S_d^+, c \in \mathbb{R}_+\},$$
$$\mathcal{C}_S := \{\psi \mid u \mapsto \langle \psi(u), x\rangle \in \mathcal{C} \text{ for all } x \in S_d^+\}.$$

The following technical statement can be obtained easily by mimicking the proofs of the corresponding statements in [16], Proposition 7.2 and Lemma 7.5:

LEMMA 5.10. *We have*:

(i) $\mathcal{C}, \mathcal{C}_S$ *are convex cones in* $C(S_d^+)$.

(ii) $\phi \in \mathcal{C}$, $\psi \in \mathcal{C}_S$ *imply* $\phi(\psi) \in \mathcal{C}$.

(iii) $\psi, \psi_1 \in \mathcal{C}_S$ *imply* $\psi_1(\psi) \in \mathcal{C}_S$.

(iv) *If* $\phi_k \in \mathcal{C}$ *converges to a continuous function* ϕ *on* S_d^+, *then* $\phi \in \mathcal{C}$. *A similar statement holds for sequences in* $\mathcal{C}_S$.

(v) *Let* $(\alpha = 0, b, \beta^{ij}, c, \gamma, m, \mu)$ *be an admissible parameter set. Then* $R^\delta \to R$ *locally uniformly as* $\delta \to 0$, *where* R^δ *corresponds to the admissible parameter set* $(\alpha = 0, b, \beta^{ij}, c, \gamma, m, \mu 1_{\{\|\xi\|\geq\delta\}})$. (*Note that there is one fixed truncation function.*)

PROPOSITION 5.11. *Let* $(\alpha = 0, b, \beta^{ij}, c, \gamma, m, \mu)$ *be an admissible parameter set. Then for all* $t \geq 0$, *the solutions* $(\phi(t, \cdot), \psi(t, \cdot))$ *of* (2.14) *and* (2.15) *lie in* $(\mathcal{C}, \mathcal{C}_S)$.

PROOF. Suppose first that[11]

$$\int_{S_d^+\setminus\{0\}} \frac{\mu_{ij}(d\xi)}{\|\xi\| \wedge 1} < \infty \tag{5.20}$$

for all $i \le j$. Then equation (2.15) is equivalent to the integral equation

$$\psi(t,u) = e^{\widetilde{B}^\top t}(u) + \int_0^t e^{\widetilde{B}^\top (t-s)} \widetilde{R}(\psi(s,u))\, ds, \tag{5.21}$$

where $R(u) = \widetilde{R}(u) + \widetilde{B}^\top(u)$ and $\widetilde{B}^\top \in \mathcal{L}(S_d)$ is given by

$$\widetilde{B}^\top(u) := B^\top(u) - \int_{S_d^+\setminus\{0\}} \frac{\langle \chi(\xi), u\rangle}{\|\xi\|^2 \wedge 1} \mu(d\xi).$$

Here, $e^{\widetilde{B}^\top t}(u)$ is the notation for the semi-group induced by $\partial_t x(t,u) = \widetilde{B}^\top(x(t, u))$, $x(0,u) = u$. Hence, the variation of constants formula yields (5.21).

Due to admissibility condition (2.11), we have that $\widetilde{B}^\top$ is a linear drift which is "inward pointing" at the boundary of S_d^+, which is equivalent to $e^{\widetilde{B}^\top t}$ being a positive semi-group, that is, $e^{\widetilde{B}^\top t}$ maps S_d^+ into S_d^+. Therefore, $e^{\widetilde{B}^\top t} \in \mathcal{C}_S$ and since $\widetilde{R}(u)$ is given by

$$\widetilde{R}(u) = \gamma - \int_{S_d^+\setminus\{0\}} \frac{(e^{-\langle u,\xi\rangle} - 1)}{\|\xi\|^2 \wedge 1} \mu(d\xi)$$

with μ satisfying (5.20), we also have

$$\widetilde{R} \in \mathcal{C}_S. \tag{5.22}$$

Using Picard's iteration and Lemma 5.10, it follows that the sequence $\psi^{(k)}$ defined as

$$\psi^{(0)}(t,u) := u,$$
$$\psi^{(k+1)}(t,u) := e^{\widetilde{B}^\top t}(u) + \int_0^t e^{\widetilde{B}^\top (t-s)} \widetilde{R}\big(\psi^{(k)}(s,u)\big)\, ds,$$

lies in $\mathcal{C}_S$, for each $t \ge 0$, hence so does its limit $\psi(t,\cdot)$. Since $F \in \mathcal{C}$, we have again by Lemma 5.10 $\phi(t,\cdot) = \int_0^t F(\psi(s,\cdot))\, ds \in \mathcal{C}$.

By an application of Lemma 5.10(v), the general case is then reduced to the former, since R^δ clearly satisfies (5.20). □

We are prepared to provide an alternative proof of Proposition 5.9 under the additional assumption $\alpha = 0$: by Proposition 5.11, $(\phi(t,\cdot), \psi(t,\cdot))$ lie in $(\mathcal{C}, \mathcal{C}_S)$. Hence for all $t \ge 0$, $x \in S_d^+$, there exists an infinitely divisible sub-stochastic kernel

[11] According to our conjecture in Section 2.1.4, this would already cover all possible jump measures if $d \ge 2$.

$p_t(x, d\xi)$ with Laplace-transform $e^{-\phi(t,u)-\langle\psi(t,u),x\rangle}$. The Chapman–Kolmogorov equations hold in view of properties (3.1) and (3.2). Whence, Proposition 5.9 follows.

REMARK 5.12. We note that the proof of statement (v) in Lemma 5.10 is much easier than the one of [16], Lemma 7.5, because $\alpha = 0$. However, for $\alpha \neq 0$ and $d \geq 2$, R cannot be locally uniformly approximated by functions R^δ of a pure jump type such as in Lemma 5.10. Indeed, otherwise one could infer as above the existence of an affine process which is infinitely decomposable and has nonvanishing diffusion component. This is in contradiction with Proposition 2.9 and in the case of pure diffusions it contradicts Example 2.8.

6. Proof of the main results.

6.1. *Proof of Theorem* 2.4. The first part is a summary of Propositions 3.4, 4.9, 4.12 and 4.18. The second part follows from Proposition 5.9.

6.2. *Proof of Theorem* 2.6. Let X be a conservative affine process. It is shown in Proposition 4.12 that $\{e^{-\langle u,\cdot\rangle} \mid u \in S_d^{++}\} \subset D(\mathcal{A})$. Hence,

$$e^{-\langle u,X_t\rangle} - e^{-\langle u,x\rangle} - \int_0^t \mathcal{A}e^{-\langle u,X_s\rangle}\,ds$$

is a $(\widetilde{\mathcal{F}}_t, \mathbb{P}_x)$-martingale with $\widetilde{\mathcal{F}}_t$ defined in (2.18). From [31], Theorem II.2.42, combined with (4.16) and Remark 2.5, it then follows that X is a semimartingale with characteristics (2.19)–(2.21). The canonical semimartingale representation ([31], Theorem II.2.34) of X is thus given by

$$\begin{aligned} X_t = x + B_t + X_t^c &+ \int_0^t \int_{S_d^+\setminus\{0\}} \chi(\xi)\big(\mu^X(ds, d\xi) - \nu(ds, d\xi)\big) \\ &+ \int_0^t \int_{S_d^+\setminus\{0\}} \big(\xi - \chi(\xi)\big)\mu^X(ds, d\xi), \end{aligned}$$

where X^c denotes the continuous martingale part, and μ^X the random measure associated with the jumps of X. In order to establish representation (2.22), we find it convenient to consider the vectorization, $\mathrm{vec}(X^c) \in \mathbb{R}^{d^2}$, of X^c. The aim is now to find a d^2-dimensional Brownian motion $\widetilde{W}$ on a possibly enlarged probability space and a $d^2 \times d^2$-matrix-valued function σ such that

$$\mathrm{vec}(X_t^c) = \int_0^t \sigma(X_s)\,d\widetilde{W}_s. \tag{6.1}$$

Thus, σ has to fulfill

$$\begin{aligned} d\langle X_{ij}^c, X_{kl}^c\rangle_t &= X_{t,ik}\alpha_{jl} + X_{t,il}\alpha_{jk} + X_{t,jk}\alpha_{il} + X_{t,jl}\alpha_{ik} \\ &= (\sigma(X_t)\sigma^\top(X_t))_{ijkl}. \end{aligned} \tag{6.2}$$

As suggested by (5.6), we define the entries of the $d^2 \times d^2$-matrix $\sigma(x)$ in terms of $\sigma^{kl}(x)$ given in (5.6) by

$$\sigma_{ijkl}(x) = \sigma_{ij}^{kl}(x) = \sqrt{x}_{ik}\Sigma_{lj} + \Sigma_{il}^\top \sqrt{x}_{jk}. \tag{6.3}$$

Note that the (kl)th column of $\sigma(x)$ is just the vectorization of the matrix $\sigma^{kl}(x)$. We thus obtain $A_{ijkl}(x) = (\sigma(x)\sigma^\top(x))_{ijkl}$. Hence, $\sigma(x)$ satisfies (6.2). Analogous to the proof of [45], Theorem 20.1, we can now build a d^2-dimensional Brownian motion $\widetilde{W}$ on an enlargement of the probability space such that (6.1) holds true. As the (ij)th entry of X^c is given by

$$\begin{aligned} X_{t,ij}^c = \operatorname{vec}(X_t^c)_{ij} &= \int_0^t \sum_{k,l=1}^d \sigma_{ijkl}(X_s)\, d\widetilde{W}_{s,kl} \\ &= \int_0^t \left(\sqrt{X_s}\, dW_s \Sigma + \Sigma^\top dW_s^\top \sqrt{X_s}\right)_{ij}, \end{aligned}$$

where W is the $d \times d$-matrix Brownian motion satisfying $\operatorname{vec}(W) = \widetilde{W}$, we obtain the desired representation.

6.3. *Proof of Theorem* 2.9. We first prove some technical lemmas.

LEMMA 6.1. *Let $g : S_d^+ \to \mathbb{R}$ be an additive function, that is, g satisfies Cauchy's functional equation*

$$g(x + y) = g(x) + g(y), \qquad x, y \in S_d^+. \tag{6.4}$$

Then g can be extended to an additive function $f : S_d \to \mathbb{R}$. Moreover, if g is measurable on S_d^+ then f is measurable on S_d. In that case, f is a continuous linear functional, that is, $f(x) = \langle c, x\rangle$ for some $c \in S_d$.

PROOF. The first part follows from Lemma 4.4.

Concerning measurability, let $E \in \mathcal{B}(\mathbb{R})$, a Borel measurable set. Then we have by the additivity of f,

$$\begin{aligned} f^{-1}(E) = \bigcup_{n=1}^\infty B_n &= \bigcup_{n=1}^\infty \{x + nI_d \mid x \in S_d, f(x) \in E, \|x\| \le n\} - nI_d \\ &= \bigcup_{n=1}^\infty \{y \in S_d \mid f(y) \in E + f(nI_d), \|y - nI_d\| \le n\} - nI_d \\ &= \bigcup_{n=1}^\infty \{y \in S_d^+ \mid g(y) \in E + g(nI_d), \|y - nI_d\| \le n\} - nI_d, \end{aligned}$$

which is again a measurable set, in view of the measurability of g on S_d^+.

For $x \in S_d$ we write $x = (x_i)_i$, where $1 \le i \le \frac{d(d+1)}{2}$. We introduce the additive functions $f_i : \mathbb{R} \to \mathbb{R}$ via $f_i(x_i) = f(0, \dots, 0, x_i, 0, \dots, 0)$. By the just proved measurability of f, we infer that all f_i are measurable functions on $\mathbb{R}$. By [1], Chapter 2, Theorem 8, any additive measurable function on the real line is a continuous linear functional. Hence for each i, we infer the existence of $c_i \in \mathbb{R}$ such that $f_i(x_i) = c_i x_i$ holds. Since $f(x) = \sum_i f_i(x_i)$, it follows that $f(x) = \langle c, x\rangle$ for some $c \in S_d$. $\square$

Also, we consider Cauchy's exponential equation for $h : S_d^+ \to \mathbb{R}_+$, that is,

$$h(x+y) = h(x)h(y), \qquad x, y \in S_d^+. \tag{6.5}$$

LEMMA 6.2. *Suppose $h : S_d^+ \to \mathbb{R}_+$ is measurable, strictly positive, and satisfies (6.5). Then $h(x) = e^{-\langle c,x\rangle}$, for some $c \in S_d$. If $h \le 1$, then $c \in S_d^+$.*

PROOF. Since h is strictly positive, its logarithm yields the well defined function $g : S_d^+ \to \mathbb{R}$, $g(x) := \log h(x)$. Clearly g is additive, hence by the first part of Lemma 6.1, there exists a unique additive extension $f : S_d \to \mathbb{R}$. Also, f is measurable on S_d^+, hence by the second assertion of Lemma 6.1 we have $f(x) = -\langle c, x\rangle$, for some $c \in S_d$. The last statement follows from the monotonicity of the exponential and the self duality of S_d^+. $\square$

REMARK 6.3. The assumption of strict positivity of h in the preceding lemma is essential. Otherwise, there exist solutions h which are not of the asserted form.

Lemma 6.2 is the main ingredient of the proof of the following characterization concerning k-fold convolutions of Markov processes.

LEMMA 6.4. *Let $(\mathbb{P}_x^{(i)})_{x\in S_d^+} \in \mathcal{P}$ $(i = 0, 1, \dots, k)$. Then*

$$\mathbb{P}_{x^{(1)}}^{(1)} * \cdots * \mathbb{P}_{x^{(k)}}^{(k)} = \mathbb{P}_x^{(0)} \qquad \forall x^{(i)} \in S_d^+, \qquad x = x^{(1)} + \cdots + x^{(k)}, \tag{6.6}$$

if and only if for all $\mathbf{t} = (t_1, \dots, t_N) \in \mathbb{R}_+^N$ and $\mathbf{u} = (u^{(1)}, \dots, u^{(N)}) \in (S_d^+)^N$, $N \in \mathbb{N}_0$, there exists $0 < \rho^{(i)}(\mathbf{t}, \mathbf{u}) \le 1$ and $\psi(\mathbf{t}, \mathbf{u}) \in S_d^+$ such that $\prod_{i=1}^k \rho^{(i)}(\mathbf{t}, \mathbf{u}) = \rho^{(0)}(\mathbf{t}, \mathbf{u})$ and

$$\mathbb{E}_x^{(j)}\left[e^{-\sum_{i=1}^N \langle u^{(i)}, X_{t_i}\rangle}\right] = \rho^{(j)}(\mathbf{t}, \mathbf{u}) e^{-\langle \psi(\mathbf{t},\mathbf{u}), x\rangle} \qquad \forall x \in S_d^+, j = 0, 1, \dots, k. \tag{6.7}$$

PROOF. We proceed similarly as in the proof of [16], Lemma 10.3. Fix $k > 1$, $N > 1$, $\mathbf{t}, \mathbf{u}$ and set

$$g^{(j)}(x) := \mathbb{E}_x^{(j)}\big[e^{-\sum_{i=1}^N \langle u^{(i)}, X_{t_i}\rangle}\big].$$

By the definition of the convolution, (6.6) is equivalent to the following:

$$g^{(1)}\big(x^{(1)}\big) \cdot \dots \cdot g^{(k)}\big(x^{(k)}\big) = g^{(0)}(x) \qquad \forall x^{(i)} \in S_d^+, \qquad x = x^{(1)} + \dots + x^{(k)}. \tag{6.8}$$

Hence, the implication (6.7) $\Rightarrow$ (6.8) is obvious. For the converse direction, we observe that $g^{(i)}$ are strictly positive on all of S_d^+. Thus, by (6.8) we have

$$g := g^{(1)}/g^{(1)}(0) = \dots = g^{(k)}/g^{(k)}(0) = g^{(0)}/g^{(0)}(0)$$

and g is a measurable, strictly positive function on S_d^+ satisfying (6.5). Hence, an application of Lemma 6.2 yields the validity of (6.7), where $\rho^{(i)}(\mathbf{t}, \mathbf{u}) = g^{(i)}(0)$. By the definition of $g^{(i)}$, it follows that $0 < \rho^{(i)}(\mathbf{t}, \mathbf{u}) \leq 1$ and $\psi(\mathbf{t}, \mathbf{u}) \in S_d^+$. $\square$

We are prepared to prove Theorem 2.9:

(i) $\Rightarrow$ (ii): due to Lemma 6.4, infinite decomposability implies that X is affine. Also, by the definition of infinite decomposability and by Lemma 6.4 we have that the kth root $(\mathbb{P}_x^{(k)})$ for each $k \geq 1$ is an affine process with state space S_d^+ with exponents $\psi(t, u)$ and $\phi(t, u)/k$. This implies that $(\mathbb{P}_x^{(k)})_{x \in S_d^+}$ has admissible parameters $(\alpha, b/k, \beta^{ij}, c/k, \gamma, m/k, \mu)$. Hence, the admissibility condition proved in Proposition 4.18 implies $b/k \succeq (d-1)\alpha \succeq 0$, for each k, which is impossible, unless $\alpha = 0$ or $d = 1$.

(ii) $\Rightarrow$ (iii): follows from Proposition 5.11, in view of the Lévy–Khintchine form of $-\phi(t, \cdot) - \langle \psi(t, \cdot), x\rangle$, for each $t > 0$.

(iii) $\Rightarrow$ (i): by definition, every transition kernel $p_t(x, d\xi)$ of X is infinitely divisible with Laplace transform $P_t e^{-\langle u, x\rangle} = e^{-\phi(t,u) - \langle x, \psi(t,u)\rangle}$. For each $k \geq 1$, the maps $\phi^{(k)} := \frac{\phi}{k}$, $\psi^{(k)} := \psi$ satisfy the properties (3.1) and (3.2). Also, infinite divisibility implies that for each $(t, x) \in \mathbb{R}_+ \times S_d^+$,

$$Q_t^{(k)} e^{-\langle u, x\rangle} := e^{-\phi^{(k)}(t,u) - \langle \psi^{(k)}(t,u), x/k\rangle}$$

is the Laplace transform of a sub-stochastic measure on S_d^+. In conjunction with Properties (3.1) and (3.2) we may conclude that $Q_t^{(k)}$ gives rise to a Feller semigroup on $C_0(S_d^+)$, which is affine in $y = x/k$. Hence, we have constructed for each $k \geq 1$ a kth root of X which is stochastic continuous by the definition of its characteristic exponents $\phi^{(k)}, \psi^{(k)}$. Thus Theorem 2.9 is proved.

APPENDIX A: EXISTENCE AND VIABILITY OF A CLASS OF JUMP-DIFFUSIONS

In this section, we study existence and viability in a nonempty closed convex set $D \subset \mathbb{R}^n$ of solutions to the equation

$$X_t = x + \int_0^t b(X_s) + \int_0^t \sigma(X_s)\, dW_s + J_t, \tag{A.1}$$

where $b(x) \in C_b(\mathbb{R}^n, \mathbb{R}^n)$, $\sigma(x) \in C_b(\mathbb{R}^n, \mathbb{R}^{n\times m})$ are Lipschitz continuous maps, W a standard m-dimensional Brownian motion and J a finite activity jump process with state-dependent, absolutely continuous compensator $K(X_t, d\xi)\, dt$. We further assume that $x \mapsto K(x, \mathbb{R}^n)$ is bounded.

We tackle this problem in three steps. First, we derive some regularity and existence results for diffusion SDEs. These results are not in the standard literature, we thus provide full proofs. Second, we prove existence of a càdlàg solution X for (A.1). Finally, we provide sufficient conditions for X to be D-valued.

A.1. Diffusion stochastic differential equations. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtered probability space satisfying the usual conditions and carrying an m-dimensional standard Brownian motion W. We consider the following diffusion SDE:

$$X_t = x + \int_0^t b(X_s) 1_{\{\theta \le s\}}\, ds + \int_0^t \sigma(X_s) 1_{\{\theta \le s\}}\, dW_s, \tag{A.2}$$

where $(\theta, x) \in [0, \infty] \times \mathbb{R}^n$ and b and σ are as above. Recall that X is a solution of (A.2) if X is continuous and (A.2) holds for all $t \ge 0$ a.s. In particular, note that this null set depends on (θ, x).

LEMMA A.1. *Fix $T > 0$ and let $p \ge 2$. Furthermore, let Θ_1, Θ_2 be stopping times and for $i = 1, 2$, U_i, $\mathcal{F}_{\Theta_i}$-measurable random variables. Consider the following equations*:

$$X_t = U_1 + \int_0^t b(X_s) 1_{\{\Theta_1 \le s\}}\, ds + \int_0^t \sigma(X_s) 1_{\{\Theta_1 \le s\}}\, dW_s,$$
$$Y_t = U_2 + \int_0^t b(Y_s) 1_{\{\Theta_2 \le s\}}\, ds + \int_0^t \sigma(Y_s) 1_{\{\Theta_2 \le s\}}\, dW_s.$$

Then there exists a constant C depending only on p, T, n, the Lipschitz constants of b and σ and $\|b\|_\infty$, $\|\sigma\|_\infty$ such that for $0 \le t \le T$,

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\le t} \|X_s - Y_s\|^p\Big] &\le C\mathbb{E}\Big[\|U_1 - U_2\|^p + |\Theta_1 \wedge t - \Theta_2 \wedge t|^{p/2} \\ &\qquad + \int_0^t \sup_{u\le s} \|X_u - Y_u\|^p\, ds\Big].\end{aligned} \tag{A.3}$$

PROOF. By the same arguments as in the proof of [45], Lemma 11.5, we first obtain the following estimate:

$$\begin{aligned}
&\sup_{s\le t}\|X_s-Y_s\|^p\\
&\quad\le 3^{p-1}\biggl(\|U_1-U_2\|^p+\biggl(\int_0^t\|b(X_s)1_{\{\Theta_1\le s\}}-b(Y_s)1_{\{\Theta_2\le s\}}\|\,ds\biggr)^p\\
&\qquad\qquad+\sup_{s\le t}\biggl\|\int_0^s\bigl(\sigma(X_u)1_{\{\Theta_1\le u\}}-\sigma(Y_u)1_{\{\Theta_2\le u\}}\bigr)\,dW_u\biggr\|^p\biggr).
\end{aligned}$$

Moreover,

$$\begin{aligned}
&\biggl(\int_0^t\|b(X_s)1_{\{\Theta_1\le s\}}-b(Y_s)1_{\{\Theta_2\le s\}}\|\,ds\biggr)^p\\
&\quad\le 2^{p-1}\biggl(\biggl(\int_{\Theta_1\wedge t}^{(\Theta_1\vee\Theta_2)\wedge t}\|b(X_s)\|\,ds\biggr)^p+\biggl(\int_{\Theta_2\wedge t}^{(\Theta_1\vee\Theta_2)\wedge t}\|b(Y_s)\|\,ds\biggr)^p\\
&\qquad\qquad+\biggl(\int_{(\Theta_1\vee\Theta_2)\wedge t}^{t}\|b(X_s)-b(Y_s)\|\,ds\biggr)^p\biggr)\\
&\quad\le 2^{p-1}\biggl(K|\Theta_1\wedge t-\Theta_2\wedge t|^p+t^{p-1}\int_0^t\|b(X_s)-b(Y_s)\|^p\,ds\biggr)\\
&\quad\le K\biggl(t^{p/2}|\Theta_1\wedge t-\Theta_2\wedge t|^{p/2}+\int_0^t\sup_{u\le s}\|X_u-Y_u\|^p\,ds\biggr).
\end{aligned}$$

For the stochastic integral part, we apply the Burkholder–Davis–Gundy inequality

$$\begin{aligned}
&\mathbb{E}\biggl[\sup_{s\le t}\biggl\|\int_0^s\bigl(\sigma(X_u)1_{\{\Theta_1\le u\}}-\sigma(Y_u)1_{\{\Theta_2\le u\}}\bigr)\,dW_u\biggr\|^p\biggr]\\
&\quad\le K\mathbb{E}\biggl[\biggl(\int_0^t\|\sigma(X_u)1_{\{\Theta_1\le u\}}-\sigma(Y_u)1_{\{\Theta_2\le u\}}\|^2\,du\biggr)^{p/2}\biggr]\\
&\quad\le K\mathbb{E}\biggl[\biggl(\int_{\Theta_1\wedge t}^{(\Theta_1\vee\Theta_2)\wedge t}\|\sigma(X_u)\|^2\,du\biggr)^{p/2}+\biggl(\int_{\Theta_2\wedge t}^{(\Theta_1\vee\Theta_2)\wedge t}\|\sigma(Y_u)\|^2\,du\biggr)^{p/2}\\
&\qquad\qquad+\biggl(\int_{(\Theta_1\vee\Theta_2)\wedge t}^{t}\|\sigma(X_u)-\sigma(Y_u)\|^2\,du\biggr)^{p/2}\biggr]\\
&\quad\le K\mathbb{E}\biggl[|\Theta_1\wedge t-\Theta_2\wedge t|^{p/2}+\int_0^t\|\sigma(X_s)-\sigma(Y_s)\|^p\,ds\biggr]\\
&\quad\le K\mathbb{E}\biggl[|\Theta_1\wedge t-\Theta_2\wedge t|^{p/2}+\int_0^t\sup_{u\le s}\|X_u-Y_u\|^p\,ds\biggr],
\end{aligned}$$

where K always denotes a constant which varies from line to line. The last estimate in both inequalities follows from the the Lipschitz continuity of b and σ. By assembling these pieces, the proof is complete. □

Here is a fundamental existence result, which is not stated in this general form in the standard literature. Therefore, we provide a full proof.

THEOREM A.2. *There exists a function* $Z:[0,\infty]\times\mathbb{R}^n\times\Omega\times\mathbb{R}_+\to\mathbb{R}^n$ *with the following properties*:

(i) $Z(\theta,x,\omega,t)$ *is continuous in* (θ,x,t) *for all* ω.

(ii) Z *is* $\mathcal{B}([0,\infty]\times\mathbb{R}^n)\otimes\mathcal{P}$*-measurable.*[12]

(iii) $Z(\theta,x,\omega,t)$ *solves* (A.2) *for all* (θ,x).

(iv) *Let* Θ *be a stopping time and* U *an* $\mathcal{F}_\Theta$ *measurable random variable*, *then* $X_t=Z(\Theta,U,t)$ *solves*

$$X_t=U+\int_0^t b(X_s)1_{\{\Theta\le s\}}\,ds+\int_0^t\sigma(X_s)1_{\{\Theta\le s\}}\,dW_s. \tag{A.4}$$

PROOF. For every $(\theta,x)\in[0,\infty]\times\mathbb{R}^n$, there exists a unique solution $X_t(\omega)=\widetilde{Z}(\theta,x,\omega,t)$ of (A.2), which is continuous in t. This is a consequence of the Lipschitz continuity of $x\mapsto b(x)1_{\{\theta\le s\}}$ and $x\mapsto\sigma(x)1_{\{\theta\le s\}}$. Uniqueness is meant modulo indistinguishability. From estimate (A.3), we can deduce for $p\ge 2$, $x,y\in[-T,T]^n$, $0\le\theta_1,\theta_2\le T$ and $0\le t\le T$,

$$\begin{aligned}&\mathbb{E}\Big[\sup_{s\le t}\|\widetilde{Z}(\theta_1,x,s)-\widetilde{Z}(\theta_2,y,s)\|^p\Big]\\&\quad\le K\bigg(\|x-y\|^{p/2}+|\theta_1-\theta_2|^{p/2}\\&\qquad\quad+\int_0^t\mathbb{E}\Big[\sup_{u\le s}\|\widetilde{Z}(\theta_1,x,u)-\widetilde{Z}(\theta_2,y,u)\|^p\Big]\,ds\bigg)\end{aligned}$$

for some constant K. Hence, by Gronwall's lemma,

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\le t}\|\widetilde{Z}(\theta_1,x,s)-\widetilde{Z}(\theta_2,y,s)\|^p\Big]&\le Ke^{KT}(\|x-y\|^{p/2}+|\theta_1-\theta_2|^{p/2})\\&\le C\|(\theta_1,x)-(\theta_2,y)\|^{p/2}.\end{aligned}$$

Let now $\mathrm{Dya}=\{j2^{-k},j\in\mathbb{Z},k\in\mathbb{N}\}$ be the set of dyadic rational numbers in $\mathbb{R}$ and $\mathrm{Dya}^n=\mathrm{Dya}\times\cdots\times\mathrm{Dya}$ the set of dyadic rational numbers in $\mathbb{R}^n$. Furthermore, we define M by $M=\mathrm{Dya}^{n+1}\cap([0,T]\times[-T,T]^n)$. By setting $p=2n+4$, we can apply Kolmogorov's lemma. Indeed, analogous to the proof of [32], Theorem 2.8, we derive for all $(\theta_1,x),(\theta_2,y)\in M$ with $0<\|(\theta_1,x)-(\theta_2,y)\|<h(\omega)$, where h is a positive valued random variable, and for all $\omega\in\Omega_T^*$, where $\Omega_T^*\in\mathcal{F}$ is some set depending on T with $\mathbb{P}(\Omega_T^*)=1$, the following estimate:

$$\sup_{s\le t}\|\widetilde{Z}(\theta_1,x,\omega,s)-\widetilde{Z}(\theta_2,y,\omega,s)\|\le\delta\|(\theta_1,x)-(\theta_2,y)\|^\gamma. \tag{A.5}$$

[12]Here, $\mathcal{P}$ denotes the predictable σ-field.

Here, $\gamma \in (0, \frac{1}{p})$ and δ is some positive constant. Let us now define Z: if $\omega \notin \Omega_T^*$, then $Z(\theta, x, \omega, t) = x$ for $0 \le t \le T$. For $\omega \in \Omega_T^*$ and $(\theta, x) \in M$, $Z(\theta, x, \omega, t) = \widetilde{Z}(\theta, x, \omega, t)$ for $0 \le t \le T$. If $(\theta, x) \in M^c$, we choose a sequence $(\theta_n, x_n)_{n\in\mathbb{N}} \subseteq M$ such that $(\theta_n, x_n) \to (\theta, x)$. By estimate (A.5), $\widetilde{Z}(\theta_n, x_n, \omega, t)$ is a Cauchy sequence converging with respect to $\sup_{s\le t}\|\cdot\|$. We can therefore set $Z(\theta, x, \omega, t) = \lim_{n\to\infty} \widetilde{Z}(\theta_n, x_n, \omega, t)$. As we have uniform convergence in t and as $\widetilde{Z}$ is continuous in t, the resulting process Z is jointly continuous in (θ, x, t). Furthermore, for every (θ, x), Z is indistinguishable from $\widetilde{Z}$, that is,

$$\mathbb{P}[Z(\theta, x, t) = \widetilde{Z}(\theta, x, t) \text{ for all } 0 \le t \le T] = 1. \tag{A.6}$$

Indeed, for $(\theta, x) \in M$, this is clear and for $(\theta, x) \in M^c$, we have for $(\theta_n, x_n)_{n\in\mathbb{N}} \subseteq M$ with $(\theta_n, x_n) \to (\theta, x)$

$$\mathbb{P}\Big[\sup_{s\le t}\|\widetilde{Z}(\theta_n, x_n, s) - \widetilde{Z}(\theta, x, s)\| \ge \varepsilon\Big] \le C\varepsilon^{-p}\|(\theta_n, x_n) - (\theta, x)\|^{p/2},$$

which implies that $\widetilde{Z}(\theta_n, x_n, t) \to \widetilde{Z}(\theta, x, t)$ in probability, uniformly in t. As $\widetilde{Z}(\theta_n, x_n, t) \to Z(\theta, x, t)$ a.s., and thus in particular in probability, it follows that $Z(\theta, x, t) = \widetilde{Z}(\theta, x, t)$ a.s. for all $0 \le t \le T$. Letting $T \to \infty$ proves assertion (i).

Statement (ii) is then a consequence of (i) and the $\mathcal{F}_t$-measurability of $\omega \mapsto Z(\theta, x, t, \omega)$, which is satisfied since $\mathcal{F}_0$ contains all null sets of $\mathcal{F}$.

Furthermore, property (A.6) implies that $Z(\theta, x, t)$ is a solution of (A.2) for all (θ, x), which yields assertion (iii).

In order to prove (iv), we proceed in two steps:

Step 1. We first assume that Θ and U take finitely many values $\theta_1, \dots, \theta_k \in [0, \infty]$ and $x_1, \dots, x_l \in \mathbb{R}^n$, respectively. Denote

$$A_j = \{\Theta = \theta_j\}, \qquad B_h = \{U = x_h\}.$$

Then

$$Z(\Theta, U, t) = \sum_{j,h} 1_{A_j \cap B_h} Z(\theta_j, x_h, t)$$

does the job. Indeed, as $A_j \cap B_h$ are disjoint and $A_j \cap B_h \in \mathcal{F}_{\theta_j}$ for all j, h, we have (see, e.g., [36], page 39)

$$\begin{aligned}
U &+ \int_0^t b(Z(\Theta, U, s))1_{\{\Theta\le s\}}\,ds + \int_0^t \sigma(Z(\Theta, U, s))1_{\{\Theta\le s\}}\,dW_s \\
&= U + \int_0^t \sum_{j,h} 1_{A_j\cap B_h} b(Z(\theta_j, x_h, s))1_{\{\theta_j\le s\}}\,ds \\
&\quad + \int_0^t \sum_{j,h} 1_{A_j\cap B_h} \sigma(Z(\theta_j, x_h, s))1_{\{\theta_j\le s\}}\,dW_s
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j,h} 1_{A_j \cap B_h} \biggl(x_h + \int_0^t b(Z(\theta_j, x_h, s)) 1_{\{\theta_j \le s\}} \, ds \\
&\qquad\qquad + \int_0^t \sigma(Z(\theta_j, x_h, s)) 1_{\{\theta_j \le s\}} \, dW_s \biggr) \\
&= \sum_{j,h} 1_{A_j \cap B_h} Z(\theta_j, x_h, t) = Z(\Theta, U, t)
\end{aligned}
$$

for all $t \ge 0$ a.s.

Step 2. For general Θ, U, approximate $\Theta^{(k)} \downarrow \Theta$ by the simple stopping times

$$
\Theta^{(k)} = \begin{cases} j2^{-k}, & (j-1)2^{-k} \le \Theta < j2^{-k}, \ j = 1, \dots, k2^k, \\ \infty, & k \le \Theta. \end{cases}
$$

Let $U^{(l)}$ be a sequence of $\mathcal{F}_\Theta$-measurable random variables, each $U^{(l)}$ taking finitely many values, and $U^{(l)} \to U$ in L^2 (such $U^{(l)}$ obviously exists). Moreover, $\{\Theta^{(k)} = \theta_j\} \cap \{U^{(l)} = x_h\} \in \mathcal{F}_{\theta_j}$ for all j, h (see [32], Chapter 1, Problem 2.24).

By Step 1, each $Z(\Theta^{(k)}, U^{(l)})$ satisfies the respective SDE. Moreover, from estimate (A.3) and Grownwall's lemma we deduce that for any $T > 0$, there exists a constant C such that

$$
\begin{aligned}
&\mathbb{E}\Bigl[\sup_{t \le T} \bigl\| Z\bigl(\Theta^{(k)}, U^{(l)}, t\bigr) - Z\bigl(\Theta^{(k')}, U^{(l')}, t\bigr) \bigr\|^2\Bigr] \\
&\qquad \le C e^{CT} \mathbb{E}\bigl[\bigl\| U^{(k)} - U^{(k')} \bigr\|^2 + \bigl|\Theta^{(k)} \wedge T - \Theta^{(k')} \wedge T\bigr|\bigr].
\end{aligned}
$$

Hence, $Z(\Theta^{(k)}, U^{(l)})$ is a Cauchy sequence and thus converging with respect to $\mathbb{E}[\sup_{t \le T} \| \cdot \|^2]$, for all $T > 0$, to some continuous process X satisfying (A.4). On the other hand, by the continuity of $(\theta, x) \mapsto Z(\theta, x, t)$, we know that

$$
Z\bigl(\Theta^{(k)}, U^{(l)}, t\bigr) \to Z(\Theta, U, t)
$$

for all ω and $t \ge 0$. Again, by continuity of $t \mapsto Z(\Theta, U, t)$, we conclude that $Z(\Theta, U) = X$ up to indistinguishability, which proves the claim. $\square$

A.2. Existence of jump-diffusions. We now provide a constructive proof for the existence of a solution of (A.1) on a specific stochastic basis which is defined as follows:

- $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \ge 0})$ is a filtered space, where $\Omega := \Omega_1 \times \Omega_2$, $\mathcal{F}_t := \mathcal{G}_t \otimes \mathcal{H}_t$ and $\mathcal{F} = \mathcal{G} \otimes \mathcal{H}$ are precisely defined below. Note that we do not have a measure on $(\Omega, \mathcal{F})$ for the moment. The generic sample element will be denoted by $\omega = (\omega_1, \omega_2) \in \Omega$.
- $(\Omega_1, \mathcal{G}, (\mathcal{G}_t)_{t \ge 0}, \mathbb{P}_1)$ is some filtered probability space satisfying the usual conditions and carrying an m-dimensional standard Brownian motion W. We shall consider the above diffusion SDE (A.2) on Ω_1 and thus obtain the respective solution $Z(\theta, x, \omega_1, t)$ satisfying the regularity properties of Theorem A.2.

- $(\Omega_2, \mathcal{H})$ is the canonical space for $\mathbb{R}^n$-valued marked point processes (see, e.g., [30]): Ω_2 consists of all càdlàg, piecewise constant functions $\omega_2 : [0, T_\infty(\omega_2)) \to \mathbb{R}^n$ with $\omega_2(0) = 0$ and $T_\infty(\omega_2) = \lim_{n\to\infty} T_n(\omega_2) \leq \infty$, where $T_n(\omega_2)$, defined by $T_0 = 0$ and

$$T_n(\omega_2) := \inf\{t > T_{n-1}(\omega_2) \mid \omega_2(t) \neq \omega_2(t-)\} \wedge \infty, \qquad n \geq 1,$$

are the successive jump times of ω_2. We denote by

$$J_t(\omega) = J_t(\omega_2) = \omega_2(t) \qquad \text{on } [0, T_\infty(\omega_2))$$

the canonical jump process, and let $\mathcal{H}_t = \sigma(J_s \mid s \leq t)$ be its natural filtration with $\mathcal{H} = \mathcal{H}_\infty$. Note that T_n are $(\mathcal{H}_t)$ and $(\mathcal{F}_t)$-stopping times if interpreted as $T_n(\omega) = T_n(\omega_2)$.

The following statement is meant to be pointwise, referring to the filtered measure space $(\Omega, \mathcal{F}, (\mathcal{F}_t))$ without reference to a probability measure.

LEMMA A.3. *Let $Z(\theta, x, \omega_1, t)$ be as of Theorem A.2. Then for an $\mathcal{F}_{T_n}$-measurable random variable $U(\omega_1, \omega_2)$ the process $Z(T_n(\omega_2), U(\omega_1, \omega_2), \omega_1, t)$ is*:

(i) *continuous in t for all (ω_1, ω_2)*,
(ii) *$\mathcal{F}_t$-adapted on $\{T_n \leq t\}$*.

PROOF. The first assertion is a consequence of Theorem A.2(i). The second one follows from the $\mathcal{B}([0, \infty] \times \mathbb{R}^n) \otimes \mathcal{P}$-measurability of $Z(\theta, x, \omega_1, t)$, as stated in Theorem A.2(ii), and the fact that T_n and U are $\mathcal{F}_t$-measurable on $\{T_n \leq t\}$. $\square$

Here is our existence result for (A.1).

THEOREM A.4. *There exists a càdlàg $\mathcal{F}_t$-adapted process X and a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ with $\mathbb{P}|_\mathcal{G} = \mathbb{P}_1$, such that X is a solution of* (A.1) *on* $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$.

PROOF. We follow the arguments in the proof of [21], Theorem 5.1, which is based on [30], Theorem 3.6, and proceed in three steps.

Step 1. We start by solving (A.1) along every path ω_2. To this end, let us define recursively: $\Delta\omega_2(0) = \Delta\omega_2(\infty) = 0$, $X_0^{(0)} = x$, and for $n \geq 1$:

$$X_t^{(n)}(\omega_1, \omega_2) = \begin{cases} Z\big(T_{n-1}(\omega_2), X_{T_{n-1}}^{(n-1)}(\omega_1, \omega_2) \\ \qquad\qquad + \Delta\omega_2(T_{n-1}), \omega_1, t\big), & t \in [0, \infty), \\ x_0, & t = \infty, \end{cases}$$

where x_0 is any fixed point in $D \subset \mathbb{R}^n$ and Z satisfies the properties of Theorem A.2. By Lemma A.3, every $X^{(n)}$ is continuous in t for all (ω_1, ω_2) and $\mathcal{F}_t$-adapted on $\{T_n \leq t\}$ since $X^{(n-1)}_{T_{n-1}}(\omega_1, \omega_2) + \Delta\omega_2(T_{n-1})$ is $\mathcal{F}_{T_n}$-measurable. Thus, the process

$$X_t(\omega_1, \omega_2) = \sum_{n \geq 1} X^{(n)}_t(\omega_1, \omega_2) 1_{\{T_{n-1} \leq t < T_n\}} \tag{A.7}$$

is càdlàg $\mathcal{F}_t$-adapted and solves (A.1) on $(\Omega_1, \mathcal{G}, (\mathcal{G}_t), \mathbb{P}_1)$ for $t \in [0, T_\infty(\omega_2))$ and any fixed path ω_2.

Step 2. It remains to show that there exists a probability measure $\mathbb{P}$ such that $K(X_t, d\xi)$ is the compensator of J and $\mathbb{P}|_{\mathcal{G}} = \mathbb{P}_1$ holds true. For this purpose, we shall make use of [30], Theorem 3.6. Let us define the following random measure ν by

$$\nu(dt, d\xi) = \begin{cases} K(X_t, d\xi)\, dt, & t < T_\infty, \\ 0, & t \geq T_\infty. \end{cases}$$

Observe that ν is predictable, since X_t is càdlàg and $\mathcal{F}_t$-adapted. Theorem 3.6 in [30] now implies that there exists a unique probability kernel $\mathbb{P}_2$ from Ω_1 to $\mathcal{H}$, such that ν is the compensator of the random measure μ associated to the jumps of J. On $(\Omega, \mathcal{F})$ we then define the probability measure $\mathbb{P}$ by $\mathbb{P}(d\omega) = \mathbb{P}_1(d\omega_1)\mathbb{P}_2(\omega_1, d\omega_2)$ whose restriction to $\mathcal{G}$ is equal to $\mathbb{P}_1$.

Step 3. We finally show that X defined by (A.7) solves (A.1) on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ for all $t \geq 0$. Note that $W(\omega) = W(\omega_1)$ is an $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$-Brownian motion. This implies that $Z(\theta, x, \omega, t) = Z(\theta, x, \omega_1, t)$ is a solution of (A.2) on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, satisfying the properties of Theorem A.2. It thus remains to show that $T_\infty = \infty$ $\mathbb{P}$-a.s. Let μ be the random measure associated with the jumps. As $x \mapsto K(x, \mathbb{R}^n)$ is bounded, we have for all $T \geq 0$,

$$\mathbb{E}_{\mathbb{P}}\big[\mu([0, T] \times \mathbb{R}^n)\big] = \mathbb{E}_{\mathbb{P}}\big[\nu([0, T] \times \mathbb{R}^n)\big] = \mathbb{E}_{\mathbb{P}}\left[\int_0^T K(X_t, \mathbb{R}^n)\, dt\right] \leq CT$$

for some constant C. This implies that $\mu([0, T] \times \mathbb{R}^n) < \infty$ a.s. for all $T \geq 0$ and hence $\mathbb{P}[T_\infty < \infty] = 0$ or equivalently $T_\infty = \infty$ a.s. □

A.3. Viability of jump-diffusions. Consider a nonempty closed convex set $D \subset \mathbb{R}^n$. We now provide sufficient conditions for the solution X in (A.7) to be D-valued. This result is based on [13], Theorem 4.1. We recall the notion of the *normal cone*

$$N_D(x) = \{u \in \mathbb{R}^n \mid \langle u, y - x\rangle \geq 0, \text{ for all } y \in D\} \tag{A.8}$$

of D at $x \in D$, consisting of inward pointing vectors. See, for example, [28], Definition III.5.2.3, except for a change of the sign.

THEOREM A.5. *Assume that σ also has a Lipschitz continuous derivative. Suppose furthermore that*

$$x + \operatorname{supp}(K(x,\cdot)) \subseteq D, \tag{A.9}$$

$$\langle \sigma^i(x), u\rangle = 0, \tag{A.10}$$

$$\left\langle b(x) - \frac{1}{2}\sum_{i=1}^n D\sigma^i(x)\sigma^i(x), u\right\rangle \geq 0, \tag{A.11}$$

for all $u \in N_D(x)$ and $x \in D$, where σ^i denotes the ith column of σ. Then, for every initial point $x \in D$, the process X defined in (A.7) is a D-valued solution of (A.1).

PROOF. We have to show that $X_t = \sum_{n\geq 1} X_t^{(n)}(\omega_1,\omega_2)1_{\{T_{n-1}\leq t<T_n\}} \in D$ a.s. for all $t \geq 0$. We proceed by induction on n. For $n = 1$, $X_t^{(1)}$, is simply given by

$$X_t^{(1)} = x + \int_0^t b(X_s^{(1)})\,ds + \int_0^t \sigma(X_s^{(1)})\,dW_s.$$

Due to [13], Theorem 4.1, conditions (A.10) and (A.11) imply that for all $t \geq 0$, $X_t^{(1)} \in D$ a.s. Let us now assume that for all $t \geq 0$, $X_t^{(n-1)} \in D$ a.s., thus in particular $X_{T_{n-1}}^{(n-1)} = X_{T_{n-1}-} \in D$ a.s. If $T_{n-1} = \infty$, then we immediately obtain

$$X_t^{(n)} = X_{T_{n-1}}^{(n-1)} + \Delta J_{T_{n-1}} = x_0 \in D.$$

Otherwise, let $f \in C_b(\mathbb{R}^n, \mathbb{R}_+)$ satisfy $\operatorname{supp}(f) \subseteq D^c$. Then,

$$\begin{aligned}\mathbb{E}\big[f\big(X_{T_{n-1}}^{(n-1)} + \Delta J_{T_{n-1}}\big)\big] &= \mathbb{E}[f(X_{T_{n-1}-} + \Delta J_{T_{n-1}})] \\ &= \mathbb{E}\left[\int_{\mathbb{R}^n\setminus\{0\}} f(X_{T_{n-1}-} + \xi)K(X_{T_{n-1}-}, d\xi)\right] = 0,\end{aligned}$$

since by (A.9), $X_{T_{n-1}-} + \operatorname{supp}(K(X_{T_{n-1}-},\cdot)) \subseteq D$ a.s. and $f(D) = 0$. Hence, $f(X_{T_{n-1}}^{(n-1)} + \Delta J_{T_{n-1}}) = 0$ a.s., implying that $X_{T_{n-1}}^{(n-1)} + \Delta J_{T_{n-1}} \notin \operatorname{supp}(f)$ a.s. As this holds true for all $f \in C_b(\mathbb{R}^n, \mathbb{R}_+)$ with $\operatorname{supp}(f) \subseteq D^c$, it follows that $X_{T_{n-1}}^{(n-1)} + \Delta J_{T_{n-1}} \in D$ a.s. Thus, again by [13], Theorem 4.1, and conditions (A.10) and (A.11)

$$X_t^{(n)} = X_{T_{n-1}}^{(n-1)} + \Delta J_{T_{n-1}} + \int_0^t b(X_s^{(n)})1_{\{T_{n-1}\leq s\}}\,ds + \int_0^t \sigma(X_s^{(n)})1_{\{T_{n-1}\leq s\}}\,dW_s$$

a.s. takes values in D, which proves the induction hypothesis. The definition of X then yields the assertion. □

APPENDIX B: AN APPROXIMATION LEMMA ON THE CONE OF POSITIVE SEMIDEFINITE MATRICES

In this section, we deliver a differentiable variant of the Stone–Weierstrass theorem for C^∞-functions on S_d^+. This approximation statement is essential for the description of the generator of an affine semigroup, as is elaborated in Section 4.2.

We employ multi-index notation in the sequel. For $n \geq 1$, a multi-index is an element $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ having length $|\boldsymbol{\alpha}| := \alpha_1 + \cdots + \alpha_n$. The factorial is defined by $\boldsymbol{\alpha}! := \prod_{i=1}^n \alpha_i!$. The partial order $\leq$ is understood componentwise, and so are the elementary operations $+, -$. That is, $\boldsymbol{\alpha} \geq \boldsymbol{\beta}$ if and only if $\alpha_i \geq \beta_i$ for $i = 1, \dots, n$. Moreover, for $\boldsymbol{\alpha} \geq \boldsymbol{\beta}$, the multinomial coefficient is defined by

$$\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} := \frac{\boldsymbol{\alpha}!}{(\boldsymbol{\alpha} - \boldsymbol{\beta})!\boldsymbol{\beta}!}.$$

We define the monomial $x^{\boldsymbol{\alpha}} := \prod_{i=1}^n x_i^{\alpha_i}$, and the differential operator $\partial^{\boldsymbol{\alpha}} := \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}}$. Corresponding to a polynomial $P(x) = \sum_{|\boldsymbol{\alpha}| \leq k} a_{\boldsymbol{\alpha}} x^{\boldsymbol{\alpha}}$, we introduce the differential operator $P(\partial) := \sum_{|\boldsymbol{\alpha}| \leq k} a_{\boldsymbol{\alpha}} \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x^{\boldsymbol{\alpha}}}$.

Let $\mathcal{S} = \mathcal{S}(S_d)$ denote the locally convex space of rapidly decreasing C^∞-functions on S_d (see [46], Chapter 7), and define the space of rapidly decreasing C^∞-functions on S_d^+ via the restriction

$$\mathcal{S}_+ = \{f = F|_{S_d^+} : F \in \mathcal{S}\}. \tag{B.1}$$

Equipped with the increasing family of semi-norms

$$p_{k,+}(f) := \sup_{x \in S_d^+, |\boldsymbol{\alpha}+\boldsymbol{\beta}| \leq k} |x^{\boldsymbol{\alpha}} \partial^{\boldsymbol{\beta}} f(x)|, \tag{B.2}$$

$\mathcal{S}_+$ becomes a locally convex vector space (see [46], Theorem 1.37).

For technical reasons, we also introduce for $\varepsilon \geq 0$ the semi-norms

$$p_{k,\varepsilon}(f) := \sup_{x \in S_d^+ + B_{\leq \varepsilon}(0), |\boldsymbol{\alpha}+\boldsymbol{\beta}| \leq k} |x^{\boldsymbol{\alpha}} \partial^{\boldsymbol{\beta}} f(x)|$$

on $C^\infty(S_d)$, where $B_{\leq r}(y) = \{z \in S_d \mid \|z - y\| \leq r\}$ denotes the closed ball with radius r and center y. Note that $p_{k,+} = p_{k,0}$. We first give an alternative description of $\mathcal{S}_+$.

LEMMA B.1. *We have*

$$\mathcal{S}_+ = \{f = G|_{S_d^+} : G \in C^\infty(S_d) \text{ and } \exists \varepsilon > 0 \text{ such that } p_{k,\varepsilon}(G) < \infty\ \forall k \geq 0\}.$$

PROOF. The inclusion $\subseteq$ is trivial. Hence, we prove $\supseteq$. So let $f = G|_{S_d^+}$ for some $G \in C^\infty(S_d)$ with $p_{k,\varepsilon}(G) < \infty$ for all $k \geq 0$ and some $\varepsilon > 0$.

We choose a standard mollifier $\rho \in C_c^\infty(S_d)$ supported in $B_{\leq \varepsilon/2}(0)$ and satisfying $\rho \geq 0$, $\int \rho = 1$. For $\delta > 0$ we introduce the neighborhoods $K_\delta := S_d^+ + B_{\leq \delta}(0)$

of S_d^+. The convolution $\varphi := \rho * 1_{K_{\varepsilon/2}} \in C^\infty(S_d)$ of the indicator function for $K_{\varepsilon/2}$ with ρ satisfies $\varphi = 1$ on S_d^+ and it vanishes outside K_ε. Furthermore, all derivatives of φ are bounded, since

$$|\partial^{\boldsymbol{\alpha}}\varphi(x)| = \left|\int_{K_{\varepsilon/2}} \partial^{\boldsymbol{\alpha}}\rho(y-x)\,dy\right| = \left|\int_{K_{\varepsilon/2}-x} \partial^{\boldsymbol{\alpha}}\rho(z)\,dz\right|$$
$$\leq \int_{B_{\leq\varepsilon/2}(0)} |\partial^{\boldsymbol{\alpha}}\rho(z)|\,dz < \infty,$$

where the last estimate holds because $\operatorname{supp}\rho \subseteq B_{\leq\varepsilon/2}(0)$.

Now we set $F := G \cdot \varphi$. By construction $F \in C^\infty(S_d)$, $F\,|_{S_d^+} = f$ and F vanishes outside K_ε, because φ does. What is left to show is that $F \in \mathcal{S}$. Since F vanishes outside K_ε, it is sufficient to deliver all estimates of its derivatives on K_ε.

Let $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^{d(d+1)/2}$, then we have by the Leibniz rule

$$x^{\boldsymbol{\alpha}}\,\partial^{\boldsymbol{\beta}}F(x) = x^{\boldsymbol{\alpha}} \sum_{0\leq\boldsymbol{\gamma}\leq\boldsymbol{\beta}} \binom{\boldsymbol{\beta}}{\boldsymbol{\gamma}} (\partial^{\boldsymbol{\beta}-\boldsymbol{\gamma}}\varphi(x))(\partial^{\boldsymbol{\gamma}}G(x))$$
$$= \sum_{0\leq\boldsymbol{\gamma}\leq\boldsymbol{\beta}} \binom{\boldsymbol{\beta}}{\boldsymbol{\gamma}} (\partial^{\boldsymbol{\beta}-\boldsymbol{\gamma}}\varphi(x))(x^{\boldsymbol{\alpha}}\,\partial^{\boldsymbol{\gamma}}G(x)).$$

By assumption $x^{\boldsymbol{\alpha}}\,\partial^{\boldsymbol{\gamma}}G$ is bounded on K_ε, and $(\partial^{\boldsymbol{\beta}-\boldsymbol{\gamma}}\varphi(x))$ is bounded on all of S_d. Hence, by the last equation, we have $\sup_{x\in S_d, |\boldsymbol{\alpha}+\boldsymbol{\beta}|\leq k}|x^{\boldsymbol{\alpha}}\,\partial^{\boldsymbol{\beta}}F(x)| < \infty$ for all $k \in \mathbb{N}_0$, which by definition means $F \in \mathcal{S}$. $\square$

LEMMA B.2. *Let $u \in S_d^{++}$. Then for each $\varepsilon \geq 0$, and for all $k \geq 0$ we have $p_{k,\varepsilon}(\exp(-\langle u, \cdot\rangle)) < \infty$. In particular, we have*

$$f_u := \exp(-\langle u, \cdot\rangle)|_{S_d^+} \in \mathcal{S}_+.$$

That is, $f_u = F_u\,|_{S_d^+}$ for some $F_u \in \mathcal{S}$.

PROOF. Since $u \in S_d^{++}$, there exists a positive constant c such that $\langle u, x\rangle \geq c\|x\|$, for all $x \in S_d^+$. Hence, we obtain by a straightforward calculation, $p_{k,+}(\exp(-\langle u, \cdot\rangle)) < \infty$, for all $k \geq 0$.

Next, let $\varepsilon > 0$, and write $x = y + z$, where $y \in S_d^+$ and $z \in B_{\leq\varepsilon}(0)$ and pick multi-indices $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^{d(d+1)/2}$. Then we have by the binomial formula

$$\begin{aligned} x^{\boldsymbol{\alpha}}\,\partial^{\boldsymbol{\beta}}e^{-\langle u,x\rangle} &= x^{\boldsymbol{\alpha}}(-1)^{|\boldsymbol{\beta}|}u^{\boldsymbol{\beta}}e^{-\langle u,x\rangle} \\ &= (y+z)^{\boldsymbol{\alpha}}(-1)^{|\boldsymbol{\beta}|}u^{\boldsymbol{\beta}}e^{-\langle u,y+z\rangle} \\ &= (-1)^{|\boldsymbol{\beta}|}u^{\boldsymbol{\beta}} \sum_{0\leq\boldsymbol{\gamma}\leq\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}} \big(y^{\boldsymbol{\alpha}}e^{-\langle u,y\rangle}\big)\big(z^{\boldsymbol{\alpha}-\boldsymbol{\gamma}}e^{-\langle u,z\rangle}\big). \end{aligned}$$

Now since z ranges in a compact set, and since $p_{k,+}(\exp(-\langle u,\cdot\rangle)) < \infty$ we see that $x^{\boldsymbol{\alpha}}\partial^{\boldsymbol{\beta}} e^{-\langle u,x\rangle}$ must be bounded uniformly in $x \in S_d^+ + B_{\le\varepsilon}(0)$. Hence, $p_{k,\varepsilon}(\exp(-\langle u,\cdot\rangle)) < \infty$, for all $k \ge 0$.

Together with Lemma B.1, this implies $f_u \in \mathcal{S}_+$. $\square$

We are now prepared to deliver the following density result for the $\mathbb{R}$-linear hull $\mathcal{M}$ of $\{f_u = \exp(-\langle u,\cdot\rangle)|_{S_d^+}, u \in S_d^{++}\}$ in $\mathcal{S}_+$.

THEOREM B.3. *$\mathcal{M}$ is dense in $\mathcal{S}_+$.*

PROOF. Denote by $\mathcal{S}' = \mathcal{S}'(S_d)$ and $\mathcal{S}'_+$ the topological dual of $\mathcal{S}$ and $\mathcal{S}_+$, respectively. The former, $\mathcal{S}'$, is known as the space of tempered distributions. The distributional action is denoted by $\langle\cdot,\cdot\rangle$ and $\langle\cdot,\cdot\rangle_+$ for $\mathcal{S}'$ and $\mathcal{S}'_+$, respectively.

Now suppose by contradiction, that $\mathcal{M}$ is not dense in $\mathcal{S}_+$. Then by [46], Theorem 3.5, there exists some $T_1 \in \mathcal{S}'_+ \setminus \{0\}$ such that $T_1 = 0$ on $\mathcal{M}$. Hence, $\langle T_1, f_u\rangle_+ = 0$, for all $u \in S_d^{++}$. The restriction $F \mapsto F|_{S_d^+}$ yields a continuous linear embedding $\mathcal{S} \hookrightarrow \mathcal{S}_+$. Hence, the restriction T of T_1 to $\mathcal{S}$, given by

$$\langle T,\varphi\rangle := \langle T_1, \varphi|_{S_d^+}\rangle_+, \qquad \varphi \in \mathcal{S}(S_d),$$

yields an element of $\mathcal{S}'$ with $\operatorname{supp}(T) \subseteq S_d^+$. Pick an $F_u \in \mathcal{S}$ according to Lemma B.2. By the definition of T, we have $\langle T, F_u\rangle = \langle T_1, f_u\rangle_+ = 0$, for all $u \in S_d^{++}$. By the Bros–Epstein–Glaser theorem (see [42], Theorem IX.15), there exists a function $G \in C(S_d)$ with $\operatorname{supp}(G) \subseteq S_d^+$, polynomially bounded [i.e., for suitable constants C, N we have $|G(x)| \le C(1+\|x\|)^N$, for all $x \in S_d^+$] and a real polynomial $P(x)$ such that $P(\partial)G = T$ in $\mathcal{S}'$. Hence, we obtain for any $u \in S_d^{++}$

$$\begin{aligned} 0 &= \langle T, F_u\rangle = \langle P(\partial)G, F_u\rangle = \langle G, P(-\partial)F_u\rangle \\ &= \int_{S_d^+} G(x)P(-\partial)F_u(x)\,dx = P(u)\int_{S_d^+} G(x)\exp(-\langle u,x\rangle)\,dx. \end{aligned}$$

But the last factor is just the Laplace transform of G. This implies $G = 0$, hence $T = 0$, which in turn implies that T_1 vanishes on all of $\mathcal{S}_+$, a contradiction. $\square$

Acknowledgments. We thank Martin Keller-Ressel and Alexander Smirnov for discussions and helpful comments.

REFERENCES

[1] ACZÉL, J. and DHOMBRES, J. (1989). *Functional Equations in Several Variables*. *Encyclopedia of Mathematics and Its Applications* **31**. Cambridge Univ. Press, Cambridge. MR1004465

[2] BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2001). Modelling by Lévy processes for financial econometrics. In *Lévy Processes* 283–318. Birkhäuser, Boston, MA. MR1833702

[3] BARNDORFF-NIELSEN, O. E. and STELZER, R. (2007). Positive-definite matrix processes of finite variation. *Probab. Math. Statist.* **27** 3–43. MR2353270

[4] BAUER, H. (1996). *Probability Theory. de Gruyter Studies in Mathematics* **23**. de Gruyter, Berlin. MR1385460

[5] BRU, M.-F. (1989). Diffusions of perturbed principal component analysis. *J. Multivariate Anal.* **29** 127–136. MR0991060

[6] BRU, M.-F. (1991). Wishart processes. *J. Theoret. Probab.* **4** 725–751. MR1132135

[7] BURASCHI, B., CIESLAK, A. and TROJANI, F. (2007). Correlation risk and the term structure of interest rates. Working paper, Univ. St. Gallen.

[8] BURASCHI, B., PORCHIA, P. and TROJANI, F. (2010). Correlation risk and optimal portfolio choice. *J. Finance* **65** 393–420.

[9] DA FONSECA, J., GRASSELLI, M. and IELPO, F. (2008). Hedging (co)variance risk with variance swaps. Working paper ESILV RR-37, Ecole Supérieure d'Ingénierie Léonard de Vinci.

[10] DA FONSECA, J., GRASSELLI, M. and IELPO, F. (2008). Estimating the Wishart affine stochastic correlation model using the empirical characteristic function. Working paper ESILV RR-35, Ecole Supérieure d'Ingénierie Léonard de Vinci.

[11] DA FONSECA, J., GRASSELLI, M. and TEBALDI, C. (2007). Option pricing when correlations are stochastic: An analytical framework. *Review of Derivatives Research* **10** 151–180.

[12] DA FONSECA, J., GRASSELLI, M. and TEBALDI, C. (2008). A multifactor volatility Heston model. *Quant. Finance* **8** 591–604. MR2457710

[13] DA PRATO, G. and FRANKOWSKA, H. (2004). Invariance of stochastic control systems with deterministic arguments. *J. Differential Equations* **200** 18–52. MR2046316

[14] DIEUDONNÉ, J. (1969). *Foundations of Modern Analysis*. Academic Press, New York. MR0349288

[15] DONATI-MARTIN, C., DOUMERC, Y., MATSUMOTO, H. and YOR, M. (2004). Some properties of the Wishart processes and a matrix extension of the Hartman–Watson laws. *Publ. Res. Inst. Math. Sci.* **40** 1385–1412. MR2105711

[16] DUFFIE, D., FILIPOVIĆ, D. and SCHACHERMAYER, W. (2003). Affine processes and applications in finance. *Ann. Appl. Probab.* **13** 984–1053. MR1994043

[17] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. MR0838085

[18] FILIPOVIĆ, D. (2005). Time-inhomogeneous affine processes. *Stochastic Process. Appl.* **115** 639–659. MR2128634

[19] FILIPOVIĆ, D. (2009). *Term-Structure Models: A Graduate Course*. Springer, Berlin. MR2553163

[20] FILIPOVIĆ, D. and MAYERHOFER, E. (2009). Affine diffusion processes: Theory and applications. In *Advanced Financial Modelling. Radon Ser. Comput. Appl. Math.* **8** 125–164. Walter de Gruyter, Berlin. MR2648460

[21] FILIPOVIĆ, D., OVERBECK, L. and SCHMIDT, T. (2009). Dynamic CDO term structure modeling. Forthcoming in mathematical finance, Ecole Polytechnique Fédérale de Lausanne.

[22] GOLUB, G. H. and VAN LOAN, C. F. (1996). *Matrix Computations*, 3rd ed. Johns Hopkins Univ. Press, Baltimore, MD. MR1417720

[23] GOURIEROUX, C., MONTFORT, A. and SUFANA, R. (2007). International money and stock market contingent claims. Working paper, CREST, CEPREMAP and Univ. Toronto.

[24] GOURIEROUX, C. and SUFANA, R. (2007). Wishart quadratic term structure models. Working paper, CREST, CEPREMAP and Univ. Toronto.

[25] GOURIEROUX, C. and SUFANA, R. (2007). Derivative pricing with Wishart multivariate stochastic volatility: Application to credit risk. Working paper, CREST, CEPREMAP and Univ. Toronto.

[26] GRASSELLI, M. and TEBALDI, C. (2008). Solvable affine term structure models. *Math. Finance* **18** 135–153. MR2380943
[27] HESTON, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. of Financial Studies* **6** 327–343.
[28] HIRIART-URRUTY, J.-B. and LEMARÉCHAL, C. (1993). *Convex Analysis and Minimization Algorithms. I. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **305**. Springer, Berlin. MR1261420
[29] HORN, R. A. and JOHNSON, C. R. (1991). *Topics in Matrix Analysis*. Cambridge Univ. Press, Cambridge. MR1091716
[30] JACOD, J. (1974/75). Multivariate point processes: Predictable projection, Radon–Nikodým derivatives, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **31** 235–253. MR0380978
[31] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877
[32] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940
[33] KATO, T. (1995). *Perturbation Theory for Linear Operators*. Springer, Berlin. MR1335452
[34] KELLER-RESSEL, M. (2009). Affine processes—theory and applications in mathematical finance. Ph.D. thesis, Vienna Univ. Technology.
[35] KELLER-RESSEL, M., SCHACHERMAYER, W. and TEICHMANN, J. (2010). Affine processes are regular. *Probab. Theory Related Fields*. DOI: 10.1007/s00440-010-0309-4. To appear.
[36] LAMBERTON, D. and LAPEYRE, B. (2008). *Introduction to Stochastic Calculus Applied to Finance*, 2nd ed. Chapman and Hall/CRC, Boca Raton, FL. MR2362458
[37] LANG, S. (1993). *Real and Functional Analysis*, 3rd ed. *Graduate Texts in Mathematics* **142**. Springer, New York. MR1216137
[38] LEIPPOLD, M. and TROJANI, F. (2008). Asset pricing with matrix affine jump diffusions. Working paper, University of Zurich—Swiss Banking Institute (ISB).
[39] MAYERHOFER, E., MUHLE-KARBE, J. and SMIRNOV, A. G. (2011). A characterization of the martingale property of exponentially affine processes. *Stochastic Process. Appl.* **121** 568–582.
[40] MAYERHOFER, E., PFAFFEL, O. and STELZER, R. (2009). On strong solutions for positive definite jump–diffusions. VIF Working Paper No. 30, Vienna Institute of Finance.
[41] NARASIMHAN, R. (1971). *Several Complex Variables*. Univ. Chicago Press, Chicago. MR0342725
[42] REED, M. and SIMON, B. (1975). *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness*. Academic Press, New York. MR0493420
[43] REVUZ, D. and YOR, M. (1991). *Continuous Martingales and Brownian Motion*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1083357
[44] ROCKAFELLAR, R. T. (1997). *Convex Analysis*. Princeton Univ. Press, Princeton, NJ. MR1451876
[45] ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol.* 2. Cambridge Univ. Press, Cambridge. MR1780932
[46] RUDIN, W. (1991). *Functional Analysis*, 2nd ed. McGraw-Hill, New York. MR1157815
[47] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. MR1739520
[48] SEMADENI, Z. (1971). *Banach Spaces of Continuous Functions. Vol. I*. PWN—Polish Scientific Publishers, Warsaw. MR0296671
[49] SKOROHOD, A. V. (1991). *Random Processes with Independent Increments*. *Mathematics and Its Applications* (*Soviet Series*) **47**. Kluwer Academic, Dordrecht. MR1155400

[50] STOKES, A. N. (1974). A special property of the matrix Riccati equation. *Bull. Austral. Math. Soc.* **10** 245–253. MR0342748

[51] STROOCK, D. W. (1975). Diffusion processes associated with Lévy generators. *Z. Wahrsch. Verw. Gebiete* **32** 209–244. MR0433614

[52] VOLKMANN, P. (1973). Über die Invarianz konvexer Mengen und Differentialungleichungen in einem normierten Raume. *Math. Ann.* **203** 201–210. MR0322305

C. CUCHIERO
J. TEICHMANN
ETH ZÜRICH
DEPARTEMENT MATHEMATIK
RÄMISTRASSE 101, 8092 ZÜRICH
SWITZERLAND
E-MAIL: christa.cuchiero@math.ethz.ch
josef.teichmann@math.ethz.ch

D. FILIPOVIĆ
ÉCOLE POLYTECHNIQUE FÉDÉRALE
DE LAUSANNE
SWISS FINANCE INSTITUTE
QUARTIER UNIL-DORIGNY
EXTRANEF BUILDING, 1015 LAUSANNE
SWITZERLAND
E-MAIL: damir.filipovic@epfl.ch

E. MAYERHOFER
VIENNA INSTITUTE OF FINANCE
HEILIGENSTÄDTER STRASSE 46–48
1190 VIENNA
AUSTRIA
E-MAIL: eberhard.mayerhofer@vif.ac.at