Adv. Appl. Prob. **52,** 825–854 (2020)
*doi:*10.1017/apr.2020.21

ERGODICITY OF AFFINE PROCESSES ON THE CONE OF SYMMETRIC POSITIVE SEMIDEFINITE MATRICES

MARTIN FRIESEN,* *University of Wuppertal*
PENG JIN,** *Shantou University*
JONAS KREMER,*** *University of Wuppertal*
BARBARA RÜDIGER,**** *University of Wuppertal*

Abstract

This article investigates the long-time behavior of conservative affine processes on the cone of symmetric positive semidefinite $d \times d$ matrices. In particular, for conservative and subcritical affine processes we show that a finite log-moment of the state-independent jump measure is sufficient for the existence of a unique limit distribution. Moreover, we study the convergence rate of the underlying transition kernel to the limit distribution: first, in a specific metric induced by the Laplace transform, and second, in the Wasserstein distance under a first moment assumption imposed on the state-independent jump measure and an additional condition on the diffusion parameter.

Keywords: Affine process; invariant distribution; limit distribution; ergodicity

2010 Mathematics Subject Classification: Primary 60J25; 37A25
Secondary 60G10; 60J75

1. Introduction

An affine process on the cone of symmetric positive semidefinite $d \times d$ matrices $\mathbb{S}_d^+$ is a stochastically continuous Markov process taking values in $\mathbb{S}_d^+$, whose log-Laplace transform depends in an affine way on the initial state of the process. *Affine processes* on the state space $\mathbb{S}_d^+$ were first systematically studied in the seminal article of Cuchiero *et al.* [12]. In their work, the $\mathbb{S}_d^+$-valued affine process is constructed and completely characterized through a set of *admissible parameters*, and the related *generalized Riccati equations* are investigated. Subsequent developments complementing the results of [12] can be found in [33], [42], [43], and [44]. Note that the notion of affine processes is not restricted to the state space $\mathbb{S}_d^+$. For affine processes on other finite-dimensional cones, particularly the canonical one $\mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$, we refer the reader to [2], [5], [6], [15], [16], [17], [28], [33], and [35]. We remark that the above list is far from complete.

Received 21 June 2019; revision received 1 March 2020.
* Postal address: School of Mathematics and Natural Sciences, University of Wuppertal, 42119 Wuppertal, Germany. E-mail address: friesen@math.uni-wuppertal.de
** Postal address: Department of Mathematics, Shantou University, Shantou, Guangdong 515063, China. E-mail address: pjin@stu.edu.cn
*** Postal address: School of Mathematics and Natural Sciences, University of Wuppertal, 42119 Wuppertal, Germany. E-mail address: kremer@math.uni-wuppertal.de
**** Postal address: School of Mathematics and Natural Sciences, University of Wuppertal, 42119 Wuppertal, Germany. E-mail address: ruediger@uni-wuppertal.de

The importance of $\mathbb{S}_d^+$-valued affine processes has been demonstrated by the rapidly growing number of their applications in mathematical finance. In particular, they provide natural models for the evolution of the covariance matrix of multi-asset prices that exhibit random dependence, such as the Wishart process [10], the jump-type Wishart process [38], and a certain class of matrix-valued Ornstein–Uhlenbeck processes driven by Lévy subordinators [7]. Among these, the Wishart process is the most popular one, and it can be used as a multivariate covariance model that extends the well-known Heston model [26]; see also [3], [9], [11], [14], [21], [22], [23], [24], and [25]. The jump-type Wishart process was introduced by Leippold and Trojani [38] to provide additional model flexibility. In [38] the jump-type Wishart process is used in multivariate option pricing, fixed-income models, and dynamic portfolio choice. For a more detailed review of financial applications of affine processes on $\mathbb{S}_d^+$ we refer the reader to the introduction of [12]; see also the references therein.

In this article, we investigate the long-time behavior of affine processes on $\mathbb{S}_d^+$; i.e., we prove the existence, uniqueness, and convergence towards the limiting distribution. Such a problem was previously studied by Alfonsi *et al.* [1] for the case of Wishart processes, using a fine analysis of the Laplace transform. The case of matrix-valued Ornstein–Uhlenbeck processes driven by Lévy subordinators was investigated by Barndorff-Nielsen and Stelzer [7], where also some results on the support of the invariant distribution were obtained; see also [45] and [47]. Based on stochastic stability criteria of Meyn and Tweedie, convergence to the unique invariant distribution in total variation was then studied for a general class of Ornstein–Uhlenbeck processes in [41], while generalized Ornstein–Uhlenbeck processes in dimension one have been investigated in [8] and [37]. First results on exponential ergodicity applicable to a class of affine processes on finite-dimensional cones were recently obtained by Mayerhofer *et al.* [44].

In contrast, we study the existence of limit distributions for general conservative, subcritical affine processes on $\mathbb{S}_d^+$ satisfying an additional first-moment condition for the state-dependent jump measure and a log-moment condition for the state-independent jump measure, and therefore extend the aforementioned results of [1] and [7] to a general class of affine processes on $\mathbb{S}_d^+$. Our proofs are inspired by Sato and Yamazato [47] (for Ornstein–Uhlenbeck processes) and rely on some technical ideas taken from [30], where affine processes on the canonical state space $\mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$ were considered. However, compared to the canonical state space, the boundary of $\mathbb{S}_d^+$ is quite complicated and prevents us from directly applying the arguments in [30]. Additional details are explained in the next section. We would like to mention that our method to show stationarity could also be applied to affine processes on more general convex cones in the spirit of [13] and [44]; however, this would require the introduction of additional notation which would make our key arguments less transparent, so we postpone the detailed discussion of this extension to Section 9.

Finally, let us mention that the long-time behavior of affine processes has previously been studied in many different settings, either based on a detailed study of the characteristic function (see e.g. [46], [39], [20], [47], [34], [32], [36], [30]), by stochastic stability criteria due to Meyn and Tweedie ([4], [29], [44]), or by coupling techniques ([19], [18], [40]). One application of such study is towards the estimation of parameters for affine models. In the case of the Wishart process, the maximum-likelihood estimator for the drift parameter was recently studied by Alfonsi *et al.* [1]. As demonstrated in their article, ergodicity helps to derive strong consistency and asymptotic normality of the estimator.

This paper is organized as follows. In Section 2, we introduce $\mathbb{S}_d^+$-valued affine processes, then formulate and discuss our main results. Section 3 is dedicated to applications of our results to specific affine models often used in finance. The proofs of the main results are then given in Sections 4–8. Finally, in Section 9 we provide an extension of our stationarity result to conservative affine processes on general convex cones.

2. Main results

In terms of terminology, we mainly follow the coordinate-free notation used in Mayerhofer [42] and Keller-Ressel and Mayerhofer [33].

Let $d \geq 2$ and denote by $\mathbb{S}_d$ the space of symmetric $d \times d$ matrices equipped with the scalar product $\langle x, y \rangle = \operatorname{tr}(xy)$ and the Frobenius norm $\|x\| := \langle x, x \rangle^{1/2}$, where tr(.) denotes the trace of a matrix. We list in Appendix A some properties of the trace and its induced norm which are repeatedly used in the remainder of the article. Denote by $\mathbb{S}_d^+$ (resp. $\mathbb{S}_d^{++}$) the cone of symmetric and positive semidefinite (resp. positive definite) real $d \times d$ matrices. We write $x \preceq y$ if $y - x \in \mathbb{S}_d^+$ and $x \prec y$ if $y - x \in \mathbb{S}_d^{++}$ for the natural partial and strict order relations introduced respectively by the cones $\mathbb{S}_d^+$ and $\mathbb{S}_d^{++}$. Let $\mathcal{B}(\mathbb{S}_d^+ \backslash \{0\})$ be the Borel-σ-algebra on $\mathbb{S}_d^+ \backslash \{0\}$. An $\mathbb{S}_d^+$*-valued measure* η on $\mathbb{S}_d^+ \backslash \{0\}$ is a $d \times d$ matrix of signed measures on $\mathbb{S}_d^+ \backslash \{0\}$ such that $\eta(A) \in \mathbb{S}_d^+$ whenever $A \in \mathcal{B}(\mathbb{S}_d^+ \backslash \{0\})$ with $0 \notin \overline{A}$.

In the following we present the notion of admissible parameters first introduced in Cuchiero *et al.* [12, Definition 2.3]. Here we mainly follow the definition given in Mayerhofer [42, Definition 3.1], with a slightly stronger condition on the linear jump coefficient.

Definition 2.1. Let $d \geq 2$. An admissible parameter set (α, b, B, m, μ) consists of the following:

(i) a linear diffusion coefficient $\alpha \in \mathbb{S}_d^+$;

(ii) a constant drift $b \in \mathbb{S}_d^+$ satisfying $b \succeq (d-1)\alpha$;

(iii) a constant jump term: a Borel measure m on $\mathbb{S}_d^+ \backslash \{0\}$ satisfying

$$\int_{\mathbb{S}_d^+ \backslash \{0\}} (\|\xi\| \wedge 1)\, m\, (\mathrm{d}\xi) < \infty;$$

(iv) a linear jump coefficient μ which is an $\mathbb{S}_d^+$-valued, sigma-finite measure on $\mathbb{S}_d^+ \backslash \{0\}$ satisfying

$$\int_{\mathbb{S}_d^+ \backslash \{0\}} \|\xi\| \operatorname{tr}(\mu)\, (\mathrm{d}\xi) < \infty, \tag{1}$$

where $\operatorname{tr}(\mu)$ denotes the measure induced by the relation $\operatorname{tr}(\mu)(A) := \operatorname{tr}(\mu(A))$ for all $A \in \mathcal{B}(\mathbb{S}_d^+ \backslash \{0\})$ with $0 \notin \bar{A}$;

(v) a linear drift B, which is a linear map $B : \mathbb{S}_d \to \mathbb{S}_d$ satisfying

$$\langle B(x), u \rangle \geq 0 \quad \text{for all } x,\ u \in \mathbb{S}_d^+ \text{ with } \langle x, u \rangle = 0.$$

According to our definition, a set of admissible parameters does not contain parameters corresponding to killing. In addition, compared with [12, Definition 2.3], our definition involves an additional first-moment assumption on the linear jump coefficient μ. Thanks to this stronger assumption and [12, Remark 2.5], the affine process we consider here is conservative.

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Theorem 2.1. (Cuchiero et al. [12].) *Let (α, b, B, m, μ) be admissible parameters in the sense of Definition 2.1. Then there exists a unique stochastically continuous transition kernel $p_t(x, \mathrm{d}\xi)$ such that $p_t(x, \mathbb{S}_d^+) = 1$ and*

$$\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,\xi\rangle} p_t(x, \mathrm{d}\xi) = \exp(-\phi(t, u) - \langle \psi(t, u), x\rangle), \qquad t \geq 0, \quad x, u \in \mathbb{S}_d^+, \tag{2}$$

where $\phi(t, u)$ and $\psi(t, u)$ in (2) are the unique solutions to the generalized Riccati differential equations; that is, for $u \in \mathbb{S}_d^+$,

$$\frac{\partial \phi(t, u)}{\partial t} = F(\psi(t, u)), \qquad \phi(0, u) = 0, \tag{3}$$

$$\frac{\partial \psi(t, u)}{\partial t} = R(\psi(t, u)), \qquad \psi(0, u) = u, \tag{4}$$

where the functions F and R are given by

$$F(u) = \langle b, u\rangle - \int_{\mathbb{S}_d^+\setminus\{0\}} (\mathrm{e}^{-\langle u,\xi\rangle} - 1) m(\mathrm{d}\xi),$$

$$R(u) = -2u\alpha u + B^\top(u) - \int_{\mathbb{S}_d^+\setminus\{0\}} (\mathrm{e}^{-\langle u,\xi\rangle} - 1)\mu(\mathrm{d}\xi).$$

Here, $B^\top$ denotes the adjoint operator on $\mathbb{S}_d$ defined by the relation $\langle u, B(\xi)\rangle = \langle B^\top(u), \xi\rangle$ for $u, \ \xi \in \mathbb{S}_d$. Under the additional moment condition (iv) of Definition 2.1, we will show in Lemma 4.2 below that $R(u)$ is continuously differentiable and thus locally Lipschitz continuous on $\mathbb{S}_d^+$. This fact, together with the absence of parameters according to killing, implies that the affine process under consideration is indeed conservative (see [12, Remark 2.5]) and, moreover, the Riccati equations have unique solutions for all $u \in \mathbb{S}_d^+$. Also, following [42], each affine process on $\mathbb{S}_d^+$, when $d \geq 2$, has jumps of finite variation, which is why we omitted, compared with [12], additional compensation in the integral against μ.

2.1. First moment

Our first result provides existence and a precise formula for the first moment of conservative affine processes on $\mathbb{S}_d^+$. For this purpose, we define the effective drift

$$\widetilde{B}(u) := B(u) + \int_{\mathbb{S}_d^+\setminus\{0\}} \xi\langle u, \mu(\mathrm{d}\xi)\rangle, \qquad \text{for } u \in \mathbb{S}_d. \tag{5}$$

Then note that $\widetilde{B}: \mathbb{S}_d \to \mathbb{S}_d$ is a linear map. We define the corresponding semigroup $(\exp(t\widetilde{B}))_{t\geq 0}$ by its Taylor series $\exp(t\widetilde{B})(u) = \sum_{n=0}^{\infty} t^n/n! \widetilde{B}^{\circ n}(u)$, where $\widetilde{B}^{\circ n}$ denotes the n-times composition of $\widetilde{B}$. For the remainder of the article we write $\mathbb{1}$ without an index for the $d \times d$-identity matrix, while $\mathbb{1}_A$ denotes the standard indicator function of a set A.

Theorem 2.2. *Let $p_t(x, \mathrm{d}\xi)$ be the transition kernel of an affine process on $\mathbb{S}_d^+$ with admissible parameters (α, b, B, m, μ) satisfying*

$$\int_{\{\|\xi\|>1\}} \|\xi\| m(\mathrm{d}\xi) < \infty. \tag{6}$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Then for each $t \geq 0$ *and* $x \in \mathbb{S}_d^+$*, the first moment of* $p_t(x, \mathrm{d}\xi)$ *exists and equals*

$$\int_{\mathbb{S}_d^+} \xi p_t(x, \mathrm{d}\xi) = \mathrm{e}^{t\widetilde{B}} x + \int_0^t \mathrm{e}^{s\widetilde{B}} \left(b + \int_{\mathbb{S}_d^+ \setminus \{0\}} \xi m(\mathrm{d}\xi) \right) \mathrm{d}s. \tag{7}$$

Although this result is not very surprising, the proof of our main result (Theorem 2.3) crucially relies on (7). For the sake of completeness we therefore provide a full and detailed proof of Theorem 2.2 in Section 4.

Based on methods of stochastic calculus, similar results were obtained for affine processes with state-space $\mathbb{R}^m_{\geq 0}$ in [5, Lemma 3.4] and on the canonical state space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$ in [19, Lemma 5.2]. For affine processes on $\mathbb{R}_{\geq 0}$, i.e., continuous-state branching processes with immigration, and also for the more general class of Dawson–Watanabe superprocesses, an alternative approach based on a fine analysis of the Laplace transform is provided in [39]. The latter approach clearly has the advantage that it is purely analytical and does not rely on the use of stochastic equations and semimartingale representations for these processes. Our proof provided in Section 4 is also purely analytical and uses some ideas taken from [39].

Remark 2.1. Note that the transition kernel $p_t(x, \cdot)$ with admissible parameters (α, b, B, m, μ) is Feller by virtue of [12, Theorem 2.4]. Therefore, there exists a canonical realization $(X, (\mathbb{P}_x)_{x \in \mathbb{S}_d^+})$ of the corresponding Markov process on the filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$, where $\Omega = \mathbb{D}(\mathbb{S}_d^+)$ is the set of all càdlàg paths $\omega : \mathbb{R}_{\geq 0} \to \mathbb{S}_d^+$, and $X_t(\omega) = \omega(t)$ for $\omega \in \Omega$. Here $(\mathcal{F}_t)_{t \geq 0}$ is the natural filtration generated by X, and $\mathcal{F} = \bigvee_{t \geq 0} \mathcal{F}_t$. For $x \in \mathbb{S}_d^+$, the probability measure $\mathbb{P}_x$ on Ω represents the law of the Markov process X given $X_0 = x$. With this notation, under the conditions of Theorem 2.2, the formula (7) reads

$$\mathbb{E}_x\big[X_t\big] = \mathrm{e}^{t\widetilde{B}} x + \int_0^t \mathrm{e}^{s\widetilde{B}} \left(b + \int_{\mathbb{S}_d^+ \setminus \{0\}} \xi m(\mathrm{d}\xi) \right) \mathrm{d}s,$$

where $\mathbb{E}_x$ denotes the expectation with respect to $\mathbb{P}_x$.

2.2. Existence and convergence to the invariant distribution

In this subsection we formulate our main result. Let $p_t(x, \cdot)$ be the transition kernel of an affine process on $\mathbb{S}_d^+$. Motivated by Theorem 2.2 it is reasonable to relate the long-time behavior of the process to the spectrum $\sigma(\widetilde{B})$ of $\widetilde{B}$. More precisely, an affine process on $\mathbb{S}_d^+$ with admissible parameters (α, b, B, m, μ) is said to be *subcritical* if

$$\sup \left\{ \mathrm{Re}\, \lambda \in \mathbb{C} : \lambda \in \sigma\left(\widetilde{B}\right) \right\} < 0. \tag{8}$$

Under the condition (8), it is well-known that there exist constants $M \geq 1$ and $\delta > 0$ such that

$$\left\| \mathrm{e}^{t\widetilde{B}} \right\| \leq M \mathrm{e}^{-\delta t}, \qquad t \geq 0. \tag{9}$$

The next remark provides a sufficient condition for (9).

Remark 2.2. According to [44, Theorem 2.7], (9) is satisfied if and only if there exists a $v \in \mathbb{S}_d^{++}$ such that $-\widetilde{B}^\top(v) \in \mathbb{S}_d^{++}$. However, in many application the linear drift is of the form $\widetilde{B}(x) = \beta x + x\beta^\top$, where β is a real-valued $d \times d$ matrix; see Section 3. In this case, it follows from [44, Corollary 5.1] that (9) is satisfied if and only if

$$\sup \{ \mathrm{Re}\, \lambda \in \mathbb{C} : \lambda \in \sigma\,(\beta) \} < 0,$$

which in turn holds true if and only if there exists a $v \in \mathbb{S}_d^{++}$ such that $-(\beta^\top v + v\beta) \in \mathbb{S}_d^{++}$.

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Let $\mathcal{P}(\mathbb{S}_d^+)$ be the space of all Borel probability measures on $\mathbb{S}_d^+$. We call $\pi \in \mathcal{P}(\mathbb{S}_d^+)$ an invariant distribution if

$$\int_{\mathbb{S}_d^+} p_t(x, \mathrm{d}\xi)\pi(\mathrm{d}x) = \pi(\mathrm{d}\xi), \qquad t \geq 0.$$

The following is our main result.

Theorem 2.3. *Let $p_t(x, \mathrm{d}\xi)$ be the transition kernel of a subcritical affine process on $\mathbb{S}_d^+$ with admissible parameters (α, b, B, m, μ). Suppose that the measure m satisfies*

$$\int_{\{\|\xi\|>1\}} \log \|\xi\| \, m(\mathrm{d}\xi) < \infty. \tag{10}$$

Then there exists a unique invariant distribution π. Moreover, for each $x \in \mathbb{S}_d^+$, one has $p_t(x, \cdot) \to \pi$ weakly as $t \to \infty$, and π has Laplace transform

$$\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,x\rangle} \pi(\mathrm{d}x) = \exp\left(-\int_0^\infty F(\psi(s, u))\, \mathrm{d}s\right), \qquad u \in \mathbb{S}_d^+. \tag{11}$$

The proof of Theorem 2.3 is postponed to Section 6. Let us make a few comments. Note that in dimension $d = 1$ it holds that $\mathbb{S}_1^+ = \mathbb{R}_{\geq 0}$, and affine processes on this state space coincide with the class of continuous-state branching processes with immigration introduced by Kawazu and Watanabe [31]. In this case, the long-time behavior has been extensively studied in the articles [36, Theorem 3.16] and [34, Theorem 2.6] and in the monograph [39, Theorem 3.20 and Corollary 3.21]. This is why we restrict ourselves to the case $d \geq 2$.

Theorem 2.3 establishes sufficient conditions for the existence, uniqueness, and convergence to the invariant distribution. For affine processes on the canonical state space $\mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$ a similar statement was recently shown in [30]. However, the method of [30] cannot be applied to the state space $\mathbb{S}_d^+$. The reason is as follows. To apply the arguments of [30] to $\mathbb{S}_d^+$-valued affine processes requires the decomposition

$$p_t(x, \cdot) = r_t(x, \cdot) * p_t(0, \cdot), \tag{12}$$

where $r_t(x, \cdot)$ is the transition kernel of an affine process on $\mathbb{S}_d^+$ whose Laplace transform is given by

$$\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,\xi\rangle} r_t(x, \mathrm{d}\xi) = \exp(-\langle \psi(t, u), x\rangle)\,; \tag{13}$$

that is, $r_t(x, \cdot)$ should have admissible parameters $(\alpha, b = 0, B, m = 0, \mu)$. Unfortunately, such a transition kernel $r_t(x, \cdot)$ is well-defined if and only if $(\alpha, b = 0, B, m = 0, \mu)$ are admissible parameters in the sense of Definition 2.1. This in turn is true if and only if $\alpha = 0$, which is a consequence of the particular structure of the boundary $\partial\mathbb{S}_d^+$. To overcome this difficulty, we use the cone structure to show that $\psi(t, u) \preceq \exp\{t\widetilde{B}^\top\}(u)$, which requires us to study the first moment of the affine process with admissible parameters $(\alpha, b, B, m = 0, \mu)$; see (29). Note that the first moment of affine processes is already studied in Theorem 2.2, where the first moment condition for μ is also explicitly used. Uniform exponential stability for $\psi(t, u)$ is then an immediate consequence of the assumption that the affine process is subcritical, i.e. that (9) holds. Compared with [30], our proof does not rely on stability results for ordinary differential equations; moreover, it has the advantage that it also applies to more general convex cones other than $\mathbb{S}_d^+$ (see Section 9). On the other hand, since this argument is crucially based on the proper closed convex cone structure, it cannot be applied to affine processes on the canonical state space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$.

Based on strong solutions to the stochastic equation and the construction of monotone couplings, an alternative approach for the study of the long-time behavior of affine processes on the canonical state space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$ was recently given in [19]. This approach certainly does not apply here since it is still an open problem whether each affine process on $\mathbb{S}_d^+$ can be obtained as a strong solution to a certain stochastic equation driven by Brownian motions and Poisson random measures (it is actually not even known what such a stochastic equation in the general case should look like). We refer the reader to [43] for some related results. In addition, we do not know if a comparison principle for such processes would be available.

For dimension $d = 1$ it is known that (10) is not only sufficient, but also necessary for the convergence to some limiting distribution; see, e.g., [39, Theorem 3.20 and Corollary 3.21]. To our knowledge, extensions of this result to higher-dimensional state spaces have not yet been obtained. In this context, we have the following partial result for subcritical affine processes on $\mathbb{S}_d^+$.

Proposition 2.1. *Let* $p_t(x, \mathrm{d}\xi)$ *be the transition kernel of a subcritical affine process on* $\mathbb{S}_d^+$ *with admissible parameters* (α, b, B, m, μ)*. Suppose that there exist* $x \in \mathbb{S}_d^+$ *and* $\pi \in \mathcal{P}(\mathbb{S}_d^+)$ *such that* $p_t(x, \cdot) \to \pi$ *weakly as* $t \to \infty$*. If* $\alpha = 0$ *and if there exists a constant* $K > 0$ *satisfying*

$$K\xi + B(\xi) \succeq 0, \quad \xi \in \mathbb{S}_d^+, \tag{14}$$

then (10) holds.

In order to prove Proposition 2.1 we first establish in Section 5 a precise lower bound for $\psi(t, u)$. Since in dimension $d \geq 2$ different components of the process interact through the drift B in a nontrivial manner on $\mathbb{S}_d^+$, the proof of the lower bound is deduced from the additional conditions $\alpha = 0$ and (14), which guarantee that these components are coupled in a *well-behaved way*. Let us remark that, in contrast to the proof of Theorem 2.3, the proof of Proposition 2.1 does not explicitly use the first moment condition for μ. Indeed, it suffices to assume that μ satisfies the condition $\int_{\mathbb{S}_d^+ \setminus \{0\}} (1 \wedge \|\xi\|) \mathrm{tr}(\mu)(\mathrm{d}\xi) < \infty$, which is weaker than (1), and, in addition, that the affine process is conservative.

We note that any linear map $B : \mathbb{S}_d \to \mathbb{S}_d$ with $B(\mathbb{S}_d^+) \subset \mathbb{S}_d^+$ satisfies the condition (14) for each $K > 0$. As an example of such a map, let $B(x) = \beta x \beta^\top$ for $x \in \mathbb{S}_d$, where β is a real-valued $d \times d$ matrix. Obviously, B defined in this way is admissible in the sense of Definition 2, and $B(\mathbb{S}_d^+) \subset \mathbb{S}_d^+$. However, such an example will not lead to a subcritical affine process. Just to understand this point, we can think of the simplest case when $d = 1$ (although we always assume $d \geq 2$ in the other parts of the paper): $B(x) = \beta^2 x$ and it is impossible for B to fulfill the subcritical condition (8).

The simplest example of an admissible B that produces subcriticality and at the same time satisfies (14) is $B(x) = -x$, $x \in S_d$. However, there are also other examples like this one. For instance, consider $d = 2$ and the linear map B on S_2 defined by

$$B\left(\begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix}\right) = \begin{pmatrix} -x_1 & -\kappa x_2 \\ -\kappa x_2 & -x_3 \end{pmatrix},$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

where $\kappa > 1$ is a constant. First, such a B is admissible in the sense of Definition 2.1, since for $u, x \in S_2^+$ with $\langle x, u\rangle = 0$,

$$\langle B(x), u\rangle = -\kappa\,\langle x, u\rangle + (\kappa - 1)\left\langle \begin{pmatrix} x_1 & 0 \\ 0 & x_3 \end{pmatrix}, u \right\rangle = (\kappa - 1)\left\langle \begin{pmatrix} x_1 & 0 \\ 0 & x_3 \end{pmatrix}, u \right\rangle \geq 0.$$

It is easy to see that the spectrum $\sigma(B) = \{-1, -\kappa\}$, so that we can choose a relatively small μ to make $\tilde{B}$ defined in (5) subcritical; moreover, (14) holds for $K = \kappa$.

We close this section with a useful moment result regarding the invariant distribution.

Corollary 2.1. *Let $p_t(x, \mathrm{d}\xi)$ be the transition kernel of a subcritical affine process on $\mathbb{S}_d^+$ with admissible parameters (α, b, B, m, μ) satisfying (6). Let π be the unique invariant distribution. Then*

$$\lim_{t\to\infty} \int_{\mathbb{S}_d^+} y p_t(x, \mathrm{d}y) = \int_{\mathbb{S}_d^+} y\pi(\mathrm{d}y) = -\widetilde{B}^{-1}\left(b + \int_{\mathbb{S}_d^+\setminus\{0\}} \xi m(\mathrm{d}\xi)\right).$$

2.3. Study of the convergence rate

Consider a subcritical affine process on $\mathbb{S}_d^+$ with admissible parameters (α, b, B, m, μ) and let δ be defined by (9). In particular, δ is strictly positive, and we will see that it appears naturally in the convergence rate towards the invariant distribution. In order to measure this rate of convergence we introduce

$$d_L(\eta, \nu) := \sup_{u\in\mathbb{S}_d^+\setminus\{0\}} \frac{1}{\|u\|}\left|\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,x\rangle}\eta(\mathrm{d}x) - \int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,x\rangle}\nu(\mathrm{d}x)\right|, \quad \eta,\ \nu \in \mathcal{P}(\mathbb{S}_d^+).$$

Note that this supremum is not necessarily finite. However, it is finite for elements from

$$\mathcal{P}_1(\mathbb{S}_d^+) = \left\{\varrho \in \mathcal{P}(\mathbb{S}_d^+) : \int_{\mathbb{S}_d^+} \|x\|\varrho(\mathrm{d}x) < \infty\right\},$$

which essentially follows from

$$\begin{aligned}\left|\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,x\rangle}\eta(\mathrm{d}x) - \int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,x\rangle}\nu(\mathrm{d}x)\right| &= \left|\int_{\mathbb{S}_d^+\times\mathbb{S}_d^+} \left(\mathrm{e}^{-\langle u,x\rangle} - \mathrm{e}^{-\langle u,y\rangle}\right)\eta(\mathrm{d}x)\nu(\mathrm{d}y)\right| \\ &\leq \|u\| \int_{\mathbb{S}_d^+\times\mathbb{S}_d^+} \|x-y\|\eta(\mathrm{d}x)\nu(\mathrm{d}y) \\ &\leq \|u\|\left(\int_{\mathbb{S}_d^+}\|x\|\eta(\mathrm{d}x) + \int_{\mathbb{S}_d^+}\|y\|\nu(\mathrm{d}y)\right).\end{aligned} \tag{15}$$

It is easy to see that d_L is a metric on $\mathcal{P}_1(\mathbb{S}_d^+)$; moreover, $\left(\mathcal{P}_1(\mathbb{S}_d^+), d_L\right)$ is complete. Using well-known properties of Laplace transforms, it can be shown that convergence with respect to d_L implies weak convergence. The next result provides an exponential rate in d_L distance.

Theorem 2.4. *Let $p_t(x, \mathrm{d}\xi)$ be the transition kernel of a subcritical affine process on $\mathbb{S}_d^+$ with admissible parameters (α, b, B, m, μ). Suppose that (10) holds and denote by π the unique invariant distribution. Then there exists a constant $C > 0$ such that*

$$d_L\left(p_t(x, \cdot), \pi\right) \leq C\,(1 + \|x\|)\,\mathrm{e}^{-\delta t}, \quad t \geq 0, \quad x \in \mathbb{S}_d^+. \tag{16}$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

The proof of this result is given in Section 7. Although under the given conditions $p_t(x, \cdot)$ and π do not necessarily belong to $\mathcal{P}_1(\mathbb{S}_d^+)$, the proof of (16) reveals that $d_L(p_t(x, \cdot), \pi)$ is well-defined. Let us mention that the main purpose of this work is dedicated to Theorem 2.3. The above convergence rate in the d_L distance is a simple byproduct of the estimates derived in the process of proving Theorem 2.3.

We turn to investigate the convergence rate from the affine transition kernel to the invariant distribution in the Wasserstein-1-distance introduced below. Given $\varrho, \widetilde{\varrho} \in \mathcal{P}_1(\mathbb{S}_d^+)$, a coupling H of $(\varrho, \widetilde{\varrho})$ is a Borel probability measure on $\mathbb{S}_d^+ \times \mathbb{S}_d^+$ which has marginals ϱ and $\widetilde{\varrho}$, respectively. We denote by $\mathcal{H}(\varrho, \widetilde{\varrho})$ the collection of all such couplings. We define the *Wasserstein distance* on $\mathcal{P}_1(\mathbb{S}_d^+)$ by

$$W_1(\varrho, \widetilde{\varrho}) = \inf \left\{ \int_{\mathbb{S}_d^+ \times \mathbb{S}_d^+} \|x - y\| H(\mathrm{d}x, \mathrm{d}y) \ : H \in \mathcal{H}(\varrho, \widetilde{\varrho}) \right\}.$$

Since ϱ and $\widetilde{\varrho}$ belong to $\mathcal{P}_1(\mathbb{S}_d^+)$, it holds that $W_1(\varrho, \widetilde{\varrho})$ is finite. According to [48, Theorem 6.16], we have that $(\mathcal{P}_1(\mathbb{S}_d^+), W_1)$ is a complete separable metric space. Note that, by an argument similar to (15), one easily finds that $d_L(\eta, \nu) \le W_1(\eta, \nu)$ for all $\eta, \nu \in \mathcal{P}_1(\mathbb{S}_d^+)$. Exponential ergodicity in different Wasserstein distances for affine processes on the canonical state space $\mathbb{R}_{\ge 0}^m \times \mathbb{R}^n$ was very recently studied in [19]. Below we provide a corresponding result for affine processes on $\mathbb{S}_d^+$.

Theorem 2.5. *Let* $p_t(x, \mathrm{d}\xi)$ *be the transition kernel of a subcritical affine process on* $\mathbb{S}_d^+$ *with admissible parameters* (α, b, B, m, μ) *satisfying (6). If* $\alpha = 0$*, then*

$$W_1(p_t(x, \cdot), \pi) \le \sqrt{d} M e^{-\delta t} \left(\|x\| + \int_{\mathbb{S}_d^+} \|y\| \pi(\mathrm{d}y) \right), \quad t \ge 0, \quad x \in \mathbb{S}_d^+. \tag{17}$$

The proof of Theorem 2.5, which is given in Section 8, largely follows some ideas of [19]. In contrast to the latter work, for the study of affine processes on $\mathbb{S}_d^+$ we encounter similar difficulties as in the proof of Theorem 2.3. Namely, the affine property allows us only to use the convolution argument (12) when $\alpha = 0$. Moreover, since it is still an open problem whether each affine process on $\mathbb{S}_d^+$ can be obtained as a strong solution to a certain stochastic equation driven by Brownian motions and Poisson random measures, we are not able to improve our Theorem 2.5 to other variants of Wasserstein distances as used in [19].

3. Applications and examples

Let $(W_t)_{t\ge 0}$ be a $d \times d$ matrix of independent standard Brownian motions. Denote by $(J_t)_{t\ge 0}$ an $\mathbb{S}_d^+$-valued Lévy subordinator with Lévy measure m. Suppose that these two processes are independent of each other. Following [43], the stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = \left(b + \beta X_t + X_t \beta^\top\right) \mathrm{d}t + \sqrt{X_t} \mathrm{d}W_t \Sigma + \Sigma^\top \mathrm{d}W_t^\top \sqrt{X_t} + \mathrm{d}J_t, \quad t \ge 0, \\ X_0 = x \in \mathbb{S}_d^+, \end{cases} \tag{18}$$

has a unique weak solution if $b \succeq (d-1)\Sigma^\top \Sigma$ and Σ, β are real-valued $d \times d$ matrices. Moreover, according to [43, Corollary 3.2], if $b \succeq (d+1)\Sigma^\top \Sigma$ and $x \in \mathbb{S}_d^{++}$, then there also

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

exists an $\mathbb{S}_d^{++}$-valued strong solution. The corresponding Markov process $X=(X_t)_{t\geq 0}$ is a conservative affine process with admissible parameters $(\alpha, b, B, m, 0)$ with diffusion $\alpha=\Sigma^\top\Sigma$ and linear drift $B(x)=\beta x+x\beta^\top$. The functions F and R are given by

$$F(u)=\langle b, u\rangle+\int_{\mathbb{S}_d^+\backslash\{0\}}(1-\mathrm{e}^{-\langle u,\xi\rangle})\, m(\mathrm{d}\xi)$$

and

$$R(u)=-2u\alpha u+u\beta+\beta^\top u.$$

The generalized Riccati equations are now given by

$$\partial_t\phi(t, u)=\langle b, \psi(t, u)\rangle+\int_{\mathbb{S}_d^+\backslash\{0\}}(1-\mathrm{e}^{-\langle\psi(t,u),\xi\rangle})\, m(\mathrm{d}\xi),$$

$$\partial_t\psi(t, u)=-2\psi(t, u)\alpha\psi(t, u)+\psi(t, u)\beta+\beta^\top\psi(t, u),$$

with initial conditions $\phi(0, u)=0$ and $\psi(0, u)=u$. Let $\sigma_t^\beta:\mathbb{S}_d^+\to\mathbb{S}_d^+$ be given by

$$\sigma_t^\beta(x):=2\int_0^t \mathrm{e}^{\beta s}x\mathrm{e}^{\beta^\top s}\mathrm{d}s, \qquad t\geq 0.$$

According to [42, Section 4.3], we have

$$\phi(t, u)=\langle b, \int_0^t\psi(s, u)\mathrm{d}s\rangle+\int_0^t\int_{\mathbb{S}_d^+\backslash\{0\}}(1-\mathrm{e}^{-\langle\psi(s,u),\xi\rangle})\, m(\mathrm{d}\xi)\mathrm{d}s,$$

$$\psi(t, u)=\mathrm{e}^{\beta^\top t}(u^{-1}+\sigma_t^\beta(\alpha))^{-1}\mathrm{e}^{\beta t}.$$

Since $\widetilde{B}(x)=B(x)$, Remark 2.2 implies that X is subcritical, provided β has only eigenvalues with negative real parts. If the Lévy measure m satisfies (10), then Theorem 2.3 implies existence, uniqueness, and convergence to the invariant distribution π whose Laplace transform satisfies

$$\int_0^\infty \mathrm{e}^{-\langle u,x\rangle}\pi\,(\mathrm{d}x)=\langle b, \int_0^\infty\psi(s, u)\mathrm{d}s\rangle\exp\left(-\int_0^\infty\int_{\mathbb{S}_d^+\backslash\{0\}}(1-\mathrm{e}^{-\langle\psi(s,u),\xi\rangle})\, m(\mathrm{d}\xi)\mathrm{d}s\right).$$

Moreover, if in addition $\int_{\{\|\xi\|\geq 1\}}\|\xi\| m(\mathrm{d}\xi)<\infty$, then we infer from Corollary 2.1 that

$$\lim_{t\to\infty}\mathbb{E}_x[X_t]=\int_0^\infty \mathrm{e}^{s\beta^\top}\left(b+\int_{\mathbb{S}_d^+\backslash\{0\}}\xi m(\mathrm{d}\xi)\right)\mathrm{e}^{s\beta}\mathrm{d}s=\int_{\mathbb{S}_d^+}y\pi(\mathrm{d}y).$$

We end this section by considering the following examples.

Example 3.1. *(The matrix-variate basic affine jump-diffusion and Wishart process.)* Take $b=2k\Sigma^\top\Sigma$ with $k\geq d-1$ in (18). This process is called matrix-variate basic affine jump-diffusion on $\mathbb{S}_d^+$ (MBAJD for short); see [42, Section 4]. Following [42, Section 4.3], $\phi(t, u)$ is precisely given by

$$\phi(t, u)=k\log\det\left(\mathbb{1}+u\sigma_t^\beta(\alpha)\right)\int_0^t\int_{\mathbb{S}_d^+\backslash\{0\}}\left(1-\mathrm{e}^{-\langle\psi(s,u),\xi\rangle}\right)m(\mathrm{d}\xi)\mathrm{d}s,$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

and Theorem 2.3 implies that the unique invariant distribution is given by

$$\int_0^\infty e^{-\langle u,x\rangle}\pi\,(dx) = \Big(\det\big(\mathbb{1} + \sigma_\infty^\beta(\alpha)u\big)\Big)^{-k} \exp\bigg(-\int_0^\infty \int_{\mathbb{S}_d^+\setminus\{0\}} (1-e^{-\langle\psi(s,u),\xi\rangle})\, m(d\xi)ds\bigg),$$

where $\sigma_\infty^\beta(\alpha) = \int_0^\infty \exp(s\beta)\alpha\exp(s\beta^\top)ds$.

The well-known Wishart process, introduced by Bru [10], is a special case of the MBAJD with $m=0$. Existence of a unique distribution was then obtained in [1, Lemma C.1]. In this case π is a Wishart distribution with shape parameter k and scale parameter $\sigma_\infty^\beta(\alpha)$.

Example 3.2. *(Matrix-variate Ornstein–Uhlenbeck-type processes.)* For $b=0$ and $\Sigma=0$, we call the solutions to the stochastic differential equation (18) matrix-variate Ornstein–Uhlenbeck-type (OU-type) processes; see [7]. Properties of the stationary matrix-variate OU-type processes were investigated in [45]. Provided that

$$\int_{\{\|\xi\|\geq 1\}} \|\xi\| m(d\xi) < \infty,$$

Theorem 2.5 implies that the matrix-variate OU-type process is also exponentially ergodic in the Wasserstein-1-distance.

Note that in [41] ergodicity in total variation was studied for matrix-variate OU-type processes. The proof crucially relies on particular properties of Ornstein–Uhlenbeck processes and hence cannot be extended to cases where the parameter μ for state-dependent jumps does not vanish. In contrast, our method would also apply for non-vanishing μ.

4. Proof of Theorem 2.2

In this section we study the first moment of a conservative affine process on $\mathbb{S}_d^+$. In particular, we prove Theorem 2.2. Essential to the proof is the space-differentiability of the functions F and R as well as of ϕ and ψ. To simplify the notation we introduce $L(\mathbb{S}_d, \mathbb{S}_d)$ as the space of all linear operators $\mathbb{S}_d \to \mathbb{S}_d$, and similarly $L(\mathbb{S}_d, \mathbb{R})$ stands for the space of all linear functionals $\mathbb{S}_d \to \mathbb{R}$. For a function $G : \mathbb{S}_d \to \mathbb{S}_d$ we denote its derivative at $u \in \mathbb{S}_d$, if it exists, by $DG(u) \in L(\mathbb{S}_d, \mathbb{S}_d)$. Similarly, we denote the derivative of $H : \mathbb{S}_d \to \mathbb{R}$ by $DH(u) \in L(\mathbb{S}_d, \mathbb{R})$. We equip $L(\mathbb{S}_d, \mathbb{S}_d)$ and $L(\mathbb{S}_d, \mathbb{R})$ with the corresponding norms

$$\|DG(u)\| = \sup_{\|x\|=1} \|DG(u)(x)\| \quad \text{and} \quad \|DH(u)\| = \sup_{\|x\|=1} \|DH(u)(x)\|\,.$$

Let F and R be as in Theorem 2.1 According to [12, Lemma 5.1] the function R is analytic on $\mathbb{S}_d^{++}$. Below we study the differentiability of F and R on the entire cone $\mathbb{S}_d^+$.

We first give a lemma that slightly extends [42, Lemma 3.3].

Lemma 4.1. *Let g be a measurable function on $\mathbb{S}_d^+$ with $\int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\,\mathrm{tr}(\mu)(d\xi) < \infty$. Then $\int_{\mathbb{S}_d^+\setminus\{0\}} g(\xi)\mu(d\xi)$ is finite and*

$$\left\| \int_{\mathbb{S}_d^+\setminus\{0\}} g(\xi)\mu(d\xi) \right\| \leq \int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\,\mathrm{tr}(\mu)(d\xi).$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Proof. Let $\mu = (\mu_{ij})$ and $\mu_{ij} = \mu_{ij}^+ - \mu_{ij}^-$ be the Jordan decomposition of μ_{ij}. Suppose $\int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)| \operatorname{tr}(\mu)(\mathrm{d}\xi) < \infty$. Then [42, Lemma 3.3] implies that $\int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\mu(\mathrm{d}\xi)$ is finite and

$$\left\| \int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\mu(\mathrm{d}\xi) \right\| \le \int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)| \operatorname{tr}(\mu)(\mathrm{d}\xi).$$

Since the ijth entry of $\int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\mu(\mathrm{d}\xi)$ is given by

$$\int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\mu_{ij}^+(\mathrm{d}\xi) - \int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\mu_{ij}^-(\mathrm{d}\xi),$$

which is finite, we must have

$$\int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\mu_{ij}^+(\mathrm{d}\xi) < \infty \qquad \text{and} \qquad \int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)|\mu_{ij}^-(\mathrm{d}\xi) < \infty, \quad \forall i, j \in \{1, \ldots d\}.$$

So $\int_{\mathbb{S}_d^+\setminus\{0\}} g(\xi)\mu(\mathrm{d}\xi)$ is finite. Again by [42, Lemma 3.3],

$$\begin{aligned}
\left\| \int_{\mathbb{S}_d^+\setminus\{0\}} g(\xi)\mu(\mathrm{d}\xi) \right\| &= \left\| \int_{\mathbb{S}_d^+\setminus\{0\}} g^+(\xi)\mu(\mathrm{d}\xi) - \int_{\mathbb{S}_d^+\setminus\{0\}} g^-(\xi)\mu(\mathrm{d}\xi) \right\| \\
&\le \left\| \int_{\mathbb{S}_d^+\setminus\{0\}} g^+(\xi)\mu(\mathrm{d}\xi) \right\| + \left\| \int_{\mathbb{S}_d^+\setminus\{0\}} g^-(\xi)\mu(\mathrm{d}\xi) \right\| \\
&\le \int_{\mathbb{S}_d^+\setminus\{0\}} g^+(\xi)\operatorname{tr}(\mu)(\mathrm{d}\xi) + \int_{\mathbb{S}_d^+\setminus\{0\}} g^-(\xi)\operatorname{tr}(\mu)(\mathrm{d}\xi) \\
&\le \int_{\mathbb{S}_d^+\setminus\{0\}} |g(\xi)| \operatorname{tr}(\mu)(\mathrm{d}\xi).
\end{aligned}$$

The lemma is proved.

Lemma 4.2. *The following statements hold:*

(a) *For $u \in \mathbb{S}_d^{++}$, $h \in \mathbb{S}_d$, we have*

$$DR(u)(h) = -2\,(u\alpha h + h\alpha u) + B^\top(h) + \int_{\mathbb{S}_d^+\setminus\{0\}} \langle h, \xi\rangle \mathrm{e}^{-\langle u, \xi\rangle}\mu(\mathrm{d}\xi). \tag{19}$$

Moreover, through (19), $DR(u)$ is continuously extended to $u \in \mathbb{S}_d^+$. In particular, $R \in C^1(\mathbb{S}_d^+)$ and (19) holds true for all $u \in \mathbb{S}_d^+$, $h \in \mathbb{S}_d$.

(b) *If (6) is satisfied, then for $u \in \mathbb{S}_d^{++}$, $h \in \mathbb{S}_d$,*

$$DF(u)(h) = \langle b, h\rangle + \int_{\mathbb{S}_d^+\setminus\{0\}} \langle h, \xi\rangle \mathrm{e}^{-\langle u, \xi\rangle} m(\mathrm{d}\xi). \tag{20}$$

Moreover, through (20), $DF(u)$ is continuously extended to $u \in \mathbb{S}_d^+$. In particular, $F \in C^1(\mathbb{S}_d^+)$ and (20) holds true for all $u \in \mathbb{S}_d^+$, $h \in \mathbb{S}_d$.

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Proof. (a) Let $u \in \mathbb{S}_d^{++}$. Consider $h \in \mathbb{S}_d$ with sufficiently small $\|h\|$ such that $u + h \in \mathbb{S}_d^+$. An easy calculation shows that

$$R(u+h) - R(u) = DR(u)(h) + r(u, h),$$

where

$$r(u, h) := -2h\alpha h + \int_{\mathbb{S}_d^+\setminus\{0\}} \mathrm{e}^{-\langle u,\xi\rangle} \left(1 - \mathrm{e}^{-\langle h,\xi\rangle} - \langle h, \xi\rangle\right) \mu(\mathrm{d}\xi).$$

Let us prove that $\lim_{0\neq\|h\|\to 0} \|r(u, h)\|/\|h\| = 0$. Assume $\|h\| \neq 0$. First, note that

$$\frac{\|2h\alpha h\|}{\|h\|} \le 2\|\alpha\| \frac{\|h\|^2}{\|h\|} \le 2\|\alpha\|\,\|h\|.$$

Let $M > 0$. For $\|\xi\| \le M$, we have

$$\begin{aligned}
\left|\mathrm{e}^{-\langle u,\xi\rangle} \left(1 - \mathrm{e}^{-\langle h,\xi\rangle} - \langle h, \xi\rangle\right)\right| &= \left|\langle h, \xi\rangle \left(\int_0^1 \mathrm{e}^{-\langle u+sh,\xi\rangle}\mathrm{d}s - \mathrm{e}^{-\langle u,\xi\rangle}\right)\right| \\
&= |\langle h, \xi\rangle| \cdot \left|\int_0^1 \left(\mathrm{e}^{-\langle u+sh,\xi\rangle} - \mathrm{e}^{-\langle u,\xi\rangle}\right) \mathrm{d}s\right| \\
&\le |\langle h, \xi\rangle|^2 , \qquad (21)
\end{aligned}$$

where we used the fact that $\langle u + sh, \xi\rangle \ge 0$ and the Lipschitz continuity of $[0, \infty) \in x \mapsto \exp(-x)$ to get the last inequality. Similarly, for $\|\xi\| > M$,

$$\left|\mathrm{e}^{-\langle u,\xi\rangle} \left(1 - \mathrm{e}^{-\langle h,\xi\rangle} - \langle h, \xi\rangle\right)\right| \le \left|\mathrm{e}^{-\langle u,\xi\rangle} - \mathrm{e}^{-\langle u+h,\xi\rangle}\right| + \left|\mathrm{e}^{-\langle u,\xi\rangle}\langle h, \xi\rangle\right| \le 2\,|\langle h, \xi\rangle| . \qquad (22)$$

Combining (21) and (22) and applying Lemma 4.1, we get

$$\begin{aligned}
&\frac{1}{\|h\|} \left\|\int_{\mathbb{S}_d^+\setminus\{0\}} \mathrm{e}^{-\langle u,\xi\rangle} \left(1 - \mathrm{e}^{-\langle h,\xi\rangle} - \langle h, \xi\rangle\right) \mu(\mathrm{d}\xi)\right\| \\
&\qquad \le \frac{1}{\|h\|} \int_{\mathbb{S}_d^+\setminus\{0\}} \left|\mathrm{e}^{-\langle u,\xi\rangle} \left(1 - \mathrm{e}^{-\langle h,\xi\rangle} - \langle h, \xi\rangle\right)\right| \mathrm{tr}(\mu)(\mathrm{d}\xi) \\
&\qquad \le \|h\| \int_{\{\|\xi\|\le M\}} \|\xi\|^2 \mathrm{tr}(\mu)(\mathrm{d}\xi) + 2 \int_{\{\|\xi\|>M\}} \|\xi\| \mathrm{tr}(\mu)(\mathrm{d}\xi).
\end{aligned}$$

So

$$\frac{\|r(u, h)\|}{\|h\|} \le \left(2\|\alpha\| + \int_{\{\|\xi\|\le M\}} \|\xi\|^2 \mathrm{tr}(\mu)(\mathrm{d}\xi)\right) \|h\| + 2 \int_{\{\|\xi\|>M\}} \|\xi\| \mathrm{tr}(\mu)(\mathrm{d}\xi).$$

Note that $\int_{\mathbb{S}_d^+\setminus\{0\}} \|\xi\| \mathrm{tr}(\mu)(\mathrm{d}\xi) < \infty$ by virtue of Definition 2.1(iv). Let $\varepsilon > 0$ be arbitrary and fix some $M = M(\varepsilon) > 0$ large enough so that $\int_{\{\|\xi\|>M\}} \|\xi\| \mathrm{tr}(\mu)(\mathrm{d}\xi) < \varepsilon/4$. Define

$$\delta = \delta(\varepsilon) := \left(1 + 2\|\alpha\| + \int_{\{\|\xi\|\le M\}} \|\xi\|^2 \mathrm{tr}(\mu)(\mathrm{d}\xi)\right)^{-1} \frac{\varepsilon}{2}.$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Then, for $\|h\| \leq \delta$, we see that

$$\frac{\|r(u,h)\|}{\|h\|} \leq \left(2\|\alpha\| + \int_{\{\|\xi\|\leq M\}} \|\xi\|^2 \mathrm{tr}(\mu)(\mathrm{d}\xi)\right)\delta + \frac{\varepsilon}{2} \leq \varepsilon.$$

This proves (19) for $u \in \mathbb{S}_d^{++}$. Finally, the continuity of $u \mapsto DR(u)$ in $\mathbb{S}_d^+$ can be easily obtained from the dominated convergence theorem.

(b) Similarly as before, we derive $F(u+h) - F(u) = DF(u)(h) + r(u,h)$ with

$$r(u,h) := \int_{\mathbb{S}_d^+\setminus\{0\}} \exp\left(-\langle u, \xi\rangle\right)(1 - \exp\left(\langle h, \xi\rangle\right) - \langle h, \xi\rangle)m(\mathrm{d}\xi).$$

Let $\|h\| \neq 0$. By essentially the same reasoning as in (a), we obtain that

$$\frac{\|r(u,h)\|}{\|h\|} \leq \|h\| \int_{\{\|\xi\|\leq M\}} \|\xi\|^2 m(\mathrm{d}\xi) + 2\int_{\{\|\xi\|>M\}} \|\xi\| m(\mathrm{d}\xi),$$

and the second integral on the right-hand side is now finite by (6). Hence, we may follow the same steps as in (a) to see that $\|r(u,h)\|/\|h\| \to 0$ as $\|h\| \to 0$ and to obtain the continuity of $DF(u)$ in $\mathbb{S}_d^+$.

Let ϕ and ψ be as in Theorem 2.1. We know from [12, Lemma 3.2 (iii)] that $\phi(t,u)$ and $\psi(t,u)$ are jointly continuous on $\mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$ and, moreover, $u \mapsto \phi(t,u)$ and $u \mapsto \psi(t,u)$ are analytic on $\mathbb{S}_d^{++}$ for $t \geq 0$.

Proposition 4.1. *The following statements hold:*

(a) *$D\psi$ has a jointly continuous extension on $\mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$.*

(b) *If (6) is satisfied, then $D\phi$ has a jointly continuous extension on $\mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$.*

Proof. (a) Noting that $s \mapsto DR(\psi(s,u)) \in L(\mathbb{S}_d, \mathbb{S}_d)$ is continuous, we may define $f_u(t)$ as the unique solution in $L(\mathbb{S}_d, \mathbb{S}_d)$ to

$$f_u(t) = \mathbb{1} + \int_0^t DR\left(\psi(s,u)\right) f_u(s)\mathrm{d}s.$$

Further, we then define the extension of $D\psi$ onto $\mathbb{R}_{\geq 0} \times \partial\mathbb{S}_d^+$ simply by

$$D\psi(t,u) = f_u(t), \qquad (t,u) \in \mathbb{R}_{\geq 0} \times \partial\mathbb{S}_d^+.$$

It remains to verify the joint continuity of $D\psi(t,u)$ on $\mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$ extended in this way. By the Riccati differential equation (4) we have

$$D\psi(t,u) = \mathbb{1} + \int_0^t DR\left(\psi(s,u)\right) D\psi(s,u)\mathrm{d}s, \qquad t \geq 0,\ u \in \mathbb{S}_d^+.$$

Using that $u \mapsto R(u)$ is continuous on $\mathbb{S}_d^+$ and ψ is jointly continuous on $\mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$, we have that for all $T > 0$ and $M > 0$, there exists a constant $C(T,M) > 0$ such that

$$\sup_{s\in[0,T],\, u\in\mathbb{S}_d^+,\, \|u\|\leq M} \|DR\left(\psi(s,u)\right)\| =: C(T,M) < \infty.$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Hence, for each $u \in \mathbb{S}_d^+$ with $\|u\| \le M$, we obtain

$$\|D\psi(t, u)\| \le 1 + C(T, M) \int_0^t \|D\psi(s, u)\| \mathrm{d}s.$$

Applying Gronwall's inequality yields

$$\|D\psi(t, u)\| \le \mathrm{e}^{C(T,M)T} =: K(T, M) < \infty$$

for all $t \in [0, T]$ and $u \in \mathbb{S}_d^+$ with $\|u\| \le M$. Because $D\psi$ is jointly continuous in $\mathbb{R}_{\ge 0} \times \mathbb{S}_d^{++}$, it is enough to prove continuity at some fixed point $(t, u) \in \mathbb{R}_{\ge 0} \times \partial\mathbb{S}_d^+$, where $\partial\mathbb{S}_d^+ := \mathbb{S}_d^+ \backslash \mathbb{S}_d^{++}$.

Without loss of generality we assume $t \in [0, T]$ and $u \in \partial\mathbb{S}_d^+$ with $\|u\| \le M$. Let $s \in \mathbb{R}_{\ge 0}$ and $v \in \mathbb{S}_d^+$ with $s \in [0, T]$ and $\|v\| \le M$. We have

$$\|D\psi(t, u) - D\psi(s, v)\| \le \|D\psi(t, u) - D\psi(s, u)\| + \|D\psi(s, u) - D\psi(s, v)\| . \tag{23}$$

We estimate the first term on the right-hand side of (23) by

$$\begin{aligned}\|D\psi(t, u) - D\psi(s, u)\| &\le \left\| \int_0^t DR\left(\psi(r, u)\right) D\psi(r, u)\mathrm{d}r - \int_0^s DR\left(\psi(r, u)\right) D\psi(r, u)\mathrm{d}r \right\| \\ &\le C(T, M) \int_{[s,t]\cup[t,s]} \|D\psi(r, u)\| \, \mathrm{d}r \\ &\le C(T, M)K(T, M)|t - s|. \end{aligned} \tag{24}$$

Turning to the second term, for $v \in \mathbb{S}_d^{++}$ with $\|v\| \le M$, $D\psi(s, u) = f_u(s)$, and $D\psi(r, u) = f_u(r)$, we obtain

$$\begin{aligned}\|D\psi(s, u) - D\psi(s, v)\| &\le \int_0^s \|DR\left(\psi(r, u)\right) D\psi(r, u) - DR\left(\psi(r, v)\right) D\psi(r, v)\| \, \mathrm{d}r \\ &\le \int_0^s \|DR(\psi(r, u)) - DR\left(\psi(r, v)\right)\| \, \|D\psi(r, v)\| \, \mathrm{d}r \\ &\quad + \int_0^s \|DR\left(\psi(r, u)\right)\| \, \|D\psi(r, u) - D\psi(r, v)\| \, \mathrm{d}r \\ &\le K(T, M) \int_0^T \|DR(\psi(r, u)) - DR\left(\psi(r, v)\right)\| \, \mathrm{d}r \\ &\quad + C(T, M) \int_0^s \|D\psi(r, u) - D\psi(r, v)\| \, \mathrm{d}r \\ &= K(T, M)a_T(v, u) + C(T, M) \int_0^s \|D\psi(r, u) - D\psi(r, v)\| \, \mathrm{d}r, \end{aligned}$$

where $a_T(v, u) := \int_0^T \|DR(\psi(r, u)) - DR(\psi(r, v))\|\mathrm{d}r$. Using once again Gronwall's inequality, we deduce

$$\|D\psi(s, u) - D\psi(s, v)\| \le K(T, M)a_T(v, u)\mathrm{e}^{C(T,M)T}. \tag{25}$$

Noting that $R \in C^1(\mathbb{S}_d^+)$ and $\psi(r, 0) = 0$ by [12, Remark 2.5], by the dominated convergence theorem we see that $a_T(v, u)$ tends to zero as $v \to u$. Consequently, the right-hand side of (25)

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

tends to zero as $v \to u$. Combining (23) with (24) and (25), we conclude that $D\psi$ extended in this way is jointly continuous in $(t, u) \in \mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$.

(b) We know from the generalized Riccati equation (3) that $\phi(t, u) = \int_0^t F(\psi(s, u))\mathrm{d}s$. Noting that $F \in C^1(\mathbb{S}_d^+)$ thanks to (6), the chain rule combined with the dominated convergence theorem implies the assertion.

We are ready to prove Theorem 2.2.

Proof of Theorem 2.2. Let $\varepsilon > 0$. We have

$$\begin{aligned}\frac{\partial}{\partial\varepsilon}\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle \varepsilon u, \xi\rangle} p_t(x, \mathrm{d}\xi) = \frac{\partial}{\partial\varepsilon}\mathrm{e}^{-\phi(t,\varepsilon u) - \langle x, \psi(t,\varepsilon u)\rangle} \\ = -\left(D\phi(t, \varepsilon u)(u) + \langle x, D\psi(t, \varepsilon u)(u)\rangle\right) \mathrm{e}^{-\phi(t,\varepsilon u) - \langle x, \psi(t,\varepsilon u)\rangle} \\ \to -\left(D\phi(t, 0)(u) + \langle x, D\psi(t, 0)(u)\rangle\right) \quad \text{as } \varepsilon \to 0,\end{aligned}$$

where we used that the functions $D\phi$ and $D\psi$ have a jointly continuous extension on $\mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$ in accordance with Proposition 4.1. On the other hand, noting that

$$|\langle u, \xi\rangle \exp\left(-\langle \varepsilon u, \xi\rangle\right)| \leq \varepsilon^{-1}\mathrm{e}^{-1}$$

and applying the dominated convergence theorem, we get

$$\frac{\partial}{\partial\varepsilon}\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle \varepsilon u, \xi\rangle} p_t(x, \mathrm{d}\xi) = -\int_{\mathbb{S}_d^+} \langle u, \xi\rangle \mathrm{e}^{-\langle \varepsilon u, \xi\rangle} p_t(x, \mathrm{d}\xi) \to -\int_{\mathbb{S}_d^+} \langle u, \xi\rangle p_t(x, \mathrm{d}\xi) \quad \text{as } \varepsilon \to 0.$$

Note that the limit on the right-hand side is finite. Indeed, using Fatou's lemma, we obtain

$$\int_{\mathbb{S}_d^+} \langle u, \xi\rangle p_t(x, \mathrm{d}\xi) \leq \liminf_{\varepsilon\to 0} \int_{\mathbb{S}_d^+} \langle u, \xi\rangle \mathrm{e}^{-\langle \varepsilon u, \xi\rangle} p_t(x, \mathrm{d}\xi) = D\phi(t, 0)(u) + \langle x, D\psi(t, 0)(u)\rangle < \infty$$

for all $u \in \mathbb{S}_d^+$. So

$$\int_{\mathbb{S}_d^+} \langle u, \xi\rangle p_t(x, \mathrm{d}\xi) = D\phi(t, 0)(u) + \langle x, D\psi(t, 0)(u)\rangle. \tag{26}$$

In what follows, we compute the derivatives $D\phi(t, 0)$ and $D\psi(t, 0)$ explicitly. By means of the generalized Riccati equation (4), we have

$$\psi(t, u) - u = \int_0^t R\left(\psi(s, u)\right) ds, \quad t \geq 0, \ u \in \mathbb{S}_d^+.$$

According to Lemma 4.2 and Proposition 4.1 we are allowed to differentiate both sides of the latter equation with respect to $u \in \mathbb{S}_d^+$ and evaluate at $u = 0$; thus, using the dominated convergence theorem,

$$D\psi(t, u)|_{u=0} - \mathrm{Id} = \int_0^t DR\left(\psi(s, u)\right) D\psi(s, u)|_{u=0}\, \mathrm{d}s, \quad t \geq 0,$$

where Id denotes the identity map on $\mathbb{S}_d^+$. From [12, Lemma 3.2(iii)] we know that $\psi(t, u)$ is continuous in $\mathbb{R}_{\geq 0} \times \mathbb{S}_d^+$, and noting that $\psi(s, 0) = 0$ (see [12, Remark 2.5]), we get

$$D\psi(t, 0) - \mathrm{Id} = \int_0^t DR\left(0\right) D\psi(s, 0)\mathrm{d}s, \quad t \geq 0.$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

From this and the precise formula for $\phi(t, h)$ we deduce that

$$D\psi(t, 0) = \mathrm{e}^{tDR(0)} \quad \text{and} \qquad D\phi(t, 0) = \int_0^t DF(0)\mathrm{e}^{sDR(0)}\mathrm{d}s.$$

We use Lemma 4.2 to get that

$$DR(0)(u) = \widetilde{B}^\top(u) \quad \text{and} \quad DF(0)(u) = \langle b + \int_{\mathbb{S}_d^+\setminus\{0\}} \xi m(\mathrm{d}\xi), u\rangle.$$

Finally, combining this with (26) yields

$$\begin{aligned}\int_{\mathbb{S}_d^+} \langle u, \xi\rangle p_t(x, \mathrm{d}\xi) &= \int_0^t (DF(0))\, \mathrm{e}^{sDR(0)}(u)\mathrm{d}s + \langle x, \mathrm{e}^{tDR(0)}(u)\rangle \\ &= \int_0^t \langle \mathrm{e}^{s\widetilde{B}} \left(b + \int_{\mathbb{S}_d^+\setminus\{0\}} \xi m(\mathrm{d}\xi)\right), u\rangle \mathrm{d}s + \langle \mathrm{e}^{t\widetilde{B}} x, u\rangle.\end{aligned}$$

Since the equality holds for each $u \in \mathbb{S}_d^+$, the assertion is proved.

5. Estimates on $\psi(t, u)$

We fix an admissible parameter set (α, b, B, m, μ) and let ψ be the unique solution to (4). In this section we study upper and lower bounds for ψ. Let us start with an upper bound for $\psi(t, u)$.

Propostion 5.1. *Let ψ be the unique solution to (4). Then*

$$\|\psi\ (t, u)\| \le M\|u\|\mathrm{e}^{-t\delta}, \qquad t \ge 0, \tag{27}$$

where M and δ are given by (9).

Proof. The proof is divided into three steps.

Step 1: Denote by $q_t(x, \mathrm{d}\xi)$ the unique transition kernel of an affine process on $\mathbb{S}_d^+$ with admissible parameters $(\alpha, b, B, m = 0, \mu)$; that is, for each $u,\ x \in \mathbb{S}_d^+$, we have

$$\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,\xi\rangle} q_t(x, \mathrm{d}\xi) = \exp\left(-\int_0^t \langle b, \psi(s, u)\rangle \mathrm{d}s - \langle x, \psi(t, u)\rangle\right), \qquad t \ge 0. \tag{28}$$

Applying Jensen's inequality to the convex function $t \mapsto \exp(-t)$ yields

$$\begin{aligned}\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u,\xi\rangle} q_t(x, \mathrm{d}\xi) &\ge \exp\left(-\int_{\mathbb{S}_d^+} \langle u, \xi\rangle q_t(x, \mathrm{d}\xi)\right) \\ &= \exp\left(-\int_0^t \langle \mathrm{e}^{s\widetilde{B}} b, u\rangle \mathrm{d}s - \langle \mathrm{e}^{t\widetilde{B}} x, u\rangle\right),\end{aligned} \tag{29}$$

where the last identity is a special case of Theorem 2.2. Using (28) we obtain

$$\langle x, \psi(t, u)\rangle + \int_0^t \langle b, \psi(t, u)\rangle \mathrm{d}s \le \langle \mathrm{e}^{t\widetilde{B}} x, u\rangle + \int_0^t \langle \mathrm{e}^{s\widetilde{B}} b, u\rangle \mathrm{d}s, \qquad \text{for all } u,\ x \in \mathbb{S}_d^+,\ t \ge 0. \tag{30}$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Step 2: Let $\alpha \in \mathbb{S}_d^+$ be fixed. We claim that (30) holds not only for $b \succeq (d-1)\alpha$ but also for any $b \in \mathbb{S}_d^+$. Aiming for a contradiction, suppose that there exist $t_0 > 0$ and $\xi,\ x_0,\ u_0 \in \mathbb{S}_d^+$ such that

$$I := \langle x_0, \psi(t_0, u_0)\rangle + \int_0^{t_0} \langle \xi, \psi(s, u_0)\rangle \mathrm{d}s - \langle x_0, \mathrm{e}^{t_0 \widetilde{B}^\top} u_0\rangle - \int_0^{t_0} \langle \xi, \mathrm{e}^{s\widetilde{B}^\top} u_0\rangle \mathrm{d}s > 0.$$

We now take an arbitrary but fixed $b_0 \succeq (d-1)\alpha$. Noting that

$$\Delta := \int_0^{t_0} \langle b_0, \psi(s, u_0)\rangle \mathrm{d}s - \int_0^{t_0} \langle b_0, \mathrm{e}^{s\widetilde{B}^\top} u_0\rangle \mathrm{d}s$$

is finite, we find a constant $K > 0$ large enough so that $KI + \Delta > 0$, i.e.,

$$\langle Kx_0, \psi(t_0, u_0)\rangle + \int_0^{t_0} \langle b_0 + K\xi, \psi(s, u_0)\rangle \mathrm{d}s > \langle Kx_0, \mathrm{e}^{t_0\widetilde{B}^\top} u_0\rangle + \int_0^{t_0} \langle b_0 + K\xi, \mathrm{e}^{s\widetilde{B}^\top} u_0\rangle \mathrm{d}s. \tag{31}$$

Now, since $b_0 + K\xi \succeq (d-1)\alpha$, we see that (31) contradicts (30) if we choose $b = b_0 + K\xi$, $x = Kx_0$, $u = u_0$, and $t = t_0$. Hence (30) holds for all $b \in \mathbb{S}_d^+$.

Step 3: According to Step 2, we are allowed to choose $b = 0$ in (30), which implies

$$\langle x, \psi(t, u)\rangle \le \langle x, \mathrm{e}^{t\widetilde{B}^\top} u\rangle$$

for all $t \ge 0$ and $x,\ u \in \mathbb{S}_d^+$. This completes the proof.

We continue with a lower bound for $\psi(t, u)$.

Proposition 5.2. *Let ψ be the unique solution to (4) and suppose that $\alpha = 0$ and (14) is satisfied. Then, for each $u,\ \xi \in \mathbb{S}_d^+$,*

$$\langle \xi, \psi(t, u)\rangle \ge \mathrm{e}^{-Kt}\langle \xi, u\rangle, \qquad t \ge 0. \tag{32}$$

Proof. Fix $u \in \mathbb{S}_d^+$ and define $W_t(u) := \psi(t, u) - \exp(-Kt)u$. Using that $\exp(-Kt)u = \psi(t, u) - W_t(u)$ we obtain

$$\frac{\partial W_t(u)}{\partial t} = R(\psi(t, u)) + K\psi(t, u) - KW_t(u).$$

Since $W_0(u) = 0$, the latter implies

$$W_t(u) = \int_0^t \mathrm{e}^{-K(t-s)} \left(K\psi(s, u) + R(\psi(s, u))\right) \mathrm{d}s.$$

Fix $\xi \in \mathbb{S}_d^+$; then

$$\langle \xi, W_t(u)\rangle = \int_0^t \mathrm{e}^{-K(t-s)} \left(K\langle \xi, \psi(s, u)\rangle + \langle \xi, R(\psi(s, u))\rangle\right) \mathrm{d}s. \tag{33}$$

In the following we estimate the integrand. For this, we write $\langle \xi, R(\psi(s, u))\rangle = I_1 + I_2$, where

$$I_1 = \langle \xi, B^\top(\psi(s, u))\rangle \quad \text{and} \quad I_2 = -\int_{\mathbb{S}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \psi(s,u), \zeta\rangle} - 1\right)\langle \xi, \mu(\mathrm{d}\zeta)\rangle,$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

and estimate I_1 and I_2 separately. For I_1, by (14) we get

$$I_1 = \langle B(\xi),\, \psi(s,u)\rangle \geq -K\langle \xi,\, \psi(s,u)\rangle,$$

where we used the self-duality of the cone $\mathbb{S}_d^+$ (see [27, Theorem 7.5.4]). Turning to I_2, we simply have

$$I_2 = \int_{\mathbb{S}_d^+\setminus\{0\}} \left(1 - \mathrm{e}^{-\langle \psi(s,u),\zeta\rangle}\right) \langle \xi,\, \mu\,(\mathrm{d}\zeta)\rangle \geq 0.$$

Collecting now the estimates for I_1 and I_2, we see that

$$(K\langle \xi,\, \psi(s,u)\rangle + \langle \xi,\, R(\psi(s,u))\rangle) \geq 0,$$

and thus $\langle \xi,\, W_t(u)\rangle \geq 0$ by (33). This proves the assertion.

6. Proofs of the main results

In this section we will prove Theorem 2.2, Proposition 2.1, and Corollary 2.1. Let $p_t(x, \mathrm{d}\xi)$ be the transition kernel of a subcritical affine process on $\mathbb{S}_d^+$ with admissible parameters (α, b, B, m, μ), and let $\delta > 0$ be given by (9).

We note that $F(u) \geq 0$ for all $u \in \mathbb{S}_d^+$. Based on the estimates on $\psi(t, u)$ that we derived in the previous section, we easily obtain the following lemma.

Lemma 6.1. *Suppose that (10) holds. Then there exists a constant* $C > 0$ *such that*

$$F(\psi(s,u)) \leq C\|u\|\mathrm{e}^{-s\delta}, \qquad s \geq 0, \quad u \in \mathbb{S}_d^+. \tag{34}$$

Consequently,

$$\int_0^\infty F(\psi(s,u))\mathrm{d}s \leq \frac{C}{\delta}\|u\|, \qquad u \in \mathbb{S}_d^+. \tag{35}$$

Proof. We know that

$$\begin{aligned} F(\psi(s,u)) &= \langle b,\, \psi(s,u)\rangle + \int_{\mathbb{S}_d^+\setminus\{0\}} \left(1 - \mathrm{e}^{-\langle \psi(s,u),\xi\rangle}\right) m(\mathrm{d}\xi) \\ &=: \langle b,\, \psi(s,u)\rangle + I(u). \end{aligned}$$

Now, first note that, by (27),

$$\langle b,\, \psi(s,u)\rangle \leq \|b\|\,\|\psi(s,u)\| \leq \|b\|\,\|u\|\mathrm{e}^{-s\delta}. \tag{36}$$

We turn to the estimate of $I(u)$. Once again using (27), we obtain

$$\begin{aligned} I(u) &= \int_{\mathbb{S}_d^+\setminus\{0\}} \left(1 - \mathrm{e}^{-\langle \psi(s,u),\xi\rangle}\right) m(\mathrm{d}\xi) \\ &\leq \int_{\mathbb{S}_d^+\setminus\{0\}} \min\{1,\, \langle \psi(s,u),\, \xi\rangle\}\, m(\mathrm{d}\xi) \\ &\leq \int_{\mathbb{S}_d^+\setminus\{0\}} \min\left\{1,\, \|\xi\|\,\|u\|\mathrm{e}^{-s\delta}\right\} m(\mathrm{d}\xi). \end{aligned}$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

For all $a \geq 0$ it holds that $1 \wedge a \leq \log(2)^{-1} \log(1+a)$; hence

$$I(u) \leq \frac{1}{\log(2)} \int_{\{\|\xi\| \leq 1\}} \|\xi\| \|u\| \mathrm{e}^{-s\delta} m(\mathrm{d}\xi) + \frac{1}{\log(2)} \int_{\{\|\xi\|>1\}} \log\left(1 + \|\xi\| \|u\| \mathrm{e}^{-s\delta}\right) m(\mathrm{d}\xi)$$
$$=: J_1(u) + J_2(u).$$

Let $C > 0$ be a generic constant which may vary from line to line. Since $m(\mathrm{d}\xi)$ integrates $\|\xi\| \mathbb{1}_{\{\|\xi\| \leq 1\}}$ by definition, we have

$$J_1(u) \leq C\|u\| \mathrm{e}^{-s\delta}.$$

Moreover, noting that $m(\mathrm{d}\xi)$ integrates $\log \|\xi\| \mathbb{1}_{\{\|\xi\|>1\}}$ by assumption, for $J_2(u)$ we use the elementary inequality (see [19, Lemma 8.5])

$$\begin{aligned}\log(1 + a \cdot c) &\leq C \min\{\log(1+a), \log(1+c)\} + C \log(1+a) \log(1+c) \\ &\leq C \log(1+a) + Ca \log(1+c) \\ &\leq Ca\,(1 + \log(1+c))\end{aligned}$$

for $a = \|u\| \exp(-s\delta)$ and $c = \|\xi\|$ to get

$$J_2(u) \leq C\|u\| \mathrm{e}^{-s\delta} \int_{\{\|\xi\|>1\}} (1 + \log(1 + \|\xi\|))\, m(\mathrm{d}\xi) \leq C\|u\| \mathrm{e}^{-s\delta}.$$

Combining the estimates for $J_1(u)$ and $J_2(u)$ yields

$$I(u) = J_1(u) + J_2(u) \leq C\|u\| \mathrm{e}^{-s\delta}. \tag{37}$$

So, by (36) and (37), we have (34), which proves the assertion.

We are now able to prove Theorem 2.3.

Proof of Theorem 2.3. Fix $x \in \mathbb{S}_d^+$. By means of Proposition 5.1, we see that

$$\begin{aligned}\lim_{t\to\infty} \int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u, \xi \rangle} p_t(x, \mathrm{d}\xi) &= \lim_{t\to\infty} \exp\left(-\int_0^t F(\psi(s, u))\mathrm{d}s - \langle x, \psi(t, u)\rangle\right) \\ &= \exp\left(-\int_0^\infty F(\psi(s, u))\mathrm{d}s\right),\end{aligned}$$

and the limit on the right-hand side is finite by Lemma 6. Clearly, by (35), we also have that $u \mapsto \int_0^\infty F(\psi(s, u)))\mathrm{d}s$ is continuous at $u = 0$. Now, Lévy's continuity theorem (cf. [12, Lemma 4.5]) implies that $p_t(x, \cdot) \to \pi$ weakly as $t \to \infty$. Moreover, π has Laplace transform (11). It remains to verify that π is the unique invariant distribution.

Invariance. Fix $u \in \mathbb{S}_d^+$ and let $t \geq 0$ be arbitrary. Then

$$\begin{aligned}\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u, \xi \rangle} \left(\int_{\mathbb{S}_d^+} p_t(x, \mathrm{d}\xi) \pi(\mathrm{d}x) \right) &= \int_{\mathbb{S}_d^+} \left(\int_{\mathbb{S}_d^+} \mathrm{e}^{-\langle u, \xi \rangle} p_t(x, \mathrm{d}\xi) \right) \pi(\mathrm{d}x) \\ &= \mathrm{e}^{-\phi(t,u)} \int_{\mathbb{S}_d^+} \exp(-\langle x, \psi(t, u)\rangle)\, \pi(\mathrm{d}x).\end{aligned}$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Note that ψ satisfies the semi-flow equation by [12, Lemma 3.2]; that is, $\psi(t+s, u) = \psi(s, \psi(t, u))$ for all $t, s \geq 0$. Using that the Laplace transform of π is given by (11), for each $u \in \mathbb{S}_d^+$ we obtain

$$\begin{aligned}\int_{\mathbb{S}_d^+} e^{-\langle u,\xi\rangle} \left(\int_{\mathbb{S}_d^+} p_t(x, d\xi)\pi(dx)\right) &= e^{-\phi(t,u)} \exp\left(-\int_0^\infty F(\psi(s, \psi(t, u)))\, ds\right)\\ &= e^{-\phi(t,u)} \exp\left(-\int_0^\infty F(\psi(t+s, u))\, ds\right)\\ &= e^{-\phi(t,u)} \exp\left(-\int_t^\infty F(\psi(s, u))\, ds\right)\\ &= \exp\left(-\int_0^\infty F(\psi(s, u))\, ds\right)\\ &= \int_{\mathbb{S}_d^+} e^{-\langle x,u\rangle}\pi(dx).\end{aligned}$$

Consequently, π is invariant.

Uniqueness. Let π' be another invariant distribution. For fixed $u \in \mathbb{S}_d^+$ and $t \geq 0$ we have

$$\begin{aligned}\int_{\mathbb{S}_d^+} e^{-\langle x,u\rangle}\pi'(dx) &= \int_{\mathbb{S}_d^+} e^{-\langle u,\xi\rangle} \left(\int_{\mathbb{S}_d^+} p_t(x, d\xi)\pi'(dx)\right)\\ &= \int_{\mathbb{S}_d^+} \exp(-\phi(t, u) - \langle x, \psi(t, u)\rangle)\, \pi'(dx).\end{aligned}$$

Letting $t \to \infty$ shows that π' also satisfies (11). By uniqueness of the Laplace transforms, it holds that $\pi' = \pi$.

Proof of Proposition 2.1. Let $x \in \mathbb{S}_d^+$ and $\pi \in \mathcal{P}(\mathbb{S}_d^+)$ be such that $p_t(x, \cdot) \to \pi$ weakly as $t \to \infty$. It follows that

$$\lim_{t\to\infty} \int_{\mathbb{S}_d^+} e^{-\langle u,\xi\rangle} p_t(x, d\xi) = \int_{\mathbb{S}_d^+} e^{-\langle u,y\rangle}\pi(dy), \qquad u \in \mathbb{S}_d^+,$$

and we obtain from (2) that

$$\lim_{t\to\infty} \exp\left(-\int_0^t F(\psi(s, u))ds\right) = \lim_{t\to\infty} e^{\langle x,\psi(t,u)\rangle} \int_{\mathbb{S}_d^+} e^{-\langle u,\xi\rangle} p_t(x, d\xi) = \int_{\mathbb{S}_d^+} e^{-\langle u,y\rangle}\pi(dy).$$

In particular, this implies

$$\int_0^\infty F(\psi(s, u))ds < \infty, \qquad u \in \mathbb{S}_d^+.$$

Fix $u \in \mathbb{S}_d^{++}$. Assume that $\alpha = 0$ and (14) holds. By definition of F we have $F(u) \geq \int_{\mathbb{S}_d^+} (1 - \exp(-\langle u, \xi\rangle))m(d\xi)$ and thereby

$$F(\psi(s, u)) \geq \int_{\{\langle \xi,u\rangle>1\}} \left(1 - e^{-e^{-Ks}\langle \xi,u\rangle}\right) m(d\xi),$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

where we used (32). Integrating over $[0, \infty)$ and using a change of variable $r := \exp(-Ks)\langle \xi, u \rangle$ with $ds = -1/K \cdot dr/r$ yields

$$\int_0^\infty F(\psi(s, u))ds \geq \frac{1}{K} \int_{\{\langle \xi, u \rangle > 1\}} \int_0^{\langle \xi, u \rangle} \frac{1 - e^{-r}}{r} dr m(d\xi)$$
$$\geq \frac{1}{K} \int_{\{\langle \xi, u \rangle > 1\}} \int_1^{\langle \xi, u \rangle} \frac{1 - e^{-r}}{r} dr m(d\xi)$$
$$\geq \frac{1 - e^{-1}}{K} \int_{\{\langle \xi, u \rangle > 1\}} \log(\langle \xi, u \rangle) m(d\xi),$$

where we used in the last inequality that $1 - \exp(-r) \geq 1 - \exp(-1) > 0$ for $r \geq 1$. This leads to the estimate

$$\int_{\{\langle \xi, u \rangle > 1\}} \log(\langle \xi, u \rangle) m(d\xi) \leq \frac{K}{1 - e^{-1}} \int_0^\infty F(\psi(s, u))ds < \infty.$$

Letting $u = \mathbb{1} \in \mathbb{S}_d^{++}$ gives $\langle \xi, \mathbb{1} \rangle = \mathrm{tr}(\xi) \geq \|\xi\|$, so that

$$\int_{\{\|\xi\| > 1\}} \log(\|\xi\|)\, m(d\xi) \leq \int_{\{\langle \xi, \mathbb{1} \rangle > 1\}} \log(\langle \xi, \mathbb{1} \rangle)\, m(d\xi) < \infty.$$

This completes the proof.

Proof of Corollary 2.1. Using that $\|\exp(t\widetilde{B})\| \leq M \exp(-\delta t)$, where δ is given by (9), we have

$$\lim_{t\to\infty} \int_{\mathbb{S}_d^+} y p_t(x, dy) = \int_0^\infty e^{s\widetilde{B}} \left(b + \int_{\mathbb{S}_d^+} \xi m(d\xi) \right) ds \in \mathbb{S}_d^+.$$

By direct computation we find that

$$\int_0^\infty e^{s\widetilde{B}} \left(b + \int_{\mathbb{S}_d^+ \setminus \{0\}} \xi m(d\xi) \right) ds = -\widetilde{B}^{-1} \left(b + \int_{\mathbb{S}_d^+ \setminus \{0\}} \xi m(d\xi) \right),$$

and hence it suffices to prove that $\lim_{t\to\infty} \int_{\mathbb{S}_d^+} y p_t(x, dy) = \int_{\mathbb{S}_d^+} y \pi(dy)$. To do so, we can proceed as in the proof of Theorem 2.2. Indeed, by Lemma A.1, we estimate

$$\sup_{t\geq 0} \int_{\mathbb{S}_d^+} \|y\| p_t(x, dy) \leq \sup_{t\geq 0} \mathrm{tr} \left(\int_{\mathbb{S}_d^+} y p_t(x, dy) \right) \leq \sqrt{d} \sup_{t\geq 0} \left\| \int_{\mathbb{S}_d^+} y p_t(x, dy) \right\| < \infty.$$

Therefore, applying Fatou's lemma yields

$$\int_{\mathbb{S}_d^+} \|y\| \pi(dy) \leq \sup_{t\geq 0} \int_{\mathbb{S}_d^+} \|y\| p_t(x, dy) < \infty.$$

So $\pi \in \mathcal{P}_1(\mathbb{S}_d^+)$. Now, let $\varepsilon > 0$. By the dominated convergence theorem, we see that

$$\lim_{\varepsilon \searrow 0} \int_{\mathbb{S}_d^+} \frac{1 - e^{-\langle \varepsilon u, y \rangle}}{\varepsilon} \pi(dy) = \int_{\mathbb{S}_d^+} \langle u, y \rangle \pi(dy).$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

Moreover, noting that, by Proposition 5.1,

$$1-\mathrm{e}^{-\langle\psi(s,\varepsilon u),\xi\rangle}\le\langle\psi(s,\varepsilon u),\xi\rangle\le\|\xi\|\|\varepsilon u\|\mathrm{e}^{-\delta t},$$

we can use once again the dominated convergence theorem to obtain

$$\begin{aligned}\lim_{\varepsilon\searrow 0}\int_{\mathbb{S}_d^+}\frac{1-\mathrm{e}^{-\langle\varepsilon u,y\rangle}}{\varepsilon}\pi(\mathrm{d}y)&=\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\int_0^\infty F(\psi(s,\varepsilon u))\mathrm{d}s\\&=\lim_{\varepsilon\searrow 0}\int_0^\infty\left(\langle b,\frac{\psi(s,\varepsilon u)}{\varepsilon}\rangle\mathrm{d}s+\int_{\mathbb{S}_d^+\setminus\{0\}}\frac{1-\mathrm{e}^{-\langle\psi(s,\varepsilon u),\xi\rangle}}{\varepsilon}m(\mathrm{d}\xi)\right)\mathrm{d}s\\&=\int_0^\infty\langle b,D\psi(s,0)(u)\rangle\mathrm{d}s+\int_0^\infty\int_{\mathbb{S}_d^+\setminus\{0\}}\langle D\psi(s,0)(u),\xi\rangle m(\mathrm{d}\xi)\mathrm{d}s\\&=\int_0^\infty\langle\mathrm{e}^{s\widetilde{B}}\left(b+\int_{\mathbb{S}_d^+\setminus\{0\}}\xi m(\mathrm{d}\xi)\right),u\rangle\mathrm{d}s,\end{aligned}$$

where we used that $D\psi(s,0)(u)=\exp(s\widetilde{B}^\top)u$ (see the proof of Theorem 2.2). Since the latter identity holds for all $u\in\mathbb{S}_d^+$, this concludes the proof.

7. Proof of Theorem 2.4

Proof of Theorem 2.4. Suppose that (10) holds. By definition of d_L, we have

$$\begin{aligned}&d_L\,(p_t(x,\mathrm{d}\xi),\pi(\mathrm{d}\xi))\\&=\sup_{u\in\mathbb{S}_d^+\setminus\{0\}}\frac{1}{\|u\|}\left|\int_{\mathbb{S}_d^+}\mathrm{e}^{-\langle u,\xi\rangle}p_t(x,\mathrm{d}\xi)-\int_{\mathbb{S}_d^+}\mathrm{e}^{-\langle u,\xi\rangle}\pi(\mathrm{d}\xi)\right|\\&=\sup_{u\in\mathbb{S}_d^+\setminus\{0\}}\frac{1}{\|u\|}\left|\exp\left(-\int_0^t F\,(\psi(s,u))\,\mathrm{d}s-\langle x,\psi(t,u)\rangle\right)-\exp\left(-\int_0^\infty F\,(\psi(s,u))\,\mathrm{d}s\right)\right|.\end{aligned}\tag{38}$$

Let $C>0$ be a generic constant that may vary from line to line. Then, using (34), for each $t\ge 0$ we have

$$\begin{aligned}&\left|\exp\left(-\int_0^t F\,(\psi(s,u))\,\mathrm{d}s-\langle x,\psi(t,u)\rangle\right)-\exp\left(-\int_0^\infty F\,(\psi(s,u))\,\mathrm{d}s\right)\right|\\&\quad\le|\exp(-\langle x\psi(t,u)\rangle)-1|\cdot\left|\exp\left(-\int_0^t F\,(\psi(s,u))\,\mathrm{d}s\right)\right|\\&\qquad+\left|\exp\left(-\int_0^t F\,(\psi(s,u))\,\mathrm{d}s\right)-\exp\left(-\int_0^\infty F\,(\psi(s,u))\,\mathrm{d}s\right)\right|\\&\quad\le|\langle x,\psi(s,u)\rangle|+\left|\int_t^\infty F\,(\psi(s,u))\,\mathrm{d}s\right|\\&\quad\le M\|x\|\|u\|\mathrm{e}^{-t\delta}+C\|u\|\int_t^\infty\mathrm{e}^{-s\delta}\mathrm{d}s\\&\quad\le C\,(1+\|x\|)\,\|u\|\mathrm{e}^{-t\delta},\end{aligned}$$

which when plugged back into (38) implies (16).

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

8. Proof of Theorem 2.5

Proof of Theorem 2.5. Note that $\pi \in \mathcal{P}_1(\mathbb{S}_d^+)$ by Corollary 2.1. Let $q_t(x, \mathrm{d}\xi)$ be a transition kernel for the conservative, subcritical affine process with admissible parameters $(\alpha = 0, b = 0, B, m = 0, \mu)$. Using the particular form of the Laplace transform for $p_t(x, \cdot)$ (see (2)), it is not difficult to see that $p_t(x, \cdot) = q_t(x, \cdot) * p_t(0, \cdot)$, where '$*$' denotes the convolution of measures. Let H be any coupling with marginals δ_x and π, i.e., $H \in \mathcal{H}(\delta_x, \pi)$. Using the invariance of π, together with the convexity of W_1 (see [48, Theorem 4.8] and [18, Lemma 2.3]), we find that

$$
\begin{aligned}
W_1\,(p_t(x,\cdot),\pi) &= W_1\left(\int_{\mathbb{S}_d^+} p_t(y,\cdot)\delta_x(\mathrm{d}y), \int_{\mathbb{S}_d^+} p_t(y',\cdot)\pi(\mathrm{d}y')\right) \\
&\le \int_{\mathbb{S}_d^+\times\mathbb{S}_d^+} W_1\left(p_t(y,\cdot), p_t(y',\cdot)\right) H(\mathrm{d}y, \mathrm{d}y') \\
&\le \int_{\mathbb{S}_d^+\times\mathbb{S}_d^+} W_1\left(q_t(y,\cdot), q_t(y',\cdot)\right) H(\mathrm{d}y, \mathrm{d}y').
\end{aligned}
$$

The integrand can now be estimated as follows:

$$
\begin{aligned}
W_1\left(q_t(y,\cdot), q_t(y',\cdot)\right) &\le \int_{\mathbb{S}_d^+\times\mathbb{S}_d^+} \|z - z'\| G(\mathrm{d}z, \mathrm{d}z') \\
&\le \int_{\mathbb{S}_d^+} \|z\| q_t(y, \mathrm{d}z) + \int_{\mathbb{S}_d^+} \|z'\| q_t(y', \mathrm{d}z') \\
&\le M\sqrt{d}\mathrm{e}^{-t\delta}\left(\|y\| + \|y'\|\right),
\end{aligned}
$$

where G is any coupling of $(q_t(y,\cdot), q_t(y',\cdot))$ and we have used Lemma A.1 to obtain

$$
\begin{aligned}
\int_{\mathbb{S}_d^+} \|z\| q_t(y, \mathrm{d}z) &\le \operatorname{tr}\left(\int_{\mathbb{S}_d^+} z q_t(y, \mathrm{d}z)\right) \\
&= \operatorname{tr}\left(\mathrm{e}^{t\widetilde{B}} y\right) \\
&\le \sqrt{d}\left\|\mathrm{e}^{t\widetilde{B}} y\right\| \\
&\le M\sqrt{d}\mathrm{e}^{-t\delta}\|y\|.
\end{aligned}
$$

Combining these estimates, we obtain

$$
\begin{aligned}
W_1(p_t(x,\cdot),\pi) &\le M\sqrt{d}\mathrm{e}^{-t\delta}\int_{\mathbb{S}_d^+\times\mathbb{S}_d^+}\left(\|y\| + \|y'\|\right) H(\mathrm{d}y, \mathrm{d}y') \\
&\le M\sqrt{d}\mathrm{e}^{-t\delta}\left(\|x\| + \int_{\mathbb{S}_d^+}\|y\|\pi(\mathrm{d}y)\right),
\end{aligned}
$$

which yields (17).

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

9. Extension to affine processes on finite-dimensional convex cones

In this section we outline how our main result, Theorem 2.3, can be extended to the more general framework of affine processes on convex cones. Below we briefly introduce the necessary concepts, while additional details, proofs, and references can be found in [13].

Let V be a finite-dimensional Euclidean space with scalar product $\langle\cdot,\cdot\rangle$ and associated norm $\|\cdot\|$. Let $K\subset V$ be a closed convex cone and suppose that it is proper, i.e. $K\cap(-K)=\{0\}$, and generating, i.e. $V=K-K$. Note that its closed dual cone

$$K^*=\{u\in V \mid \langle x, u\rangle\geq 0,\quad \forall x\in K\}$$

is then also generating and proper. The partial order on V is, for $u, v\in V$, defined by $u\preceq v \Longleftrightarrow v-u\in K^*$. The following is the definition of affine transition semigroups due to [13, Definition 2.1].

Definition 9.1. *(Cuchiero et al. [13].)* A Markov transition kernel $(p_t(x,\mathrm{d}\xi))_{t\geq 0, x\in K}$ is called affine if the following hold:

(i) $p_t(x,\cdot)\longrightarrow p_s(x,\cdot)$ weakly as $t\to s$ for every $s\geq 0$ and $x\in K$.

(ii) There exist functions $\phi:\mathbb{R}_+\times K^*\longrightarrow\mathbb{R}$ and $\psi:\mathbb{R}_+\times K^*\longrightarrow V$ such that

$$\int_K \mathrm{e}^{-\langle u,\xi\rangle}p_t(x,\mathrm{d}\xi)=\mathrm{e}^{-\phi(t,u)-\langle\psi(t,u),x\rangle},\qquad x\in K,\quad (t,u)\in\mathbb{R}_+\times K^*. \tag{39}$$

Note that this definition allows non-conservative affine transition kernels. However, having in mind that we investigate existence, uniqueness, and convergence of $p_t(x,\cdot)$ to the invariant distribution, we restrict our study to conservative kernels; that is, $p_t(x,K)=1$ for all $t\geq 0$ and $x\in K$. The following is a summary of the main results obtained in [13] on the existence and structure of conservative affine transition kernels. We emphasize that these results hold in a more general setting that allows killing and explosion of the process; see [13] for more details.

Theorem 9.1. (Cuchiero et al. [13])

(a) *Let $p_t(x,\mathrm{d}\xi)$ be a conservative affine transition kernel. Then the associated transition semigroup is C_0-Feller, and the functions ϕ,ψ are differentiable with respect to t and satisfy the generalized Riccati equations for $u\in K^*$; that is,*

$$\partial_t\phi(t,u)=F(\psi(t,u)),\qquad \phi(0,u)=0, \tag{40}$$

$$\partial_t\psi(t,u)=R(\psi(t,u)),\qquad \psi(0,u)=u\in K^*,$$

where $F(u)=\partial_t\phi(t,u)|_{t=0}$ and $R(u)=\partial_t\psi(t,u)|_{t=0}$. Moreover, there exists a parameter set (Q,b,B,m,μ) such that the functions F and R are of the form

$$F(u)=\langle b,u\rangle-\int_{K\setminus\{0\}}\left(\mathrm{e}^{-\langle u,\xi\rangle}-1\right)m(\mathrm{d}\xi),$$

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

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

where

(i) $b \in K$;

(ii) *m is a Borel measure on* $K\backslash\{0\}$ *with*

$$\int_{K\backslash\{0\}} (1 \wedge \|\xi\|)\, m(\mathrm{d}\xi) < \infty;$$

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

(iv) μ *is a* K^**-valued* σ*-finite Borel measure on* $K\backslash\{0\}$ *satisfying*

$$\int_{K\backslash\{0\}} \left(1 \wedge \|\xi\|^2\right) \langle x, \mu(\mathrm{d}\xi)\rangle < \infty, \qquad x \in K;$$

(v) $B{:}V \longrightarrow V$ *is a linear map satisfying, for all* $u \in K^*$ *and* $x \in K$ *with* $\langle u, x\rangle = 0$,

$$\int_{K\backslash\{0\}} \mathbb{1}_{\{\|\xi\|\le 1\}} \langle \xi, u\rangle \langle x, \mu(\mathrm{d}\xi)\rangle \le \langle x, B^\top(u)\rangle < \infty.$$

(b) *Conversely, let* $(Q = 0, b, B, m, \mu)$ *be a parameter set satisfying the above conditions, and suppose in addition that*

$$\int_{K\backslash\{0\}} \mathbb{1}_{\{\|\xi\|>1\}} \|\xi\| \langle x, \mu(\mathrm{d}\xi)\rangle < \infty, \qquad x \in K. \tag{41}$$

Then there exists a unique conservative affine transition semigroup on K such that (39) holds for all $(t, u) \in \mathbb{R}_+ \times K^*$, *where* $\phi(t, u)$ *and* $\psi(t, u)$ *are given by (40).*

Note that, for a given parameter set (Q, b, B, m, μ), existence of an affine transition kernel is only shown for the case $Q = 0$, while conservativeness is a consequence of the first moment condition (41) (see [13, p. 373] and follow the argument given in [16, Section 9]).

Assuming, in addition, that K is an irreducible symmetric cone (see [13, Definition 2.6]), there exists a conservative affine transition semigroup for any parameter set (Q, b, B, m, μ) fulfilling the above conditions (i)–(v) and (41); see [13, Theorem 2.19]. Moreover, for dimensions larger than 2 it has been shown that the jumps are of finite variation and the drift satisfies $b \succeq 4^{-1} d(r-1)Q(e, e)$, where e denotes the identity element for the multiplication on V, and r denotes the rank and d the Peirce invariant of V; see [13, Theorem 2.13] and [13, Appendix] for additional details. Note that if we let $V = S_d$ and $K = S_d^+$, we obtain the case of positive semidefinite matrices as a particular case of these results.

Below we provide an extension of Theorem 2.3 for a conservative affine transition kernel on a proper closed convex cone K, which is generating. Since we do not suppose that this cone is symmetric, existence of such a kernel is a priori not clear and should be established separately.

Theorem 9.2. *Suppose that there exists an affine transition kernel* $p_t(x, \mathrm{d}\xi)$ *with parameters* (Q, b, B, m, μ) *satisfying (41) (which makes the kernel* p_t *necessarily conservative). Define* $\widetilde{B}{:}V \longrightarrow V$ *by*

$$\widetilde{B}(u) := B(u) + \int_{K\backslash\{0\}} \mathbb{1}_{\{\|\xi\|>1\}} \xi \langle u, \mu(\mathrm{d}\xi)\rangle, \qquad u \in V,$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

and suppose that $\sigma(\widetilde{B}) \subset \{\lambda \in \mathbb{C} \mid \mathrm{Re}(\lambda) < 0\}$. *Moreover, assume that*

$$\int_{K\setminus\{0\}} \mathbb{1}_{\{\|\xi\|>1\}} \log(\|\xi\|) m(\mathrm{d}\xi) < \infty.$$

Then the conservative affine transition kernel $p_t(x, \mathrm{d}\xi)$ *has a unique invariant distribution* π. *Moreover, for each* $x \in K$, *one has* $p_t(x, \cdot) \longrightarrow \pi$ *weakly as* $t \to \infty$, *and* π *has Laplace transform*

$$\int_K \mathrm{e}^{-\langle u,\xi\rangle} \pi(\mathrm{d}\xi) = \exp\left(-\int_0^\infty F(\psi(s, u))\mathrm{d}s\right), \qquad u \in K^*,$$

where the latter integral is absolutely convergent.

Proof. We follow the proof of our main result, Theorem 2.3.

Step 1. Let us prove that there exist constants $M \geq 1$ and $\delta > 0$ such that

$$\|\psi(t, u)\| \leq M\|u\|\mathrm{e}^{-\delta t}, \qquad t \geq 0, \quad u \in K^*. \tag{42}$$

Indeed, one could check that Theorem 2.2 also holds in this situation (provided that m has finite first moment). One may think of deducing the assertion by arguments similar to those for Proposition 5.1. Unfortunately, the latter is built on a convolution argument, and it requires the existence of a conservative affine transition kernel with parameters $(Q, b, B, m = 0, \mu)$, which is currently not known. So we provide below an alternative and direct proof which avoids the convolution trick.

First observe that, for $u \in K^*$, it holds that

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

Let $y(t, u) = e^{t\widetilde{B}^\top}(u)$ be the unique solution to

$$\partial_t y(t, u) = \widetilde{B}^\top(y(t, u)), \quad y(0, u) = u \in K^*.$$

Note that R is quasi-monotone increasing on K^* by [13, Proposition 3.12]. Then

$$0 = \partial_t\psi(t, u) - R(\psi(t, u)) = \partial_t y(t, u) - \widetilde{B}^\top(y(t, u)) \preceq \partial_t y(t, u) - R(y(t, u)),$$

and hence by Volkmann's comparison theorem for ordinary differential equations (see [13, Theorem 3.13]) we conclude that

$$\psi(t, u) \preceq y(t, u) = \mathrm{e}^{t\widetilde{B}^\top}(u);$$

that is,

$$\langle\psi(t, u), x\rangle \leq \langle\mathrm{e}^{t\widetilde{B}^\top}(u), x\rangle, \quad x \in K.$$

Now we can repeat the argument in [13, p. 387]: since K^* is a finite-dimensional proper closed convex cone, it is normal; i.e., there exists a constant γ_{K^*} such that for $x, y \in K^*$,

$$0 \preceq x \preceq y \implies \|x\| \leq \gamma_{K^*}\|y\|.$$

We thus have

$$\|\psi(t, u)\| \leq \gamma_{K^*}\left\|\mathrm{e}^{t\widetilde{B}^\top}(u)\right\|, \quad u \in K^*.$$

https://doi.org/10.1017/apr.2020.21 Published online by Cambridge University Press

The property (42) is now an immediate consequence of the assumption that the spectrum of $\widetilde{B}$ is contained in the left half-plane of $\mathbb{C}$.

Step 2. Following exactly the same arguments as in Lemma 6.1, we find a constant $C > 0$ such that

$$\int_0^\infty F(\psi(t, u))\mathrm{d}t \leq C\|u\|, \qquad u \in K^*.$$

Step 3. Using [13, Proposition 3.1(i)–(ii)] combined with a version of Lévy's continuity theorem [13, Lemma 3.7], the assertion can be deduced by literally the same arguments as given in the proof of Theorem 2.3.

Appendix A. Matrix Calculus

For a $d \times d$ square matrix x, recall that $\mathrm{tr}(x) = \sum_{i=1}^d x_{ii}$. The Frobenius norm of x is given by $\|x\| = \mathrm{tr}(xx)^{1/2} = \big(\sum_{i,j=1}^d |x_{ij}|^2\big)^{1/2}$. Let us recollect one property of this norm.

Lemma A.1 *Let $x \in \mathbb{S}_d^+$; then*

$$\|x\| \leq \mathrm{tr}(x) \leq \sqrt{d}\|x\|.$$

Proof. Write $x = u^\top \kappa u$, where u is orthogonal and κ is diagonal with its entries given by $\lambda_i(x)$, $i = 1, \ldots, d$, the eigenvalues of x. We have

$$\|x\|^2 = \mathrm{tr}\left(u^\top \kappa^2 u\right) = \sum_{i=1}^d \lambda_i(x)^2.$$

Since $x \in \mathbb{S}_d^+$, it holds that $\lambda_i(x) \geq 0$, $i = 1, \ldots, d$. Then

$$\|x\| = \left(\sum_{i=1}^d \lambda_i^2(x)\right)^{1/2} \leq \sum_{i=1}^d \lambda_i(x) \leq \sqrt{d}\left(\sum_{i=1}^d \lambda_i^2(x)\right)^{1/2} = \sqrt{d}\|x\|.$$

Acknowledgements

The authors would like to thank the referees for a careful reading of this work, useful comments, and pointing out the reference [41]. Peng Jin is supported by the STU Scientific Research Foundation for Talents (No. NTF18023) and NFSC (No. 11861029).

References

[1] Alfonsi, A., Kebaier, A. and Rey, C. (2016). Maximum likelihood estimation for Wishart processes. *Stoch. Proc. Appl.* **126**, 3243–3282.

[2] Andersen, L. B. G. and Piterbarg, V. V. (2007). Moment explosions in stochastic volatility models. *Finance Stoch.* **11**, 29–50.

[3] Baldeaux, J. and Platen, E. (2013). *Functionals of Multidimensional Diffusions with Applications to Finance* (Bocconi & Springer Series **5**). Springer, Cham.

[4] Barczy, M., Döring, L., Li, Z. and Pap, G. (2014). Stationarity and ergodicity for an affine two-factor model. *Adv. Appl. Prob.* **46**, 878–898.

[5] Barczy, M., Li, Z. and Pap, G. (2015). Stochastic differential equation with jumps for multi-type continuous state and continuous time branching processes with immigration. *Lat. Am. J. Prob. Math. Statist.* **12**, 129–169.

[6] Barczy, M., Li, Z. and Pap, G. (2016). Moment formulas for multitype continuous state and continuous time branching process with immigration. *J. Theoret. Prob.* **29**, 958–995.
[7] Barndorff-Nielsen, O. E. and Stelzer, R. (2007). Positive-definite matrix processes of finite variation. *Prob. Math. Statist.* **27**, 3–43.
[8] Behme, A., Lindner, A. and Maller, R. (2011). Stationary solutions of the stochastic differential equation $dV_t = V_{t-} dU_t + dL_t$ with Lévy noise. *Stoch. Proc. Appl.* **121**, 91–108.
[9] Biagini, F., Gnoatto, A. and Härtel, M. (2018). Long-term yield in an affine HJM framework on S_d^+. *Appl. Math. Optimization* **77**, 405–441.
[10] Bru, M.-F. (1991). Wishart processes. *J. Theoret. Prob.* **4**, 725–751.
[11] Chiarella, C., Da Fonseca, J. and Grasselli, M. Pricing range notes within Wishart affine models. *Insurance Math. Econom.* **58**, 193–203.
[12] Cuchiero, C., Filipović, D., Mayerhofer, E. and Teichmann, J. (2011). Affine processes on positive semidefinite matrices. *Ann. Appl. Prob.* **21**, 397–463.
[13] Cuchiero, C., Keller-Ressel, M., Mayerhofer, E. and Teichmann, J. (2016). Affine processes on symmetric cones. *J. Theoret. Prob.* **29**, 359–422.
[14] Da Fonseca, J., Grasselli, M. and Tebaldi, C. (2007). Option pricing when correlations are stochastic: an analytical framework. *Rev. Derivatives Res.* **10**, 151–180.
[15] Dawson, D. A. and Li, Z. (2006). Skew convolution semigroups and affine Markov processes. *Ann. Prob.* **34**, 1103–1142.
[16] Duffie, D., Filipović, D. and Schachermayer, W. (2003). Affine processes and applications in finance. *Ann. Appl. Prob.* **13**, 984–1053.
[17] Filipović, D. and Mayerhofer, E. (2009). Affine diffusion processes: theory and applications. In *Advanced Financial Modelling*, Walter de Gruyter, Berlin, pp. 125–164.
[18] Friesen, M., Jin, P., Kremer, J. and Rüdiger, B. (2019). Exponential ergodicity for stochastic equations of nonnegative processes with jumps. Preprint. Available at https://arxiv.org/abs/1902.02833.
[19] Friesen, M., Jin, P. and Rüdiger, B. (2019). Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes. To appear in *Ann. Appl. Prob.* Preprint available at https://arxiv.org/abs/1901.05815.
[20] Glasserman, P. and Kim, K.-K. (2010). Moment explosions and stationary distributions in affine diffusion models. *Math. Finance* **20**, 1–33.
[21] Gnoatto, A. (2012). The Wishart short rate model. *Internat. J. Theoret. Appl. Finance* **15**, 1–24.
[22] Gnoatto, A. and Grasselli, M. (2014). An affine multicurrency model with stochastic volatility and stochastic interest rates. *SIAM J. Financial Math.* **5**, 493–531.
[23] Gourieroux, C. and Sufana, R. (2010). Derivative pricing with Wishart multivariate stochastic volatility. *J. Business Econom. Statist.* **28**, 438–451.
[24] Gourieroux, C. and Sufana, R. (2011). Discrete time Wishart term structure models. *J. Econom. Dynam. Control* **35**, 815–824.
[25] Grasselli, M. and Miglietta, G. (2016). A flexible spot multiple-curve model. *Quant. Finance* **16**, 1465–1477.
[26] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Studies* **6**, 327–343.
[27] Horn, R. A. and Johnson, C. R. (2013). *Matrix Analysis*, 2nd edn. Cambridge University Press.
[28] Jena, R. P., Kim, K.-K. and Xing, H. (2012). Long-term and blow-up behaviors of exponential moments in multi-dimensional affine diffusions. *Stoch. Proc. Appl.* **122**, 2961–2993.
[29] Jin, P., Kremer, J. and Rüdiger, B. (2017). Exponential ergodicity of an affine two-factor model based on the α-root process. *Adv. Appl. Prob.* **49**, 1144–1169.
[30] Jin, P., Kremer, J. and Rüdiger, B. (2020). Existence of limiting distribution for affine processes. *J. Math. Anal. Appl.* **486**, 123912.
[31] Kawazu, K. and Watanabe, S. (1971). Branching processes with immigration and related limit theorems. *Teor. Veroyat. Primen.* **16**, 34–51.
[32] Keller-Ressel, M. (2011). Moment explosions and long-term behavior of affine stochastic volatility models. *Math. Finance* **21**, 73–98.
[33] Keller-Ressel, M. and Mayerhofer, E. (2015). Exponential moments of affine processes. *Ann. Appl. Prob.* **25**, 714–752.
[34] Keller-Ressel, M. and Mijatović, A. (2012). On the limit distributions of continuous-state branching processes with immigration. *Stoch. Proc. Appl.* **122**, 2329–2345.
[35] Keller-Ressel, M., Schachermayer, W. and Teichmann, J. (2011). Affine processes are regular. *Prob. Theory Rel. Fields* **151**, 591–611.
[36] Keller-Ressel, M. and Steiner, T. (2008). Yield curve shapes and the asymptotic short rate distribution in affine one-factor models. *Finance Stoch.* **12**, 149–172.

[37] KEVEI, P. (2018). Ergodic properties of generalized Ornstein–Uhlenbeck processes. *Stoch. Proc. Appl.* **128**, 156–181.
[38] LEIPPOLD, M. AND TROJANI, F. (2010). Asset pricing with matrix jump diffusions. Working paper.
[39] LI, Z. (2011). *Measure-Valued Branching Markov Processes*. Springer, Heidelberg.
[40] LI, Z. AND MA, C. (2015). Asymptotic properties of estimators in a stable Cox–Ingersoll–Ross model. *Stoch. Proc. Appl.* **125**, 3196–3233.
[41] MASUDA, H. (2004). On multidimensional Ornstein–Uhlenbeck processes driven by a general Lévy process. *Bernoulli* **10**, 97–120.
[42] MAYERHOFER, E. (2012). Affine processes on positive semidefinite $d \times d$ matrices have jumps of finite variation in dimension $d > 1$. *Stoch. Proc. Appl.* **122**, 3445–3459.
[43] MAYERHOFER, E., PFAFFEL, O. AND STELZER, R. (2011). On strong solutions for positive definite jump diffusions. *Stoch. Proc. Appl.* **121**, 2072–2086.
[44] MAYERHOFER, E., STELZER, R. AND VESTWEBER, J. (2020). Geometric ergodicity of affine processes on cones. *Stoch. Proc. Appl.* **130**, 4141–4173.
[45] PIGORSCH, C. AND STELZER, R. (2009). On the definition, stationary distribution and second order structure of positive semidefinite Ornstein–Uhlenbeck type processes. *Bernoulli* **15**, 754–773.
[46] PINSKY, M. A. (1972). Limit theorems for continuous state branching processes with immigration. *Bull. Amer. Math. Soc.* **78**, 242–244.
[47] SATO, K. AND YAMAZATO, M. (1984). Operator-self-decomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type. *Stoch. Proc. Appl.* **17**, 73–100.
[48] VILLANI, C. (2009). *Optimal Transport: Old and New*. Springer, Berlin, Heidelberg.