Stationary covariance regime for affine stochastic covariance models in Hilbert spaces

Martin Friesen[1] · Sven Karbach[2]

Abstract

This paper introduces stochastic covariance models in Hilbert spaces with stationary affine instantaneous covariance processes. We explore the applications of these models in the context of forward curve dynamics within fixed-income and commodity markets. The affine instantaneous covariance process is defined on positive self-adjoint Hilbert–Schmidt operators, and we prove the existence of a unique limit distribution for subcritical affine processes, provide convergence rates of the transition kernels in the Wasserstein distance of order $p \in [1, 2]$, and give explicit formulas for the first two moments of the limit distribution. Our results allow us to introduce affine stochastic covariance models in the stationary covariance regime and to investigate the behaviour of the implied forward volatility for large forward dates in commodity forward markets.

Keywords Affine processes · Invariant measure · Stationarity · Ergodicity · Stochastic covariance · Implied forward volatility · Generalised Feller semigroups

Mathematics Subject Classification 60J25 · 37A25 · 60G10

JEL Classification C60 · G12 · G13

1 Introduction

1.1 Motivation

Affine processes are highly valued in the field of mathematical finance thanks to their exceptional analytical tractability, making them a popular choice for stochastic modelling. In the last two decades, numerous authors have studied affine processes and their applications on various state spaces, including the canonical state

✉ S. Karbach
  sven@karbach.org

1   School of Mathematical Sciences, Dublin City University, Dublin 9, Ireland

2   Kortweg-de Vries Institute for Mathematics & Informatics Institute, University of Amsterdam, LAB42 - Science Park 900, 1098 XH, Amsterdam, The Netherlands

space $\mathbb{R}_+^d \times \mathbb{R}^n$ studied in Duffie et al. [25], Keller-Ressel [48, Part I], Dawson and Li [23], Filipović and Mayerhofer [28], Keller-Ressel and Mayerhofer [50], Kallsen and Muhle-Karbe [44], and the cone of symmetric positive semi-definite $(d \times d)$-matrices $\mathbb{S}_+^d$ studied in Cuchiero et al. [21], Cuchiero [20, Chap. 5]. Affine diffusions on canonical state spaces in infinite-dimensional Hilbert spaces have been studied in Schmidt et al. [64], Yu [68, Chap. 2] with applications to interest rate modelling. Moreover, infinite-dimensional affine processes also emerge in Cuchiero and Teichmann [22] as Markovian lifts of affine Volterra processes, which are used to model rough volatility. Besides the general framework discussed therein, examples of Markovian lifts are either constructed on a space of signed measures or on the Filipović space which also plays a central role for our applications; see also Sect. 4.

In this work, we focus on infinite-dimensional affine stochastic covariance models that are used to describe the forward curve dynamics in fixed-income or commodity markets in Benth et al. [9], Benth and Simonsen [11], Benth and Sgarra [10], Benth et al. [8], Cox et al. [18], Karbach [46, Part I]. These models are developed within the Heath, Jarrow, Morton and Musiela (HJMM) framework in Carmona and Tehranchi [13] and Benth and Krühner [7], and consist of a hyperbolic stochastic partial differential equation that describes the evolution of the forward curve, an inherently infinite-dimensional object modelled in the Filipović space, and a separate operator-valued *instantaneous covariance process* that captures the covariance structure of the forward curve dynamics.

The main challenge in these models lies in specifying the instantaneous covariance process in a way that is both accurate and feasible for tasks in statistical inference, simulation or the pricing of financial derivatives. To address the challenges associated with these models, the analytical tractability of affine processes plays a crucial role. In fact, in finite dimension, the instantaneous covariance process has been modelled as an affine process on symmetric positive semi-definite matrices in Cuchiero et al. [21], Leippold and Trojani [54], Barndorff-Nielsen and Stelzer [4], Gouriéroux and Sufana [36]. Motivated by the need to capture high-dimensional noise, for example to model maturity-specific risk (see Carmona and Tehranchi [13, Chap. 2], Benth and Krühner [7] and Koekebakker and Ollmar [52] for the case of electricity markets), it was proposed in Benth et al. [9] and Cox et al. [17, 18] to model the instantaneous covariance process of the forward curves by affine processes on positive Hilbert–Schmidt operators. These new affine stochastic covariance models provide great flexibility in modelling high-dimensional noise as they allow infinite-rank operator-valued instantaneous covariance while maintaining sufficient tractability due to their affine structure.

Empirical data suggests that stochastic covariance, similarly to stochastic volatility, enjoys the stylised fact of mean-reversion. The latter allows us to define e.g. long-term averages and the speed of mean-reversion for stochastic covariance; see e.g. Teng et al. [65]. Mathematically, mean-reversion is captured by the long-time behaviour, stationarity and ergodicity of the process. In this article, we study these properties for the class of affine processes on positive Hilbert–Schmidt operators and henceforth provide a mathematical justification of mean-reversion. Additionally to the study of mean-reversion, stationarity also plays an important role in parameter estimation and model calibration of the instantaneous covariance process as noted

in Alfonsi et al. [1]. Ergodicity, on the other hand, is critical for the L^1-convergence of the general method of moments in stochastic covariance models, as discussed in Pigorsch and Stelzer [62]. In the broader picture, a better understanding of the long-time behaviour of forward curve models is also critical for managing the risk associated with portfolios of forward contracts and evaluating long-term investment decisions. In the following, we introduce our setting and present our main results more formally.

1.2 Main result and applications

Following the literature, we realise the evolution of the forward curves as a function-valued process with values in some Filipović space; see Sect. 4 below. Thus let us consider a real and separable Hilbert space $(H, (\cdot, \cdot))$. Accordingly, we model the instantaneous covariance process of the forward curve dynamics on $\mathcal{H}_+$, the set of all positive self-adjoint Hilbert–Schmidt operators defined on H. Note that $\mathcal{H}_+$ is a convex cone in the Hilbert space of all self-adjoint Hilbert–Schmidt operators equipped with the trace inner product $\langle x, y \rangle := \mathrm{Tr}(yx)$ for $x, y \in \mathcal{H}$. Given a time-homogeneous Markov process $(X_t)_{t \geq 0}$ with values in $\mathcal{H}_+$, we denote its transition kernels by $(p_t(x, \cdot))_{t \geq 0}$, and following the general terminology of affine processes, we call $(X_t)_{t \geq 0}$ *affine* whenever the Laplace transform of X_t, for every $t \geq 0$, is of an exponential-affine form in the initial value $x \in \mathcal{H}_+$, i.e.,

$$\int_{\mathcal{H}_+} \mathrm{e}^{-\langle \xi, u \rangle} p_t(x, \mathrm{d}\xi) = \mathrm{e}^{-\phi(t,u) - \langle x, \psi(t,u) \rangle}, \qquad t \geq 0, u \in \mathcal{H}_+, \qquad (1.1)$$

for some functions $\phi \colon \mathbb{R}_+ \times \mathcal{H}_+ \to \mathbb{R}_+$ and $\psi \colon \mathbb{R}_+ \times \mathcal{H}_+ \to \mathcal{H}_+$. Typically, the functions ϕ and ψ are solutions of an associated pair of *generalised Riccati equations* that can be solved numerically. Affine processes on the state-space $\mathcal{H}_+$, as considered in this work, have been first introduced and studied in Cox et al. [17]. Their construction is summarised in Theorem 2.2 below; see also [17, Theorem 2.8]. We want to emphasise that such processes can be considered as the natural infinite-dimensional analogues of affine processes on symmetric positive semi-definite $(d \times d)$-matrices $\mathbb{S}_+^d$ studied in Cuchiero et al. [21]. Indeed, for $H = \mathbb{R}^d$, the self-adjoint Hilbert–Schmidt operators on H are precisely the symmetric matrices equipped with the Frobenius norm, and $\mathcal{H}_+ = \mathbb{S}_+^d$ holds in this case.

To carry out our study of long-time behaviour for affine stochastic covariance models, we first prove that the transition kernels $(p_t(x, \cdot))_{t \geq 0}$ of the affine instantaneous covariance process converge weakly to some probability measure π on $\mathcal{H}_+$ as $t \to \infty$. This limit distribution is then shown to be the unique invariant measure of the process. Afterwards, we provide an exponential rate of convergence in the Wasserstein distance of order $p \in [1, 2]$ and finally construct the corresponding stationary process that has the invariant measure π as its time-marginals. Below, we provide a short version of our main result. We call an affine process on $\mathcal{H}_+$ *subcritical* if its state-dependent drift is negative (i.e., all eigenvalues have strictly negative real parts).

Theorem 1.1 *Let $(X_t)_{t \geq 0}$ be a subcritical affine process on the cone $\mathcal{H}_+$ with transition kernels $(p_t(x, \cdot))_{t \geq 0}$ and the existence of which is guaranteed by Theorem* 2.2 *below. Then the following hold true*:

i) *There exists a unique invariant measure π of $(p_t(x, \cdot))_{t \geq 0}$ for all $x \in \mathcal{H}_+$.*

ii) *For every $x \in \mathcal{H}_+$ and $p \in [1, 2]$, the sequence $(p_t(x, \cdot))_{t \geq 0}$ converges exponentially fast to π in the Wasserstein distance of order p as $t \to \infty$.*

iii) *There exists a Markov process $(X_t^\pi)_{t \geq 0}$ with transition kernels $(p_t(x, \cdot))_{t \geq 0}$ such that the distribution of X_t^π is equal to π for all $t \geq 0$.*

In Sect. 1.3 below, we give an in-depth overview of our methodology and the relevance of our work in the context of the existing literature and methods. In the following, we emphasise the relevance of our results for applications in stochastic covariance modelling of forward curve dynamics.

Indeed, having understood the long-time behaviour of affine instantaneous covariance processes of Hilbert-valued stochastic covariance models, we can introduce and study these models in the *stationary covariance regime*, where we replace the affine instantaneous covariance process by its corresponding stationary affine process having the same transition probabilities and whose existence is guaranteed by part iii) of the result above (which is Proposition 3.5 below). The notion of the stationary covariance regime is inspired by the univariate case, i.e., the *stationary variance regime* in Keller-Ressel [49]. In Proposition 4.2 below, we first show that the affine stochastic covariance model in the stationary covariance regime satisfies an affine transform formula, making it a tractable model for e.g. the pricing of options on forwards; see Sect. 4.3 below. Since the operator Barndorff-Nielsen–Shephard (BNS) model of Benth et al. [9] is known to be included in the affine class, we derive the characteristic function of the operator BNS model in the stationary covariance regime. This complements the literature on infinite-dimensional BNS-type models by their long-time behaviour, extending the finite-dimensional matrix-valued case in Barndorff-Nielsen and Stelzer [4], Pigorsch and Stelzer [62].

As an application of such stochastic covariance models in the stationary regime, we study in Sect. 4.3 the implied forward volatility for forward-start options written on commodity forwards. In particular, we examine the implied forward volatility in the geometric affine stochastic covariance model and show in Proposition 4.4 below that it converges to the implied spot volatility of a European call option written on the forward, but this time modelled in the stationary covariance regime. This extends a result in Keller-Ressel [49, Proposition 5.2] for forward-start options in (univariate) affine stochastic volatility models to an infinite-dimensional setting.

1.3 Related literature and methodology

The long-time behaviour of affine processes on the finite-dimensional state spaces $\mathbb{R}_+^d \times \mathbb{R}^n$ and $\mathbb{S}_+^d$ for $d, n \in \mathbb{N}$ is now mostly well understood. More precisely, based on the representation by strong solutions of stochastic differential equations, ergodicity was studied for different Wasserstein distances in Friesen et al. [34]. By using the regularity of transition densities with respect to Lebesgue measure combined with the Meyn–Tweedie stability theory, the ergodicity in total variation distances has been

studied in Barczy et al. [3], Friesen and Jin [31], Jin et al. [41, 42], Mayerhofer et al. [60], Friesen et al. [33]. Finally, coupling techniques for affine processes are studied in Wang [67], Li and Ma [58]. Unfortunately, all these methods implicitly either use the dimension of the state space or are closely related to the particular structure of the state space and hence do not allow an extension to our infinite-dimensional setting. Indeed, for general affine processes on $\mathcal{H}_+$, a pathwise construction is still absent from the literature which rules out the methods developed in Friesen et al. [34]. The absence of an infinite-dimensional Lebesgue measure prevents us from effectively using the Meyn–Tweedie stability theory (in terms of estimates of the density). Finally, although there exist some extensions of the coupling techniques to infinite-dimensional state-spaces (see Li [57] for measure-valued branching processes), these methods seem to be closely related to the measure-valued structure of the process and hence are not suitable for our Hilbert space framework.

The most promising tool for the study of the long-time behaviour for affine processes in infinite-dimensional settings is thus based on the convergence of Fourier–Laplace transforms. This requires, in view of the affine transform formula (1.1), to study stability for the solutions ϕ, ψ of the generalised Riccati equations. For finite-dimensional state spaces, these ideas have been developed in Glasserman and Kim [35], Keller-Ressel [49], Keller-Ressel and Mijatović [51], Jin et al. [43], Pigorsch and Stelzer [62], Friesen et al. [32]. In these works, the existence of an invariant distribution (as well as weak convergence of transition probabilities) is obtained from Lévy's continuity theorem. More recently, such techniques also have been applied in Friesen [30] to Dawson–Watanabe superprocesses with immigration, which form a specific class of affine processes on the state space of (possibly tempered) measures (so-called measure-valued Markov processes).

Unfortunately, in infinite-dimensional settings, there exists no direct analogue of Lévy's continuity theorem since additional technical tightness conditions need to be verified. For Ornstein–Uhlenbeck processes on Hilbert spaces, such a problem can be avoided by taking advantage of their infinite divisibility; see Chojnowska-Michalik [15]. Note that such processes form a subclass of affine processes; see also Example 3.7 below. Apart from this, the long-time behaviour of affine processes in infinite dimensions has not been investigated in a systematic way. The present work provides a first general treatment of this topic, while our methodology also allows the treatment of other infinite-dimensional models being so far out of reach.

Namely, as a first step, we construct a candidate for the invariant measure π as a weak limit of the transition probabilities when $t \to \infty$, and show that this limit is indeed the unique invariant measure of the process. To this end, we follow ideas taken from Friesen et al. [32] where the state-space $\mathbb{S}_+^d$ was studied and show that for subcritical affine processes, the limits $\lim_{t\to\infty} \phi(t, u)$ and $\lim_{t\to\infty} \psi(t, u)$ exist for every $u \in \mathcal{H}_+$. Consequently, the Fourier–Laplace transform of the process (see (1.1)) converges when $t \to \infty$. In contrast to the finite-dimensional case, this alone does not suffice to obtain the existence of a limit distribution π. As a key tool, we propose to utilise the generalised Feller semigroup approach for the process (see Appendix A for a definition). More precisely, we provide uniform bounds on the operator norm of the transition semigroup $(P_t)_{t\geq 0}$ which allows us to prove that $\lim_{t\to\infty} P_t f =: \ell(f)$ exists for a sufficiently large class of functions f. By

showing that the limit ℓ is a continuous linear functional, we can apply a variant of Riesz' representation theorem for generalised Feller semigroups to show that ℓ has a representation $\ell(f) = \int_{\mathcal{H}_+} f(y)\pi(\mathrm{d}y)$. The measure π obtained in this way is then shown to be the desired unique invariant probability measure. As a by-product, we also obtain weak convergence of the transition probabilities. Due to the infinite-dimensional nature of the process and the use of generalised Feller semigroups, such a convergence a priori only holds in the weak topology on $\mathcal{H}_+$.

In the second step, we strengthen the above convergence from the weak topology to the norm topology by establishing bounds on the Wasserstein distance of order $p \in [1, 2]$ for the transition probabilities with respect to the invariant measure π. In contrast to the finite-dimensional results in Friesen et al. [34, 32], our new bounds are dimension-free while the convergence rate is shown to be still exponential. Finally, the corresponding stationary process is obtained via Kolmogorov's consistency theorem which is applicable since π is shown to be inner regular (a condition that is always satisfied when working in finite-dimensional cases). As a side product of our considerations, we also obtain that the first two moments of the process converge to the first two moments of the invariant measure. Using generalised Feller semigroups, we provide explicit formulas for these.

We should like to mention that our proposed method for the study of the long-time behaviour of stochastic processes on Hilbert spaces via the generalised Feller semigroup can also be applied to other settings. For instance, in the recent work Jacquier et al. [38] that appeared at the same time as this article, the authors established a Krylov–Bogoliubov-type theorem for generalised Feller semigroups and subsequently applied it to multi-dimensional rough Heston models. Their key method, a uniform bound on the operator norm of P_t, also appears as an auxiliary step in our result (see Lemma 5.5). However, our approach does not require such a Krylov–Bogoliubov-type result and also allows us to prove the uniqueness of limit distributions. Following Cuchiero and Teichmann [22] and Benth et al. [6], general (affine) Volterra processes can be realised as Markovian lifts on infinite-dimensional spaces including the Filipović space. Therefore, the methods of this work will be also useful for the study of such Markovian lifts associated with affine Volterra processes.

1.4 Layout of the article

In Sect. 2, we introduce the class of affine processes on $\mathcal{H}_+$ established in Cox et al. [17] and recall some preliminary results. Subsequently, in Sect. 3, we present and discuss our main results in full detail. Afterwards, in Sect. 4, we discuss applications of our results in the context of affine stochastic covariance models for forward curves. Finally, the proofs are contained in Sect. 5 which is subdivided into several subsections: We first consider the long-time behaviour of the solutions of the generalised Riccati equations in Sect. 5.1, then prove the existence of a unique invariant measure in Sect. 5.2, derive the convergence rates in Sect. 5.3, show the existence of stationary affine processes in Sect. 5.4 and finally prove the moment formulas of the invariant measure in Sect. 5.5. We end the article with a brief conclusion and outlook in Sect. 6. For the reader's convenience, we added some background information on generalised Feller semigroups in Appendix A, where we give in particular a version

of Kolmogorov's extension theorem that is tailored to our needs. In Appendix B, we give a convolution property for the Wasserstein distance of order 2.

2 Preliminaries: affine processes on $\mathcal{H}_+$

We set $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For a complex number $z \in \mathbb{C}$, we denote its real and imaginary parts by $\Re(z)$ and $\Im(z)$. For a vector space X and a subset $U \subseteq X$, we denote the linear span of U in X by $\mathrm{lin}(U)$. For a topological vector space (X, τ) and a subset $S \subseteq X$, we denote the *Borel-σ-algebra* generated by the relative topology on S by $\mathcal{B}(S)$. We write $C_b(S)$ for the space of real-valued bounded functions on S that are continuous with respect to the relative topology; then $C_b(S)$ is a Banach space when endowed with the supremum norm $\| \cdot \|_{C(S)}$. For a Banach space X with norm $\| \cdot \|_X$, we denote by $\mathcal{L}(X)$ the space of all *bounded linear operators* on X, which becomes a Banach space when equipped with the operator norm $\| \cdot \|_{\mathcal{L}(X)}$.

Throughout this article, we let $(H, \langle \cdot, \cdot \rangle_H)$ be a separable real Hilbert space and denote by $\mathcal{H}$ the set of all self-adjoint *Hilbert–Schmidt operators* from H to H. This is a Hilbert space when endowed with the *trace inner* product

$$\langle A, B \rangle = \sum_{n=1}^{\infty} \langle A f_n, B f_n \rangle_H$$

for $A, B \in \mathcal{H}$ and where $(f_n)_{n \in \mathbb{N}}$ is an orthonormal basis for H. Note that $\langle \cdot, \cdot \rangle$ is independent of the choice of the orthonormal basis. We denote by $\| \cdot \|$ the norm on $\mathcal{H}$ induced by $\langle \cdot, \cdot \rangle$. In addition, we denote by $\mathcal{H}_+$ the set of all positive operators in $\mathcal{H}$, i.e., $\mathcal{H}_+ := \{A \in \mathcal{H} \colon \langle Ah, h \rangle_H \geq 0 \text{ for all } h \in H\}$. Note that $\mathcal{H}_+$ is a closed subset of $\mathcal{H}$. Moreover, it is a convex cone in $\mathcal{H}$, i.e., $\mathcal{H}_+ + \mathcal{H}_+ \subseteq \mathcal{H}_+$, $\lambda \mathcal{H}_+ \subseteq \mathcal{H}_+$ for all $\lambda \geq 0$, and $\mathcal{H}_+ \cap (-\mathcal{H}_+) = \{0\}$. The cone $\mathcal{H}_+$ induces a partial ordering $\leq_{\mathcal{H}_+}$ on $\mathcal{H}$, which is defined by $x \leq_{\mathcal{H}_+} y$ whenever $y - x \in \mathcal{H}_+$. The cone $\mathcal{H}_+$ is also *generating* for $\mathcal{H}$, i.e., $\mathcal{H} = \mathcal{H}_+ - \mathcal{H}_+$. For a Hilbert space $(V, \langle \cdot, \cdot \rangle_V)$, in this article either $(H, \langle \cdot, \cdot \rangle_H)$ or $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, we denote the *adjoint* of $A \in \mathcal{L}(V)$ by A^*. For two elements x and y in V, we define the operator $x \otimes y \in \mathcal{L}(V)$ by $(x \otimes y)(h) = \langle x, h \rangle_V y$ for every $h \in V$ and write $x^{\otimes 2} := x \otimes x$.

2.1 Admissible parameters and discussion

Let $\chi \colon \mathcal{H} \to \mathcal{H}$ be given by $\chi(\xi) = \xi \mathbf{1}_{\|\xi\| \leq 1}(\xi)$. The following definition of an *admissible parameter set* stems from Cox et al. [17, Definition 2.3] and is closely related to the finite-dimensional version, i.e., *matrix-valued admissible parameter sets* known from Cuchiero et al. [21, Definition 2.3].

Definition 2.1 An *admissible parameter set* (b, B, m, μ) consists of
   i) a measure $m \colon \mathcal{B}(\mathcal{H}_+ \setminus \{0\}) \to [0, \infty]$ such that
   (a) $\int_{\mathcal{H}_+ \setminus \{0\}} \|\xi\|^2 m(\mathrm{d}\xi) < \infty$,
   (b) $\int_{\mathcal{H}_+ \setminus \{0\}} |\langle \chi(\xi), h \rangle| m(\mathrm{d}\xi) < \infty$ for all $h \in \mathcal{H}$ and there exists an element $I_m \in \mathcal{H}$ such that $\langle I_m, h \rangle = \int_{\mathcal{H}_+ \setminus \{0\}} \langle \chi(\xi), h \rangle m(\mathrm{d}\xi)$ for every $h \in \mathcal{H}$;

ii) a vector $b \in \mathcal{H}$ such that

$$\langle b, v \rangle - \int_{\mathcal{H}_+ \setminus \{0\}} \langle \chi(\xi), v \rangle m(\mathrm{d}\xi) \geq 0 \qquad \text{for all } v \in \mathcal{H}_+;$$

iii) an $\mathcal{H}_+$-valued measure $\mu \colon \mathcal{B}(\mathcal{H}_+ \setminus \{0\}) \to \mathcal{H}_+$ such that the kernel $M(x, \mathrm{d}\xi)$, defined for every $x \in \mathcal{H}_+$ on $\mathcal{B}(\mathcal{H}_+ \setminus \{0\})$ by

$$M(x, \mathrm{d}\xi) := \frac{\langle x, \mu(\mathrm{d}\xi) \rangle}{\|\xi\|^2}, \tag{2.1}$$

satisfies for all $u, x \in \mathcal{H}_+$ with $\langle u, x \rangle = 0$ that

$$\int_{\mathcal{H}_+ \setminus \{0\}} \langle \chi(\xi), u \rangle M(x, \mathrm{d}\xi) < \infty;$$

iv) an operator $B \in \mathcal{L}(\mathcal{H})$ with adjoint B^* satisfying for all $x, u \in \mathcal{H}_+$ with $\langle u, x \rangle = 0$ that

$$\langle B^*(u), x \rangle - \int_{\mathcal{H}_+ \setminus \{0\}} \langle \chi(\xi), u \rangle \frac{\langle \mu(\mathrm{d}\xi), x \rangle}{\|\xi\|^2} \geq 0. \tag{2.2}$$

In comparison to the matrix-valued affine processes in Cuchiero et al. [21], the class of affine processes studied in this article is of pure-jump type and lacks a diffusion coefficient in the admissible parameter set in the sense of Definition 2.1. This is motivated by the fact that general affine diffusions on the cone of positive Hilbert–Schmidt operators are necessarily finite-rank-valued, as can be conjectured from the matrix-valued case. Indeed, for the matrix-valued case, the diffusion coefficient α must satisfy the inequality $b \succeq_{\mathbb{S}^d_+} (d-1)\alpha$ with b being the drift coefficient and $d \in \mathbb{N}$ the dimension of the matrix, i.e., its maximal rank. When the rank approaches infinity, the diffusion coefficient must vanish to meet this requirement. In future work, this argument will be made precise and discussed in the context of Wishart processes on positive Hilbert–Schmidt operators. Thus affine processes with nonzero diffusion components on $\mathcal{H}_+$ are finite-rank-valued and on a fixed orthonormal basis behave very similarly to their matrix-valued counterparts. For this reason, we restrict our considerations to the case of vanishing diffusion components with general jumps so that no rank conditions need to be imposed.

Consequently, when assuming that the diffusion component is absent, randomness solely enters through the parameters m and μ. Here m is called *constant* and μ is called *linear jump coefficient*, respectively. Note that m is a standard measure defined on the Borel-σ-algebra on $\mathcal{H}_+ \setminus \{0\}$, whereas μ is a vector-valued measure. For a comprehensive introduction to the integration theory of vector-valued measures, we refer to Bartle et al. [5] and Lewis [56].

Note that by part (b) of Definition 2.1 i), we have

$$\int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \|\xi\| m(\mathrm{d}\xi) \leq \int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \|\xi\|^2 m(\mathrm{d}\xi) < \infty,$$

which means that the integral $\int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \xi\, m(\mathrm{d}\xi)$ is defined in the Bochner sense. Similarly, it can be seen that the map $u \mapsto \int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \langle \xi, u \rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}$ is a bounded linear operator on $\mathcal{H}$. Indeed, for $v \in \mathcal{H}$ such that $v = v^+ - v^-$ for $v^+, v^- \in \mathcal{H}_+$, we write $|\langle \mu(\mathrm{d}\xi), v \rangle| := \langle \mu(\mathrm{d}\xi), v^+ \rangle + \langle \mu(\mathrm{d}\xi), v^- \rangle$ and see that $|\langle \mu(\mathrm{d}\xi), v \rangle|$ is a positive measure for all $v \in \mathcal{H}$. We thus have

$$\left\langle \int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \langle \xi, u \rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}, v \right\rangle \leq \|u\| \left(\int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \|\xi\|^{-1} |\langle \mu(\mathrm{d}\xi), v \rangle| \right)$$

$$\leq \|u\| |\langle \mu(\mathcal{H}_+ \cap \{\|\xi\| > 1\}), v \rangle|,$$

and taking the supremum over all $v \in \mathcal{H}$ with $\|v\| = 1$ on both sides proves that the map is bounded. Note that $\|\mu(\mathcal{H}_+ \cap \{\|\xi\| > 1\})\| < \infty$ since by Definition 2.1 iii), we have $\mu(A) \in \mathcal{H}_+$ and hence $\|\mu(A)\| < \infty$ for every $A \in \mathcal{B}(\mathcal{H}_+ \setminus \{0\})$.

One noticeable difference to the matrix-valued case in Cuchiero et al. [21] is that while the diffusion component is absent, the jump measure $M(x, \mathrm{d}\xi)$ in (2.1) can have infinite variation. The latter is actually impossible in the matrix-valued case as shown in Mayerhofer [59]. Thus the jump components are more general than their finite-dimensional counterparts; see Karbach [47, Sect. 3.4] for a concrete example of an affine process of infinite variation.

2.2 Existence and moment formulas

We recall the main result in Cox et al. [17, Theorem 2.8] which ensures the existence of a broad class of affine processes on $\mathcal{H}_+$ associated with the admissible parameter set (b, B, m, μ) satisfying the conditions in Definition 2.1.

Theorem 2.2 *Let (b, B, m, μ) be an admissible parameter set according to Definition 2.1. Then there exist a conservative time-homogeneous $\mathcal{H}_+$-valued Markov process X with transition kernels $(p_t(x, \cdot))_{t \geq 0}$ and constants $K, \omega \in [1, \infty)$ such that*

$$\int_{\mathcal{H}_+} \|\xi\|^2 p_t(x, \mathrm{d}\xi) \leq K e^{\omega t} (\|x\|^2 + 1)$$

and for all $t \geq 0$ and $u, x \in \mathcal{H}_+$, we have

$$\int_{\mathcal{H}_+} e^{-\langle \xi, u \rangle} p_t(x, \mathrm{d}\xi) = e^{-\phi(t,u) - \langle x, \psi(t,u) \rangle}, \tag{2.3}$$

where $\phi(\cdot, u), \psi(\cdot, u)$ are the unique solutions to the generalised Riccati equations

$$\frac{\partial \phi(t, u)}{\partial t} = F\big(\psi(t, u)\big), \quad t > 0, \qquad \phi(0, u) = 0, \tag{2.4}$$

$$\frac{\partial \psi(t, u)}{\partial t} = R\big(\psi(t, u)\big), \quad t > 0, \qquad \psi(0, u) = u, \tag{2.5}$$

where $F: \mathcal{H}_+ \to \mathbb{R}$ and $R: \mathcal{H}_+ \to \mathcal{H}$ are given by

$$F(u) = \langle b, u \rangle - \int_{\mathcal{H}_+ \setminus \{0\}} \left(e^{-\langle \xi, u \rangle} - 1 + \langle \chi(\xi), u \rangle \right) m(\mathrm{d}\xi), \qquad (2.6)$$

$$R(u) = B^*(u) - \int_{\mathcal{H}_+ \setminus \{0\}} \left(e^{-\langle \xi, u \rangle} - 1 + \langle \chi(\xi), u \rangle \right) \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}. \qquad (2.7)$$

Define the transition semigroup $(P_t)_{t \geq 0}$ by

$$P_t f(x) = \int_{\mathcal{H}_+} f(\xi) p_t(x, \mathrm{d}\xi)$$

for bounded measurable functions $f : \mathcal{H}_+ \to \mathbb{R}$. Then $(P_t)_{t \geq 0}$ is a positive semigroup. Let $\rho(x) = 1 + \|x\|^2$ and define

$$\|f\|_\rho = \sup_{x \in \mathcal{H}_+} \frac{|f(x)|}{\rho(x)}.$$

Denote by $B_\rho(\mathcal{H}_+)$ the Banach space of all measurable functions $f : \mathcal{H}_+ \to \mathbb{R}$ for which $\|f\|_\rho$ is finite. Clearly, $(P_t)_{t \geq 0}$ extends onto $B_\rho(\mathcal{H}_+)$ and satisfies for each $f \in B_\rho(\mathcal{H}_+)$ that

$$|P_t f(x)| \leq \|f\|_\rho \int_{\mathcal{H}_+} \rho(y) p_t(x, \mathrm{d}y) \leq \|f\|_\rho (1 + K) e^{\omega t} \rho(x), \qquad x \in \mathcal{H}_+, \tag{2.8}$$

i.e., $(P_t)_{t \geq 0}$ leaves $B_\rho(\mathcal{H}_+)$ invariant. Let $\mathcal{H}_{+,\mathrm{w}}$ be the space $\mathcal{H}_+$ equipped with the weak topology and denote by $C_b(\mathcal{H}_{+,\mathrm{w}})$ the space of all bounded and weakly continuous functions $f : \mathcal{H}_+ \to \mathbb{R}$. Finally, let $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ be the closure of $C_b(\mathcal{H}_{+,\mathrm{w}})$ in $B_\rho(\mathcal{H}_+)$. It follows from Cox et al. [17] that $(P_t)_{t \geq 0}$ leaves $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ invariant and satisfies $\lim_{t \to 0+} P_t f(x) = f(x)$ for all $f \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ and $x \in \mathcal{H}_{+,\mathrm{w}}$. From (2.8), we then obtain

$$\|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}))} \leq (1 + K) e^{\omega t}, \qquad t \geq 0.$$

Hence $(P_t)_{t \geq 0}$ is a *generalised Feller semigroup* on $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ (see Appendix A for the definition and additional details).

Remark 2.3 Given the transition kernels $(p_t(x, \cdot))_{t \geq 0}$, the process $(X_t)_{t \geq 0}$ with initial value $X_0 = x \in \mathcal{H}_+$ can be constructed by a version of Kolmogorov's extension theorem in Cuchiero and Teichmann [22, Theorem 2.11]. Indeed, for every $x \in \mathcal{H}_+$, one can show the existence of a unique measure $\mathbb{P}_x$ on $\Omega := (\mathcal{H}_+)^{\mathbb{R}_+}$, equipped with the σ-algebra generated by the canonical projections $X_t : \Omega \to \mathcal{H}$ given by $X_t(\omega) = \omega(t)$ for $\omega \in \Omega$. For $x \in \mathcal{H}_+$, the probability measure $\mathbb{P}_x$ is the distribution of X with $\mathbb{P}_x[X_0 = x] = 1$. We denote the expectation with respect to $\mathbb{P}_x$ by $\mathbb{E}_x[\cdot]$.

Let (b, B, m, μ) be an admissible parameter set according to Definition 2.1 and denote by $(X_t)_{t \geq 0}$ the associated affine process on $\mathcal{H}_+$. Let us define, for $u \in \mathcal{H}$, the constant and linear *effective drift* terms $\hat{b}$ and $\hat{B}$ via

$$\hat{b} := b + \int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \xi\, m(\mathrm{d}\xi),$$

$$\hat{B}(u) := B^*(u) + \int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \langle \xi, u \rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}. \tag{2.9}$$

Then $\hat{b} \in \mathcal{H}$ and $\hat{B} \in \mathcal{L}(\mathcal{H})$ are well defined. Below we state explicit formulas for the first two moments of the process $(X_t)_{t \geq 0}$, which can be derived from differentiating the Laplace transform of $(X_t)_{t \geq 0}$ using its affine form; see Cox et al. [17, Proposition 4.7].

Proposition 2.4 *Let $(X_t)_{t \geq 0}$ be the affine process associated with the admissible parameter set (b, B, m, μ). Then for all $v, w \in \mathcal{H}_+$, we have the formulas*

$$\mathbb{E}_x \left[\langle X_t, v \rangle \right] = \int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} v \rangle\, \mathrm{d}s + \langle x, \mathrm{e}^{t\hat{B}} v \rangle \tag{2.10}$$

and

$$\mathbb{E}_x \left[\langle X_t, v \rangle \langle X_t, w \rangle \right]$$
$$= \left(\int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} v \rangle\, \mathrm{d}s + \langle x, \mathrm{e}^{t\hat{B}} v \rangle \right) \left(\int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} w \rangle\, \mathrm{d}s + \langle x, \mathrm{e}^{t\hat{B}} w \rangle \right)$$
$$+ \int_0^t \int_{\mathcal{H}_+ \setminus \{0\}} \langle \xi, \mathrm{e}^{s\hat{B}} v \rangle \langle \xi, \mathrm{e}^{s\hat{B}} w \rangle\, m(\mathrm{d}\xi)\, \mathrm{d}s$$
$$+ \int_0^t \int_0^s \left\langle \hat{b}, \mathrm{e}^{(s-u)\hat{B}} \int_{\mathcal{H}_+ \setminus \{0\}} \langle \xi, \mathrm{e}^{u\hat{B}} v \rangle \langle \xi, \mathrm{e}^{u\hat{B}} w \rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \right\rangle \mathrm{d}u\, \mathrm{d}s$$
$$+ \int_0^t \left\langle x, \mathrm{e}^{(t-s)\hat{B}} \int_{\mathcal{H}_+ \setminus \{0\}} \langle \xi, \mathrm{e}^{s\hat{B}} w \rangle \langle \xi, \mathrm{e}^{s\hat{B}} v \rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \right\rangle \mathrm{d}s. \tag{2.11}$$

3 Main results

Let $V_\tau := (V, \tau)$ be a topological vector space and denote by $\mathcal{M}(V_\tau)$ the set of all probability measures defined on the Borel-σ-algebra $\mathcal{B}(V_\tau)$. For the vector space $\mathcal{H}$ equipped with its weak topology τ_w, we write $\mathcal{H}_w = (\mathcal{H}, \tau_w)$. Note that the positive cone $\mathcal{H}_{+,w}$ is also closed in the weak topology and, moreover, the Borel-σ-algebras of the strong and weak topology coincide, i.e., $\mathcal{B}(\mathcal{H}_+) = \mathcal{B}(\mathcal{H}_{+,w})$. We say that a measure $\nu \in \mathcal{M}(\mathcal{H}_{+,\tau})$ is *inner regular* (with respect to the topology τ) whenever

$$\nu(A) = \sup \{ \nu(K) \colon K \subseteq A,\ K \text{ is } \tau\text{-compact} \}.$$

For a sequence $(\nu_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}(\mathcal{H}_+)$, we write $\nu_n \Rightarrow \nu$ as $n \to \infty$ for the weak convergence of $(\nu_n)_{n\in\mathbb{N}}$ to ν in the strong topology, i.e.,

$$\lim_{n\to\infty} \int_{\mathcal{H}_+} f(\xi)\nu_n(\mathrm{d}\xi) = \int_{\mathcal{H}_+} f(\xi)\nu(\mathrm{d}\xi) \qquad \text{for all } f \in C_b(\mathcal{H}).$$

For a family $(p_t(x, \cdot))_{t\geq 0}$ of transition kernels, we call the probability measure π an *invariant measure* whenever

$$\int_{\mathcal{H}_+} p_t(x, \mathrm{d}\xi)\pi(\mathrm{d}x) = \pi(\mathrm{d}\xi) \qquad \text{for all } t \geq 0 \text{ and } x \in \mathcal{H}_+.$$

For $\nu_1, \nu_2 \in \mathcal{M}(\mathcal{H}_+)$, we call a probability measure G on the product Borel-σ-algebra $\mathcal{B}(\mathcal{H}_+) \times \mathcal{B}(\mathcal{H}_+)$ a *coupling* of (ν_1, ν_2) whenever its marginal distributions are given by ν_1 and ν_2, respectively. We denote the set of all possible couplings of (ν_1, ν_2) by $\mathcal{C}(\nu_1, \nu_2)$. For $p \in [1, \infty)$, the *Wasserstein distance of order p* between $\nu_1 \in \mathcal{M}(\mathcal{H}_+)$ and $\nu_2 \in \mathcal{M}(\mathcal{H}_+)$ is defined as

$$W_p(\nu_1, \nu_2) = \left(\inf\left\{ \int_{\mathcal{H}_+\times\mathcal{H}_+} \|x - y\|^p G(\mathrm{d}x, \mathrm{d}y) \colon G \in \mathcal{C}(\nu_1, \nu_2) \right\} \right)^{1/p}.$$

For an introduction to Wasserstein distances, we refer to Villani [66, Sect. 6].

Now let (b, B, m, μ) be an admissible parameter set and denote the spectrum of $\hat{B}$ defined in (2.9) by $\sigma(\hat{B})$. We introduce the following central assumption:

Assumption 3.1 The spectral bound $s(\hat{B}) := \sup\{\Re(\lambda)\colon \lambda \in \sigma(\hat{B})\}$ of $\hat{B}$ is strictly negative, i.e., $s(\hat{B}) < 0$.

Definition 3.2 We call an affine process $(X_t)_{t\geq 0}$ on $\mathcal{H}_+$ associated with an admissible parameter set (b, B, m, μ) satisfying Assumption 3.1 a *subcritical* affine process on $\mathcal{H}_+$.

Recall that $\hat{B}$ is bounded and generates the operator semigroup $(\mathrm{e}^{t\hat{B}})_{t\geq 0}$ given by $\mathrm{e}^{t\hat{B}} := \sum_{n=0}^{\infty} \frac{(t\hat{B})^n}{n!}$, where the convergence of the series is understood in the $\mathcal{L}(\mathcal{H})$-norm. It is well known that $(\mathrm{e}^{t\hat{B}})_{t\geq 0}$ is a uniformly continuous semigroup, see Engel and Nagel [26, Sect. I.3], which implies that the spectral bound $s(\hat{B})$ coincides with the *growth bound* of $(\mathrm{e}^{t\hat{B}})_{t\geq 0}$, see [26, Corollary 4.2.4], i.e.,

$$s(\hat{B}) = \inf\{w \in \mathbb{R}\colon \exists M_w \geq 1 \text{ such that } \|\mathrm{e}^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})} \leq M_w\mathrm{e}^{wt} \text{ for all } t \geq 0\}.$$

Thus whenever Assumption 3.1 is satisfied, there exist $M \geq 1$ and $\delta > 0$ such that

$$\|\mathrm{e}^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})} \leq M\mathrm{e}^{-\delta t}, \tag{3.1}$$

The following theorem is a detailed version of our main result concerning the long-time behaviour of affine processes on the state space $\mathcal{H}_+$.

Theorem 3.3 *Let* (b, B, m, μ) *be an admissible parameter set satisfying Assumption* 3.1. *Denote the associated subcritical affine process on* $\mathcal{H}_+$ *by* $(X_t)_{t\geq 0}$ *and its transition kernels by* $(p_t(x, \cdot))_{t\geq 0}$. *Then the following hold true*:

i) *There exists a unique invariant measure* π *for* $(p_t(x, \cdot))_{t\geq 0}$, *and the Laplace transform of* π *is given by*

$$\int_{\mathcal{H}_+} e^{-\langle u, x \rangle} \pi(\mathrm{d}x) = \exp\left(-\int_0^\infty F\big(\psi(s, u)\big)\, \mathrm{d}s \right), \qquad u \in \mathcal{H}_+, \tag{3.2}$$

where F *and* $\psi(s, u)$ *are as in* (2.6) *and* (2.5). *Moreover,* π *is an inner regular measure on* $\mathcal{B}(\mathcal{H}_{+,\mathrm{w}})$.

ii) *For* $p \in [1, 2], t \geq 0$ *and* $x \in \mathcal{H}_+$, *we have*

$$W_p\big(p_t(x, \cdot), \pi\big) \leq C_1 e^{-\delta t}\left(\|x\| + \left(\int_{\mathcal{H}_+} \|y\|^p \pi(\mathrm{d}y) \right)^{1/p} \right) \tag{3.3}$$

$$+ C_2 e^{-\delta/2t}\left(\|x\|^{1/2} + \left(\int_{\mathcal{H}_+} \|y\|^{p/2} \pi(\mathrm{d}y) \right)^{1/p} \right), \tag{3.4}$$

where $C_1 = 2M$ *and* $C_2 = 2^{1/2} M^{3/2} \delta^{-1/2} \|\mu(\mathcal{H}_+ \setminus \{0\})\|^{1/2}$ *for* $M \geq 1$ *and* $\delta > 0$ *as in* (3.1). *In particular, we have* $p_t(x, \cdot) \Rightarrow \pi$ *as* $t \to \infty$.

Remark 3.4 For locally compact and second countable Hausdorff spaces, in particular for finite-dimensional normed spaces, every probability measure defined on the Borel-σ-algebra is regular. The last assertion of part i) in Theorem 3.3 states that the invariant measure π is an inner regular measure on $\mathcal{B}(\mathcal{H}_{+,\mathrm{w}})$, i.e., inner regular in the weak topology, although $\mathcal{H}_{+,\mathrm{w}}$ in the infinite-dimensional case is not locally compact. We see that the inner regularity is a nontrivial property of the invariant measure that is used in the proof of Proposition 3.5 below.

Setting $p = 1$, it is worthwhile to compare our result with the estimates given in Friesen et al. [32, Theorem 2.9, Equation (2.12)] for the state space $\mathbb{S}_+^d$ ($d \in \mathbb{N}$), i.e., $H = \mathbb{R}^d$ in our setting. The convergence rate there is given by

$$\sqrt{d} M e^{-\delta t}\left(\|x\| + \int_{\mathbb{S}_+^d} \|y\| \pi(\mathrm{d}y) \right), \qquad t \geq 0, x \in \mathbb{S}_+^d,$$

which is dimension-dependent. In contrast, the estimate obtained here in (3.3), (3.4) is independent of the dimension and hence holds for arbitrary $\mathcal{H}_+$ at the expense that an additional term (3.4) appears.

As a consequence of Theorem 3.3 which ensures the existence of an invariant inner regular measure π on $\mathcal{B}(\mathcal{H}_{+,\mathrm{w}})$, we obtain the existence of a stationary process with invariant measure π as formulated in the next result.

Proposition 3.5 *There exists a process* $(X_t^\pi)_{t\geq 0}$ *on* $\mathcal{H}_+$ *with the transition kernels* $(p_t(x, \cdot))_{t\geq 0}$ *such that the distribution of* X_t^π *equals* π *for all* $t \geq 0$.

Coming back to Theorem 3.3, we note that the pth moment of π shows up in the convergence rate in (3.3), where we implicitly assumed that these terms are finite. That this is indeed the case is part of the next result, where we also obtain explicit formulas for the first two moments of the invariant measure π.

Proposition 3.6 *Under the same conditions as in Theorem* 3.3 *and denoting the unique invariant measure of* $(p_t(x, \cdot))_{t \geq 0}$ *by* π, *we have* $\int_{\mathcal{H}_+} \|y\|^2 \pi(\mathrm{d}y) < \infty$,

$$\lim_{t \to \infty} \mathbb{E}_x[X_t] = \int_{\mathcal{H}_+} y \pi(\mathrm{d}y) = \int_0^\infty \mathrm{e}^{s\hat{B}} \left(b + \int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \xi m(\mathrm{d}\xi) \right) \mathrm{d}s \quad (3.5)$$

and

$$\lim_{t \to \infty} \mathbb{E}_x[X_t \otimes X_t] = \int_{\mathcal{H}_+} (y \otimes y) \pi(\mathrm{d}y)$$

$$= \int_0^\infty (\mathrm{e}^{s\hat{B}^*}\hat{b})^{\otimes 2} \, \mathrm{d}s + \int_0^\infty \int_{\mathcal{H}_+ \setminus \{0\}} (\mathrm{e}^{s\hat{B}^*}\xi)^{\otimes 2} m(\mathrm{d}\xi) \, \mathrm{d}s$$

$$+ \int_0^\infty \int_0^s \int_{\mathcal{H}_+ \setminus \{0\}} (\mathrm{e}^{u\hat{B}^*}\xi)^{\otimes 2} \left\langle \hat{b}, \mathrm{e}^{(s-u)\hat{B}} \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \right\rangle \mathrm{d}u \, \mathrm{d}s. \quad (3.6)$$

Example 3.7 As an example, we consider Lévy-driven Ornstein–Uhlenbeck processes on $\mathcal{H}_+$. For this, let m be a Lévy measure on $\mathcal{B}(\mathcal{H}_+ \setminus \{0\})$ with finite second moment and $b \in \mathcal{H}_+$ be such that part ii) of Definition 2.1 is satisfied. Let $\mu = 0$ and $B \in \mathcal{L}(\mathcal{H})$ be of the form $B(u) = Gu + uG^*$ for some $G \in \mathcal{L}(H)$; then part iv) of Definition 2.1 is satisfied, which can be seen from the fact that for every $u \in \mathcal{H}_+$, we have $\mathrm{e}^{tB}u = \mathrm{e}^{tG}u\mathrm{e}^{tG^*} \geq_{\mathcal{H}_+} 0$ for all $t \geq 0$ and hence Lemmert and Volkmann [55, Theorem 1] implies that B satisfies (2.2). Thus the tuple $(b, m, B, 0)$ is an admissible parameter set according to Definition 2.1, and the associated affine process $(X_t)_{t \geq 0}$ becomes an Ornstein–Uhlenbeck process driven by an $\mathcal{H}_+$-valued Lévy process $(L_t)_{t \geq 0}$ with characteristics $(b, 0, m)$; see Cox et al. [18, Lemma 2.3], i.e.,

$$X_t = \mathrm{e}^{tG} x \mathrm{e}^{tG^*} + \int_0^t \mathrm{e}^{(t-s)G} \mathrm{d}L_s \mathrm{e}^{(t-s)G^*}, \qquad t \geq 0.$$

Since $\sigma(B) = \sigma(G) + \sigma(G)$, see Rosenblum [63], and hence $s(B) \leq s(G)$, we see that whenever the spectral bound $s(G)$ of the operator G is negative, the same holds for $s(B)$ and hence Assumption 3.1 is satisfied. This provides an explicit and simple sufficient criterion for the Ornstein–Uhlenbeck process $(X_t)_{t \geq 0}$ to be subcritical. By Theorem 3.3, there exists a unique invariant measure π with Laplace transform

$$\int_{\mathcal{H}_+} \mathrm{e}^{-\langle u, x \rangle} \pi(\mathrm{d}x) = \exp\left(-\int_0^\infty \varphi_L(\mathrm{e}^{sG} u \mathrm{e}^{sG^*}) \, \mathrm{d}s \right),$$

where $\varphi_L \colon \mathcal{H} \to \mathbb{C}$ denotes the Laplace exponent of the Lévy process L, given by

$$\varphi_L(u) = \langle b, u \rangle - \int_{\mathcal{H}_+ \setminus \{0\}} \left(\mathrm{e}^{-\langle \xi, u \rangle} - 1 + \langle \chi(\xi), u \rangle \right) m(\mathrm{d}\xi), \qquad u \in \mathcal{H}_+. \quad (3.7)$$

Existence and uniqueness of invariant measures for Ornstein–Uhlenbeck processes were studied in Chojnowska-Michalik [15], where a similar result follows under the assumption of a finite log-moment of the Lévy measure m. In Proposition 3.6, the process $(X_t)_{t \geq 0}$ is assumed to have a finite second moment, which is a stronger assumption than that of finite log-moments. However, this stronger assumption also allows the derivation of explicit formulas for the first and second moments of π. Indeed, setting $\mu = 0$ and $B(u) = Gu + uG^*$ in (3.5) and (3.6) gives

$$\lim_{t \to \infty} \mathbb{E}_x [X_t] = \int_0^\infty e^{sG} \left(b + \int_{\mathcal{H}_+ \setminus \{0\} \cap \{\|\xi\| > 1\}} \xi m(\mathrm{d}\xi) \right) e^{sG^*} \, \mathrm{d}s$$

and

$$\lim_{t \to \infty} \mathbb{E}_x [X_t \otimes X_t] = \int_0^\infty (e^{sG} \hat{b} e^{sG^*})^{\otimes 2} \mathrm{d}s + \int_0^\infty \int_{\mathcal{H}_+ \setminus \{0\}} (e^{sG} \xi e^{sG^*})^{\otimes 2} m(\mathrm{d}\xi) \, \mathrm{d}s.$$

4 The stationary covariance regime for forward curve dynamics

In this section, we investigate applications of our findings regarding the long-time behaviour of affine processes on positive Hilbert–Schmidt operators within the context of an affine stochastic covariance model for commodity forward curves. This model characterises the evolution of the logarithmic forward curve over time through a stochastic partial differential equation, with the instantaneous volatility determined by a stationary affine process on positive Hilbert–Schmidt operators. Subsequently, we introduce the model, recall some auxiliary results and then study the implied volatility for forward-start options written on forward contracts.

4.1 A geometric affine forward curve model

Let $0 \leq T \leq \hat{T}$ and denote by $F(T, \hat{T})$ the forward price at time T with the maturity date $\hat{T}$, i.e., $F(T, \hat{T})$ represents the price at time T for one unit of an underlying spot commodity with delivery date $\hat{T}$. We follow the HJMM approach and model the dynamics of $(F(T, \hat{T}))_{T \leq \hat{T}}$ directly using a stochastic partial differential equation. To begin, we employ the Musiela parametrisation $x = T - t$ and denote by f_t the forward curve at time $t \geq 0$ in terms of time-to-maturity, i.e., $f_t(x) = F(t, t + x)$ for $x \geq 0$. The logarithmic forward prices $Y_t(x) = \ln f_t(x)$ are then described by

$$\mathrm{d}Y_t(x) = \left(\frac{\partial}{\partial x} Y_t(x) + G(t, x) \right) \mathrm{d}t + \sigma(t, x) \mathrm{d}W_t(x), \qquad t \geq 0, \qquad (4.1)$$

where for all $x \geq 0$, $(\sigma(t, x))_{t \geq 0}$ represents the instantaneous volatility process of the forward price dynamics, $(W(t, x))_{t \geq 0}$ is a collection of one-dimensional Brownian motions parametrised by x and $(G(t, x))_{t \geq 0}$ represents the drift component. To guarantee arbitrage-free pricing, this drift needs to have a particular functional form based on σ (the so-called HJM condition); see Benth and Krühner [7].

Mathematically, (4.1) is a stochastic partial differential equation, and hence we regard $(Y_t)_{t \geq 0}$ as a process taking values in a Hilbert space of functions containing

viable forward curves. Below, we introduce a weighted Sobolev space that serves as a model for such forward curves. An economic reasoning for this particular choice was given in Filipović [27, Chap. 5] for forward rate curves in fixed-income markets (thus it is henceforth called Filipović space), and in Clewlow and Strickland [16, Chap. 4] for commodity forward spaces. Namely, for $\beta > 0$, we let H_β be the space of all absolutely continuous functions $f : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\|f\|_\beta := \left(|f(0)|^2 + \int_{\mathbb{R}_+} e^{\beta x} |f'(x)|^2 dx \right)^{1/2} < \infty,$$

where f' denotes the weak derivative of f. We equip the space H_β with the inner product

$$\langle f, g \rangle_\beta := f(0)g(0) + \int_{\mathbb{R}_+} e^{\beta |x|} f'(x) g'(x) dx,$$

which makes $(H_\beta, \langle \cdot, \cdot \rangle_\beta)$ a separable Hilbert space.

It is well known that the left-shift semigroup, denoted by $(S(t))_{t \geq 0}$, is strongly continuous on H_β with the infinitesimal generator being the operator of differentiation in time-to-maturity. We denote it by $(\mathcal{A}, \text{dom}(\mathcal{A}))$, i.e., $\mathcal{A}f = f'$, where f belongs to its maximal domain $\text{dom}(\mathcal{A}) = \{f \in H_\beta : f' \in H_\beta\}$ (compare with Benth et al. [6, Sect. 4]). We further note that for every $t \geq 0$, the point evaluation maps $\delta_t : H_\beta \to \mathbb{R}$ are continuous linear functionals; see Filipović et al. [29, Theorem 2.1]. Throughout this section, we identify the map δ_t for any $t \geq 0$ with an element $u_t \in H_\beta$ such that $\langle f, u_t \rangle_\beta = \delta_t(f)$ holds. In this way, we can interpret the process $(Y_t)_{t \geq 0}$ as the (unique strong) solution to the stochastic differential equation on H_β given by

$$dY_t = \big(\mathcal{A}Y_t + G(\sigma_t)\big)dt + \sigma_t dW_t, \qquad Y_0 = y \in H_\beta, \tag{4.2}$$

where W is a cylindrical Brownian motion on H_β, $G : \mathcal{H} \to H_\beta$ a continuous affine function whose dependence on σ stems from the HJM condition, and $(\sigma_t)_{t \geq 0}$ is a positive trace-class-operator-valued process that determines the stochastic covariance structure of forwards with different times to maturity.

An affine stochastic covariance model, as introduced by Cox et al. [18], is characterised by the instantaneous volatility $\sigma_t = D^{1/2} X_t^{1/2}$, where D is a trace-class operator on H_β and $(X_t)_{t \geq 0}$ is an affine process on the cone of positive Hilbert–Schmidt operators. Here D corresponds to a constant covariance term used to ensure that the covariance process σ is integrable with respect to W; see Cox et al. [18] for a discussion. Consequently, for such a choice of σ_t, G takes the form

$$G(X_t) = -\frac{1}{2} D^{1/2} X_t D^{1/2} S^*(\cdot) u_0 \tag{4.3}$$

with $u_0 \in H_\beta$ such that $\langle f, u_0 \rangle_\beta = \delta_0(f)$, in order to satisfy the HJM condition; see Karbach [46, Lemma 4.6]. The joint process $(Y_t^y, X_t)_{t \geq 0}$ is then called an *affine stochastic covariance model on H_β*. It determines the forward curve dynamics via

$$F(T, \hat{T}) := \delta_{\hat{T}-T}\big(\exp(Y_T^y)\big) = \exp(\langle Y_T^y, u_{\hat{T}-T} \rangle_\beta). \tag{4.4}$$

This justifies the name *geometric affine forward curve model* for $(F(T, \hat{T}))_{T \leq \hat{T}}$ on H_β. Note that geometric models of this form ensure the positivity of forward prices which is crucial for many forward markets. A similar geometric model was proposed in Benth and Sgarra [10]. In their case, the underlying stochastic covariance model $(Y_t^y, X_t)_{t \geq 0}$ is the Hilbert-valued Barndorff-Nielsen–Shephard (BNS) model (see (4.6) below) with additional leverage term in the dynamics (4.2). However, in our model (4.2), we omit the leverage term (compared to [10]) for brevity, although leverage is a stylised fact in many commodity forward markets. The advantage of the more general affine geometric model class $(Y_t^y, X_t)_{t \geq 0}$ is its flexible jump structure; in particular, it allows self-exciting jumps, which has the potential to model *volatility clustering*, a key feature in many commodity markets.

Example 4.1 We give an example of an admissible parameter set (b, B, m, μ). For brevity, let the constant parameters b and m be zero. Let $g, z \in \mathcal{H}_+$ be fixed and let $\eta \colon \mathcal{B}((0, \infty)) \to [0, \infty)$ be such that $\int_0^\infty \lambda^{-2}\eta(d\lambda) < \infty$. Then define the vector-valued measure $\mu \colon \mathcal{B}(\mathcal{H}_+ \setminus \{0\}) \to \mathcal{H}_+$ by

$$\mu(A) := g\eta(\{\lambda \in \mathbb{R}_+ \colon \lambda z \in A\}),$$

which implies that jumps only occur in the direction of $z \in \mathcal{H}_+$, with the intensity determined by the measure η on $\mathbb{R}_+$, and the structure of the spillover effects between the volatility of forwards with different times to maturity determined by the operator g. If we define for $\gamma \in \mathbb{R}$,

$$B(u) := \gamma u + \int_{\mathcal{H}_+ \setminus \{0\}} \chi(\xi) \frac{\langle u, \mu(d\xi) \rangle}{\|\xi\|^2}$$

$$= \gamma u + \frac{z}{\|z\|^2} \langle g, u \rangle \int_0^{\frac{1}{\|z\|}} \lambda^{-1}\eta(d\lambda), \qquad u \in \mathcal{H}_\beta,$$

then B and μ satisfy Condition (2.2), i.e., $(0, B, 0, \mu)$ is an admissible parameter set in the sense of Definition 2.1. The associated affine process $(X_t)_{t \geq 0}$ is concentrated on the set $\{\lambda z \colon \lambda \geq 0\}$. In applications, we could choose z to be a Hilbert–Schmidt integral operator on H_β, i.e., given by some kernel function $k \colon \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$.

Let us verify that the corresponding affine process is subcritical for appropriate γ, so that our results can be applied. Indeed, it is easy to see that

$$B^*(u) = \gamma u + g \left\langle \frac{z}{\|z\|^2}, u \right\rangle \int_0^{\frac{1}{\|z\|}} \lambda^{-1}\eta(d\lambda).$$

Hence we find that the effective drift is given by

$$\widehat{B}(u) = \gamma u + g \left\langle \frac{z}{\|z\|^2}, u \right\rangle \int_0^\infty \lambda^{-1}\eta(d\lambda),$$

which is well defined provided that $\int_0^\infty \lambda^{-1}\eta(d\lambda) < \infty$. To verify that the affine process is subcritical, let us first note that the semigroup $(e^{t\widehat{B}})_{t \geq 0}$ on $\mathcal{H}_\beta$ is given by

$$e^{t\widehat{B}}u = \frac{a(u)}{a(g)}e^{t(a(g)+\gamma)}g, \qquad t \geq 0,$$

where $a(u) = \langle \frac{z}{\|z\|^2}, u \rangle \int_0^\infty \lambda^{-1} \eta(d\lambda) \geq 0$. Now suppose that

$$\langle z, g \rangle > 0 \qquad \text{and} \qquad \gamma < -\left\langle \frac{z}{\|z\|^2}, g \right\rangle \int_0^\infty \lambda^{-1} \eta(d\lambda).$$

Then $a(g) > 0$ and hence $(e^{t\widehat{B}})_{t\geq 0}$ is uniformly exponentially stable. This shows that the process is subcritical.

4.2 The stationary covariance regime for affine models

Next we proceed to construct the model introduced in (4.2) and (4.3) in the stationary covariance regime. For this purpose, let $(X_t)_{t\geq 0}$ be an affine process on $\mathcal{H}_+$ with initial value $X_0 = x$, associated with an admissible parameter set (b, B, m, μ). Let $(Y_t^y)_{t\geq 0}$ be the unique (mild) solution to (4.2) given by the variation-of-constants formula

$$Y_t^y = S(t)y + \int_0^t S(t-s)G(X_s)ds + \int_0^t S(t-s)D^{1/2}X_s^{1/2}dW_s, \qquad t \geq 0. \quad (4.5)$$

In other words, let $(Y_t^y, X_t)_{t\geq 0}$ be an affine stochastic covariance model on H_β. Note that $(Y_t^y, X_t)_{t\geq 0}$ can be considered as a stochastic process on the probability basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}_x) := (\Omega^1 \times \Omega^2, \mathcal{F}^1 \otimes \mathcal{F}^2, (\mathcal{F}_t^1 \otimes \mathcal{F}_t^2)_{t\geq 0}, \mathbb{Q} \otimes \mathbb{P}_x)$, where we denote by $(\Omega^2, \mathcal{F}^2, (\mathcal{F}_t^2)_{t\geq 0}, \mathbb{P}_x)$ the filtered probability space accommodating the affine process $(X_t)_{t\geq 0}$ (see also Remark 2.3) and $(\Omega^1, \mathcal{F}^1, (\mathcal{F}_t^1)_{t\geq 0}, \mathbb{Q})$ is another filtered probability space carrying the cylindrical Wiener process $W : [0, \infty) \times \Omega \to H_\beta$ and the solution process $(Y_t)_{t\geq 0}$ with $\mathbb{Q}[Y_0 = y] = 1$. From now on, we denote the expectation with respect to the product measure $\mathbb{Q}_x$ by $\mathbb{E}_x[\cdot]$. It was shown in Karbach [47] that affine processes $(X_t)_{t\geq 0}$ on $\mathcal{H}_+$ have versions with càdlàg paths, and in the following, we always consider a càdlàg version of $(X_t)_{t\geq 0}$. It follows from Cox et al. [18] that under these conditions, the stochastic covariance model $(Y_t^y, X_t)_{t\geq 0}$ is well defined for every initial value $(y, x) \in H_\beta \times \mathcal{H}_+$. Heuristically, the process $(Y_t^y, X_t)_{t\geq 0}$ satisfies a similar transform formula for its mixed Fourier–Laplace transform as the process $(X_t)_{t\geq 0}$ does for the Laplace transform in (1.1); see Benth et al. [9], Cox et al. [18].

If moreover Assumption 3.1 is satisfied, then by Theorem 3.3, there exists a unique invariant measure π for $(p_t(x, \cdot))_{t\geq 0}$, and Proposition 3.5 ensures the existence of the stationary process $(X_t^\pi)_{t\geq 0}$. Now if there exists a mild solution $(\tilde{Y}_t)_{t\geq 0}$ of (4.5) for $y = 0$ and where the process $(X_t)_{t\geq 0}$ is replaced by $(X_t^\pi)_{t\geq 0}$, then we call the joint process $(\tilde{Y}_t, X_t^\pi)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}_\pi)$ (with $\mathbb{Q}_\pi = \mathbb{Q} \otimes \mathbb{P}_\pi$ and where the expectation with respect to $\mathbb{Q}_\pi$ is denoted by $\mathbb{E}_\pi[\cdot]$) an affine stochastic covariance model on H in the *stationary covariance regime*. This terminology is inspired by the univariate setting in Keller-Ressel [49, Sect. 3]. In the following result, we give an affine transform formula for the affine stochastic covariance model $(Y_t^y, X_t)_{t\geq 0}$ in the stationary covariance regime, where for brevity we let $G = 0$.

Proposition 4.2 *Assume that $(Y_t^y, X_t)_{t\geq 0}$ is an affine stochastic covariance model, where the admissible parameters of the process $(X_t)_{t\geq 0}$ satisfy Assumption 3.1, and*

let $(\tilde{Y}_t, X_t^\pi)_{t \geq 0}$ be the model in the stationary covariance regime. Then for every $T \geq 0$ and $u = (u_1, u_2) \in \mathrm{i}H_\beta \times \mathcal{H}_+$, we have

$$\mathbb{E}_\pi[\mathrm{e}^{\langle \tilde{Y}_t, u_1 \rangle_\beta - \langle X_t^\pi, u_2 \rangle_\beta}] = \mathrm{e}^{-\Phi(t, u_1, u_2) - \int_0^\infty F(\psi_2(s, 0, \psi_2(t, u_1, u_2)))\mathrm{d}s}, \qquad t \in [0, T],$$

where $\Phi(\cdot, u_1, u_2)$, $\psi_1(\cdot, u_1, u_2)$ and $\psi_2(\cdot, u_1, u_2)$ are the unique solutions on $[0, T]$ of the differential equations

$$\begin{cases} \dfrac{\partial \Phi}{\partial t}(t, u) = F\big(\psi_2(t, u)\big), & \Phi(0, u) = 0, \\[2mm] \psi_1(t, u) = u_1 - \mathrm{i}\mathcal{A}^*\left(\mathrm{i} \int_0^t \psi_1(s, u)\mathrm{d}s\right), & \psi_1(0, u) = u_1, \\[2mm] \dfrac{\partial \psi_2}{\partial t}(t, u) = \mathcal{R}\big(\psi_1(t, u), \psi_2(t, u)\big), & \psi_2(0, u) = u_2, \end{cases}$$

where F is as in (2.6), $\mathcal{R}(h, u) := R(u) - \frac{1}{2}(D^{1/2}h) \otimes (D^{1/2}h)$ with R as in (2.7), and $(\mathcal{A}^*, \mathrm{dom}(\mathcal{A}^*))$ denotes the adjoint operator of the generator $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$.

The proof of Proposition 4.2 is analogous to the finite-dimensional version in Keller-Ressel [49] and is thus omitted. As a particular case, we obtain an operator-valued extension of the Barndorff-Nielsen–Shephard stochastic covariance model as introduced below.

Example 4.3 The *operator Barndorff-Nielsen–Shephard* (BNS) stochastic covariance model $(Y_t^y, X_t)_{t \geq 0}$ of Benth et al. [9] is given by the pair of stochastic differential equations

$$\begin{cases} \mathrm{d}Y_t = \mathcal{A}(Y_t)\mathrm{d}t + D^{1/2}X_t^{1/2}\mathrm{d}W_t, & Y_0 = y \in H_\beta, \\[2mm] \mathrm{d}X_t = B(X_t)\mathrm{d}t + \mathrm{d}L_t, & X_0 = x \in \mathcal{H}_+, \end{cases} \tag{4.6}$$

where B and $(L_t)_{t \geq 0}$ are as in Example 3.7. Note that $(Y_t^\pi)_{t \geq 0}$ is the same as (4.5) written in differential form with $G = 0$ and $(S(t))_{t \geq 0}$ denoting the strongly continuous semigroup generated by $(\mathcal{A}, \mathrm{dom}(\mathcal{A})$. For the $\mathcal{H}_+$-valued Ornstein–Uhlenbeck process $(X_t)_{t \geq 0}$, we already showed the existence of a unique invariant measure π of $(X_t)_{t \geq 0}$ in Example 3.7. Hence we may consider the *operator BNS model in the stationary covariance regime* and denote it by $(\tilde{Y}_t, X_t^\pi)_{t \geq 0}$. In Cox et al. [18], it was shown that the operator valued BNS model is a particular case of the SV models with affine pure-jump variance as given by (4.2) with $\sigma_t = D^{1/2}X_t^{1/2}$, where $(X_t)_{t \geq 0}$ is an affine process on $\mathcal{H}_+$. Hence we obtain from Proposition 4.2 applied to this particular case that for every $(u_1, u_2) \in \mathrm{i}H_\beta \times \mathcal{H}_+$,

$$\mathbb{E}_\pi[\mathrm{e}^{\langle \tilde{Y}_t, u_1 \rangle_\beta - \langle X_t^\pi, u_2 \rangle_\beta}] = \exp\left(-\int_0^t \varphi_L\big(\psi(s, u_1, u_2)\big)\mathrm{d}s - \int_0^\infty \varphi_L(\mathrm{e}^{sB^*}u_2)\,\mathrm{d}s\right),$$

where φ_L is given by (3.7) and $\psi(t, u_1, u_2)$ is explicitly known as

$$\psi(t, u_1, u_2) = \mathrm{e}^{sB^*}u_2 + \frac{1}{2}\int_0^s \mathrm{e}^{(s-\tau)B^*}\big(D^{1/2}S^*(\tau)u_1\big)^{\otimes 2}\mathrm{d}\tau.$$

4.3 Pricing forward-start options on forwards

In this section, we study the geometric affine forward curve model in (4.4) determined by an affine stochastic covariance model $(Y_t^y, X_t)_{t \geq 0}$. In particular, we examine the long-time behaviour of the forward implied volatility in this model. Let us assume our model is already formulated under the risk-neutral measure $\tilde{\mathbb{Q}}$, i.e., we assume $(F(T, \hat{T}))_{T \leq \hat{T}}$ is a $\tilde{\mathbb{Q}}$-martingale for all $\hat{T} \geq 0$. A *forward-start option* with forward-start date $\tau \geq 0$, forward maturity T and strike e^K written on a forward with maturity date $\hat{T}$ is a European option with payoff at time $\tau + T$ given by

$$\left(\frac{F(\tau + T, \tau + \hat{T})}{F(\tau, \tau + \hat{T})} - e^K \right)^+. \tag{4.7}$$

A forward-start option is a contract on the relative price difference of a forward contract at two times τ and $\tau + T$ in the future. In practice, it is used to price the future volatility of the underlying asset. Forward-start options are very common in commodity forward markets, and more complex derivatives such as *cliquet options* are building up on these; see e.g. Crosby [19]. Forward-start options on stocks are discussed in e.g. Jacquier and Roome [39, 40], Kruse and Nögel [53], Keller-Ressel [49]. Here we restrict ourselves to the ratio-type payoff functions in (4.7), but similar results can be obtained for difference-type payoffs, i.e., $(F(\tau + T, \tau + \hat{T}) - K F(\tau, \tau + \hat{T}))^+$; see also Kruse and Nögel [53].

We now define the *implied forward volatility* of the model (4.4). First, let us denote the price of a forward-start option with payoff (4.7) by $C_{\text{fwd}}(\tau, T, \hat{T}, K)$. Then, as a reference model for the forward prices $F(T, \hat{T})$, we take *Black's model*, see Black [12], and denote the forward prices within this model by $F^B(T, \hat{T})$. We assume that we have the spot–forward relation

$$F^B(T, \hat{T}) = s_T e^{r(\hat{T} - T)}, \qquad 0 \leq T \leq \hat{T}, \tag{4.8}$$

where $r \geq 0$ denotes the risk-free interest rate and $(s_t)_{t \geq 0}$ the spot price process of the underlying commodity, which is given by a geometric Brownian motion with volatility parameter σ. We denote by $C_{\text{fws}}^B(\tau, T, \hat{T}, K, \sigma)$ the price in Black's model of a forward-start option with an identical payoff function as in (4.7) and define the *implied forward volatility* $\sigma(\tau, T, \hat{T}, K)$ as the unique solution to

$$C_{\text{fws}}^B \big(\tau, T, \hat{T}, K, \sigma(\tau, T, \hat{T}, K) \big) = C_{\text{fwd}}(\tau, T, \hat{T}, K).$$

In the following result we show that $\sigma(\tau, T, \hat{T}, K)$ exists for all $\tau, K \geq 0$ and study its long-time behaviour as $\tau \to \infty$.

Proposition 4.4 *Let* $0 \leq T \leq \hat{T}$ *and denote by* $F(T, \hat{T})$ *the forward price at time* T *with maturity date* $\hat{T}$ *given by* (4.4), *where* $(Y_t^y, X_t)_{t \geq 0}$ *is an affine stochastic covariance model on* H_β *as defined in Sect.* 4.2 *with* $(S(t))_{t \geq 0}$ *the left-shift semigroup on* H_β. *Moreover, let* $(\tilde{Y}_t, X_t^\pi)_{t \geq 0}$ *be the model in the stationary covariance regime and*

define $\tilde{F}(T, \hat{T}) := \exp(\langle \tilde{Y}_T, u_{\hat{T}-T} \rangle_\beta)$. Suppose we model directly under the pricing measure $\tilde{\mathbb{Q}}$ so that $(F(T, \hat{T}))_{T \le \hat{T}}$ is a $\tilde{\mathbb{Q}}$-martingale for all $\hat{T} \ge 0$. Then for all $\tau, K \ge 0$, the implied forward volatility $\sigma(\tau, T, \hat{T}, K)$ exists and we have

$$\lim_{\tau \to \infty} \sigma(\tau, T, \hat{T}, K) = \tilde{\sigma}(T, \hat{T}, K), \tag{4.9}$$

where $\tilde{\sigma}(T, \hat{T}, K)$ denotes the implied volatility of a European call option with payoff function $(\tilde{F}(T, \hat{T}) - K)^+$.

Proof First, we show for $0 \le T \le \hat{T}$ and $K \ge 0$ a relation between the prices $C^{\mathrm{B}}(T, \hat{T}, K, \sigma)$ of a European call option and $C^{\mathrm{B}}_{\mathrm{fws}}(\tau, T, \hat{T}, K, \sigma)$ of the forward-start call option with forward-start date $\tau \ge 0$, both in Black's model. Namely, let $\mathbb{Q}$ denote the unique risk-neutral measure in Black's model and recall the price $C^{\mathrm{B}}_{\mathrm{fws}}(\tau, T, \hat{T}, K, \sigma)$ of a forward-start option at time zero with forward-start date τ written on the forward with maturity $\hat{T}$. Inserting (4.8) into the payoff function and by risk-neutral pricing, we have

$$C^{\mathrm{B}}_{\mathrm{fws}}(\tau, T, \hat{T}, K, \sigma) = \mathrm{e}^{-r(\tau+T)} \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{s_{\tau+T} \mathrm{e}^S r(\hat{T} - T)}{s_\tau \mathrm{e}^{r\hat{T}}} - \mathrm{e}^K \right)^+ \right].$$

It is known that in the Black–Scholes model, the forward-start call option and the European call option satisfy

$$\mathrm{e}^{-r(\tau+T)} \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{s_{\tau+T}}{s_\tau} - K \right)^+ \right] = \mathrm{e}^{-r(\tau+T)} \mathbb{E}_{\mathbb{Q}}[(s_T - K)^+],$$

see also Keller-Ressel [49], and hence

$$
\begin{aligned}
C^{\mathrm{B}}_{\mathrm{fws}}(\tau, T, \hat{T}, K, \sigma) &= \mathrm{e}^{-r(\tau+T)} \mathrm{e}^{-rT} \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{s_{\tau+T}}{s_\tau} - \mathrm{e}^K \right)^+ \right] \\
&= \mathrm{e}^{-r(\tau+T)} \mathrm{e}^{-rT} \mathbb{E}_{\mathbb{Q}}[(s_T - \mathrm{e}^{K'})^+] \\
&= \mathrm{e}^{-r(\tau+T)} C^{\mathrm{BS}}(T, K', \sigma),
\end{aligned}
$$

where $K' = K + rT$ and the superscript BS indicates that the underlying model for $(s_t)_{t \ge 0}$ is the Black–Scholes model. From this and the definition of the implied forward volatility $\sigma(\tau, T, \hat{T}, K)$, we have

$$
\begin{aligned}
C^{\mathrm{BS}}(T, K', \sigma(\tau, T, \hat{T}, K)) &= \mathrm{e}^{r(\tau+T)} C^{\mathrm{B}}_{\mathrm{fws}}(\tau, T, \hat{T}, K, \sigma(\tau, T, \hat{T}, K)) \\
&= \mathrm{e}^{r(\tau+T)} C_{\mathrm{fwd}}(\tau, T, \hat{T}, K). \tag{4.10}
\end{aligned}
$$

Next, we compute the right-hand side of (4.10). Recall that for every $t \in \mathbb{R}$ and $f \in H_\beta$, we use the identification $\delta_t(f) = \langle f, u_t \rangle$ for the evaluation functional with $u_t \in H_\beta$. Moreover, we denote the expectation with respect to the pricing measure $\tilde{\mathbb{Q}}$ by $\mathbb{E}_{\tilde{\mathbb{Q}}}[\cdot]$ (here we suppress the initial value x compared to $\mathbb{Q}_x$ above). The payoff

function of the forward-start option is given by (4.7); hence by risk-neutral pricing and inserting our model (4.4), we have

$$C_{\text{fwd}}(\tau, T, \hat{T}, K) = e^{-r(\tau+T)} \mathbb{E}_{\tilde{\mathbb{Q}}}\left[\left(\frac{F(\tau+T, \tau+\hat{T})}{F(\tau, \tau+\hat{T})} - e^K\right)^+\right]$$

$$= e^{-r(\tau+T)} \mathbb{E}_{\tilde{\mathbb{Q}}}[(e^{\langle Y^y_{\tau+T}, u_{\hat{T}-T}\rangle_\beta - \langle Y^y_\tau, u_{\hat{T}}\rangle_\beta} - e^K)^+]. \tag{4.11}$$

Note that by the definition of Y_τ, we have

$$\langle S(T)Y_\tau, u_{\hat{T}-T}\rangle_\beta = \langle S(T+\tau)y, u_{\hat{T}-T}\rangle_\beta + \left\langle \int_0^\tau S(T+\tau-s)G(X_s)ds, u_{\hat{T}-T}\right\rangle_\beta$$

$$+ \left\langle \int_0^\tau S(T+\tau-s)X_s^{1/2}dW_s, u_{\hat{T}-T}\right\rangle_\beta,$$

and the left-shift $S(T)$ satisfies $\langle S(T)Y^y_\tau, u_{\hat{T}-T}\rangle_\beta = \langle Y^y_\tau, u_{\hat{T}}\rangle_\beta$. Thus we obtain

$$\langle Y^y_{\tau+T}, u_{\hat{T}-T}\rangle_\beta - \langle Y^y_\tau, u_{\hat{T}}\rangle_\beta = \left\langle \int_\tau^{T+\tau} S(T+\tau-s)G(X_s)\,ds, u_{\hat{T}-T}\right\rangle_\beta$$

$$+ \left\langle \int_\tau^{T+\tau} S(T+\tau-s)X_s^{1/2}dW_s, u_{\hat{T}-T}\right\rangle_\beta. \tag{4.12}$$

By the independent increments property and the Markov property of $(X_t)_{t\geq0}$, the sum of the integrals inside the inner products on the right-hand side of (4.12) has the same distribution as $Y^0_T = \int_0^T S(T-s)G(X_{\tau+s})ds + \int_0^T S(T-s)X_{\tau+s}^{1/2}dW_s$. Hence for the expectation on the right-hand side in (4.11), we obtain

$$\mathbb{E}_{\tilde{\mathbb{Q}}}[(e^{\langle Y^y_{\tau+T}, u_{\hat{T}-T}\rangle_\beta - \langle Y^y_\tau, u_{\hat{T}}\rangle_\beta} - e^K)^+] = \mathbb{E}_{\tilde{\mathbb{Q}}}[\mathbb{E}[(e^{\langle Y^0_T, u_{\hat{T}-T}\rangle_\beta} - e^K)^+ \mid X_\tau]],$$

and thus we conclude that the left-hand side of (4.10) is given by

$$C^{\text{BS}}(T, K', \sigma(\tau, T, \hat{T}, K)) = \mathbb{E}_{\tilde{\mathbb{Q}}}[\mathbb{E}[(e^{\langle Y^0_T, u_{\hat{T}-T}\rangle_\beta} - e^K)^+ \mid X_\tau]].$$

Now taking the limit $\tau \to \infty$ and since $\tilde{F}(T, \hat{T}) = e^{\langle \tilde{Y}_T, u_{\hat{T}-T}\rangle_\beta}$, we obtain

$$\lim_{\tau\to\infty} C^{\text{BS}}(T, K', \sigma(\tau, T, \hat{T}, K)) = \mathbb{E}_{\tilde{\mathbb{Q}}}[(e^{\langle \tilde{Y}_T, u_{\hat{T}-T}\rangle_\beta} - e^K)^+]$$

$$= \mathbb{E}_{\tilde{\mathbb{Q}}}[(\tilde{F}(T, \hat{T}) - e^K)^+]. \tag{4.13}$$

The term $\mathbb{E}_{\tilde{\mathbb{Q}}}[(\tilde{F}(T, \hat{T}) - e^K)^+]$ on the right-hand side of (4.13) is simply e^{rT} times the price of a European call option, and the continuity of $\sigma \mapsto C^{\text{BS}}(T, K', \sigma)$ gives

$$\lim_{\tau\to\infty} C^{\text{BS}}(T, K', \sigma(\tau, T, \hat{T}, K)) = C^{\text{BS}}\left(T, K', \lim_{\tau\to\infty} \sigma(\tau, T, \hat{T}, K)\right),$$

from which we conclude (4.9), since (4.13) has a unique solution in terms of the Black–Scholes implied volatility.   □

According to Proposition 4.4, the prices of forward-start options for large forward dates τ can be accurately approximated by those of European plain vanilla options modelled in the stationary covariance regime. This result is significant because pricing options in affine stochastic covariance models, particularly in the stationary covariance regime, can be done with relatively low computational effort, as demonstrated in Karbach [46, Chap. 4]. This is due to the affine transform formula provided by Proposition 4.2, which allows computing option prices using Fourier inversion, as discussed in references such as Carr and Madan [14], Kallsen et al. [45], Hubalek et al. [37]. Our ongoing work focuses on the numerical aspects of option pricing in infinite-dimensional affine stochastic covariance models.

5 Proofs of the main results

Throughout this section, we assume that (b, B, m, μ) is an admissible parameter set according to Definition 2.1. We denote the unique subcritical affine process associated with (b, B, m, μ) through Theorem 2.2 by $(X_t)_{t \geq 0}$ and its family of transition kernels by $(p_t(x, \cdot))_{t \geq 0}$. We set $P_t f := \int_{\mathcal{H}} f(\xi) p_t(\cdot, d\xi)$ for all measurable functions f such that the integral exists. Recall from Cox et al. [17] that the transition semigroup $(P_t)_{t \geq 0}$ is a generalised Feller semigroup on the space $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$; see also Appendix A.

5.1 Some properties of the generalised Riccati equations (2.4), (2.5)

In this section, we consider the long-time behaviour of the solutions $\phi(\cdot, u)$ and $\psi(\cdot, u)$ of the generalised Riccati equations in (2.4), (2.5). The arguments in this section are analogous to the finite-dimensional setting in Friesen et al. [32] and are thus left to the reader. We recall from Cox et al. [17, Sect. 3] that for every $u \in \mathcal{H}_+$, there exists a unique and global solution $\psi(\cdot, u) \in C^1(\mathbb{R}_+, \mathcal{H}_+)$ to (2.5). Given $\psi(\cdot, u)$, we solve (2.4) by mere integration and obtain $\phi(\cdot, u) \in C^1(\mathbb{R}_+, \mathbb{R}_+)$ given by $\phi(t, u) = \int_0^t F(\psi(s, u)) \, ds$. This means that we can write the affine transform formula (2.3) as

$$\int_{\mathcal{H}_+} e^{-\langle \xi, u \rangle} p_t(x, d\xi) = \exp\left(-\int_0^t F(\psi(s, u)) \, ds - \langle x, \psi(t, u) \rangle \right).$$

Moreover, we recall that the unique solution $\psi(\cdot, u)$ to (2.5) satisfies the flow equation

$$\psi(t + s, u) = \psi(t, \psi(s, u)). \tag{5.1}$$

In the next result, we show that F and R are continuous functions on $\mathcal{H}_+$ and grow at most quadratically.

Lemma 5.1 *Let (b, B, m, μ) be an admissible parameter set according to Definition 2.1 and let F and R be given by (2.6) and (2.7), respectively. Then F and R are*

continuous on $\mathcal{H}_+$ and for all $u \in \mathcal{H}_+$, we have

$$|F(u)| \leq \left(\|b\| + \int_{\mathcal{H}_+ \setminus \{0\}} \|\xi\|^2 m(\mathrm{d}\xi) \right) (\|u\| + \|u\|^2),$$

$$\|R(u)\| \leq \left(\|B\|_{\mathcal{L}(\mathcal{H})} + \|\mu(\mathcal{H}_+ \setminus \{0\})\| \right) (\|u\| + \|u\|^2).$$

Proof See Karbach [47, Lemma 5.13].   □

Assumption 3.1 implies that the semigroup $(e^{t\hat{B}})_{t \geq 0}$ satisfies (3.1), i.e., $(e^{t\hat{B}})_{t \geq 0}$ is uniformly exponentially stable. This has the following consequence on the solution $\psi(\,\cdot\,, u)$ of the generalised Riccati equation (2.5).

Lemma 5.2 *Let (b, B, m, μ) be an admissible parameter set according to Definition 2.1 and for $u \in \mathcal{H}_+$, let $\psi(\,\cdot\,, u)$ be the unique solution to (2.5). Then*

$$\|\psi(t, u)\| \leq \|e^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})} \|u\|, \qquad \forall t \geq 0.$$

If moreover Assumption 3.1 is satisfied, then $\lim_{t \to \infty} \psi(t, u) = 0$.

5.2 Invariant measures for affine processes on $\mathcal{H}_+$

For two measures $\nu_1, \nu_2 \in \mathcal{M}(\mathcal{H}_+)$, we denote the convolution of ν_1 and ν_2 by $\nu_1 * \nu_2$. In the following result, we give an important convolution property of the transition kernels $p_t(x, \,\cdot\,)$.

Lemma 5.3 *Let $(Y_t)_{t \geq 0}$ be the unique affine process associated with the admissible parameter set $(0, B, 0, \mu)$ and transition kernels $(q_t(x, \,\cdot\,))_{t \geq 0}$. Then for every $t \geq 0$ and $x \in \mathcal{H}_+$, we have*

$$p_t(x, \,\cdot\,) = p_t(0, \,\cdot\,) * q_t(x, \,\cdot\,).$$

Proof Since $b = 0$ and $m = 0$, the function F in (2.4) vanishes, see also (2.6), and thus $\phi(t, u) = 0$ for all $t \geq 0$. Hence for every $t \geq 0$, the affine transform formula (2.3) for Y_t takes the form

$$\int_{\mathcal{H}_+} e^{-\langle u, \xi \rangle} q_t(x, \mathrm{d}\xi) = \exp\left(- \langle \psi(t, u), x \rangle \right) \qquad \text{for } u \in \mathcal{H}_+.$$

Now let $(X_t)_{t \geq 0}$ denote the unique affine process associated with the admissible parameter set (b, B, m, μ) and denote its transition kernels by $p_t(x, \,\cdot\,)$. Let $t \geq 0$ be arbitrary and $u \in \mathcal{H}_+$; then

$$\int_{\mathcal{H}_+} e^{-\langle u, \xi \rangle} \left(p_t(0, \,\cdot\,) * q_t(x, \,\cdot\,) \right)(\mathrm{d}\xi) = \int_{\mathcal{H}_+} \left(\int_{\mathcal{H}_+} e^{-\langle u, \xi_1 + \xi_2 \rangle} p_t(0, \mathrm{d}\xi_1) \right) q_t(x, \mathrm{d}\xi_2)$$

$$= e^{-\phi(t, u)} \int_{\mathcal{H}_+} e^{-\langle u, \xi_2 \rangle} q_t(x, \mathrm{d}\xi_2)$$

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

which completes the proof thanks to (2.3) and the fact that the functions $x \mapsto e^{-\langle u, x \rangle}$ characterise measures; see Cox et al. [18, Lemma A.1]. $\qquad\square$

In the next result, we show that the Laplace transform of a subcritical affine process converges pointwise as the time t tends to infinity.

Proposition 5.4 *Let* $(X_t)_{t \geq 0}$ *be an affine process associated with the admissible parameter set* (b, B, m, μ) *satisfying Assumption* 3.1. *Then for all* $u, x \in \mathcal{H}_+$, *we have*

$$\lim_{t \to \infty} \mathbb{E}_x[e^{-\langle u, X_t \rangle}] = \exp\left(- \int_0^\infty F(\psi(s, u)) \, \mathrm{d}s \right) \in [0, \infty). \tag{5.2}$$

Proof Let $u, x \in \mathcal{H}_+$. Then by Lemma 5.2 and (3.1), we have

$$|\langle \psi(t, u), x \rangle| \leq \|\psi(t, u)\| \|x\| \leq \|e^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})} \|x\| \|u\| \leq M e^{-\delta t} \|x\| \|u\|.$$

Lemma 5.1 gives

$$|F(\psi(t, u))| \leq C\big(\|\psi(t, u)\| + \|\psi(t, u)\|^2\big) \leq C M^2 e^{-\delta t} (\|u\| + \|u\|^2) \tag{5.3}$$

with $C = \|b\| + \int_{\mathcal{H}_+ \setminus \{0\}} \|\xi\|^2 m(\mathrm{d}\xi)$. For every $u \in \mathcal{H}_+$, this implies

$$\int_0^\infty |F(\psi(s, u))| \, \mathrm{d}s \leq \frac{C M^2}{\delta} (\|u\| + \|u\|^2) < \infty,$$

and hence the limit $\lim_{t \to \infty} \phi(t, u) = \int_0^\infty F(\psi(s, u)) \, \mathrm{d}s$ exists for every $u \in \mathcal{H}_+$. This, continuity of the exponential function and the fact that $\langle \psi(t, u), x \rangle \to 0$ for all $x, u \in \mathcal{H}_+$ as $t \to \infty$ by Lemma 5.2 imply (5.2). $\qquad\square$

The next result asserts uniform boundedness in time of the transition semigroup $(P_t)_{t \geq 0}$ in the operator norm on $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$.

Lemma 5.5 *Let* $(X_t)_{t \geq 0}$ *be an affine process associated with the admissible parameter set* (b, B, m, μ) *satisfying Assumption* 3.1 *and denote its transition semigroup by* $(P_t)_{t \geq 0}$. *Then we have*

$$\sup_{t \geq 0} \|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}))} < \infty. \tag{5.4}$$

Proof Recall that $\rho(x) = 1 + \|x\|^2$ and note that for every $f \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$, we have $|f(y)| \leq \rho(y) \|f\|_{\mathcal{B}_\rho}$ and hence

$$\|P_t f\|_{\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})} = \sup_{x \in \mathcal{H}_+} \rho(x)^{-1} \left| \int_{\mathcal{H}_+} f(y) p_t(x, \mathrm{d}y) \right| \leq \|f\|_{\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})} \|P_t \rho\|_{\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})},$$

which yields $\sup_{t\geq 0}\|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,w}))} \leq \sup_{t\geq 0}\|P_t\rho\|_{\mathcal{B}_\rho(\mathcal{H}_{+,w})}$. Let $(e_i)_{i\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$ and recall that by Cox et al. [17, Remark 4.6], we have $P_t\rho(x) = \mathbb{E}_x[\rho(X_t)]$ for all $t \geq 0$. Hence by Parseval's identity, we conclude that

$$0 \leq P_t\rho(x) = 1 + \mathbb{E}_x[\|X_t\|^2] = 1 + \sum_{i=1}^{\infty}\mathbb{E}_x[\langle X_t, e_i\rangle^2].$$

Using (2.11) with $v = w = e_i$ for $i \in \mathbb{N}$, we find

$$\mathbb{E}_x[\langle X_t, e_i\rangle^2] = \left(\int_0^t \langle \hat{b}, e^{s\hat{B}}e_i\rangle\, ds + \langle x, e^{t\hat{B}}e_i\rangle\right)^2$$

$$+ \int_0^t \int_{\mathcal{H}_+\backslash\{0\}} \langle \xi, e^{s\hat{B}}e_i\rangle^2 m(d\xi)\, ds$$

$$+ \int_0^t \int_0^s \left\langle \hat{b}, e^{(s-u)\hat{B}}\int_{\mathcal{H}_+\backslash\{0\}} \langle \xi, e^{u\hat{B}}e_i\rangle^2 \frac{\mu(d\xi)}{\|\xi\|^2}\right\rangle du\, ds$$

$$+ \int_0^t \left\langle x, e^{(t-s)\hat{B}}\int_{\mathcal{H}_+\backslash\{0\}} \langle \xi, e^{s\hat{B}}e_i\rangle^2 \frac{\mu(d\xi)}{\|\xi\|^2}\right\rangle ds$$

and hence

$$\sum_{i=1}^{\infty}\mathbb{E}_x[\langle X_t, e_i\rangle^2] \leq 2\left\|\int_0^t e^{s\hat{B}^*}\hat{b}\, ds\right\|^2 + 2\|e^{t\hat{B}^*}x\|^2 \tag{5.5}$$

$$+ \int_0^t \int_{\mathcal{H}_+\backslash\{0\}} \|e^{s\hat{B}^*}\xi\|^2 m(d\xi)\, ds \tag{5.6}$$

$$+ \int_0^t \int_0^s \int_{\mathcal{H}_+\backslash\{0\}} \|e^{u\hat{B}^*}\xi\|^2\left\langle \hat{b}, e^{(s-u)\hat{B}}\frac{\mu(d\xi)}{\|\xi\|^2}\right\rangle du\, ds \tag{5.7}$$

$$+ \int_0^t \int_{\mathcal{H}_+\backslash\{0\}} \|e^{s\hat{B}^*}\xi\|^2\left\langle x, e^{(t-s)\hat{B}}\frac{\mu(d\xi)}{\|\xi\|^2}\right\rangle ds. \tag{5.8}$$

In the following, we show that all four terms in (5.5)–(5.8) converge as $t \to \infty$ uniformly in x, which then yields (5.4).

Note first that the adjoint semigroup $(e^{t\hat{B}^*})_{t\geq 0}$ generated by $\hat{B}^*$, the adjoint of $\hat{B}$, is also uniformly stable as $\|e^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})} = \|e^{t\hat{B}^*}\|_{\mathcal{L}(\mathcal{H})}$ for all $t \geq 0$. For the first term on the right-hand side of (5.5), we have $\int_0^t \|e^{s\hat{B}^*}\hat{b}\|\, ds \leq \frac{M}{\delta}\|\hat{b}\|$. The second term in (5.5) vanishes as $t \to \infty$ since $(e^{t\hat{B}^*})_{t\geq 0}$ is uniformly stable. Note that $s \mapsto M^2 e^{-2\delta s}\int_{\mathcal{H}_+\backslash\{0\}}\|\xi\|^2 m(d\xi)$ is an integrable majorant for the term in (5.6) and thus the integral converges for $t \to \infty$. For (5.7), note that $\langle \hat{b}, e^{(s-u)\hat{B}}\frac{\mu(d\xi)}{\|\xi\|^2}\rangle \geq 0$ for every $s, u \in \mathbb{R}_+$, which follows from the admissible parameter conditions which

imply that $\hat{b} \in \mathcal{H}_+$ and $e^{(s-u)\hat{B}}(\mathcal{H}_+) \subseteq \mathcal{H}_+$ whenever $s \geq u$. Hence we have

$$\int_0^\infty \int_0^s \int_{\mathcal{H}_+ \setminus \{0\}} \|e^{u\hat{B}^*}\xi\|^2 \left\langle \hat{b}, e^{(s-u)\hat{B}} \frac{\mu(d\xi)}{\|\xi\|^2} \right\rangle du\, ds$$

$$\leq \frac{3}{2} \frac{M^3}{\delta^2} \|\hat{b}\| \|\mu(\mathcal{H}_+ \setminus \{0\})\|.$$

Finally, note that $\int_0^t e^{-2\delta s} e^{-\delta(t-s)} ds = \frac{1}{\delta}(e^{-\delta t} - e^{-2\delta t})$ and hence the last term in (5.8) vanishes as $t \to \infty$, which can be seen from

$$\int_0^t \int_{\mathcal{H}_+ \setminus \{0\}} \|e^{s\hat{B}^*}\xi\|^2 \left\langle x, e^{(t-s)\hat{B}} \frac{\mu(d\xi)}{\|\xi\|^2} \right\rangle ds$$

$$\leq M^3 \|\mu(\mathcal{H}_+ \setminus \{0\})\| \|x\| \int_0^t e^{-2\delta s} e^{-\delta(t-s)} ds$$

$$\leq \frac{M^3}{\delta} \|\mu(\mathcal{H}_+ \setminus \{0\})\| \|x\| (e^{-\delta t} - e^{-2\delta t}). \tag{5.9}$$

Thus we obtain $\sup_{t \geq 0} \sup_{x \in \mathcal{H}_+} \mathbb{E}_x [\rho(X_t)] < \infty$, which proves the statement. $\qquad \square$

In the next result, we show first that for every $f \in \mathcal{B}_\rho(\mathcal{H}_{+,w})$, the transition semigroup $(P_t)_{t \geq 0}$ converges in $\mathcal{B}_\rho(\mathcal{H}_{+,w})$ as $t \to \infty$, and subsequently, we use this to define a continuous linear functional on $\mathcal{B}_\rho(\mathcal{H}_{+,w})$ given by the limits.

Proposition 5.6 *For all $f \in \mathcal{B}_\rho(\mathcal{H}_{+,w})$, the limit $\lim_{t \to \infty} P_t f$ in $\mathcal{B}_\rho(\mathcal{H}_{+,w})$ exists, and $\pi(f) := \lim_{t \to \infty} P_t f(x)$ defines a continuous linear functional on $\mathcal{B}_\rho(\mathcal{H}_{+,w})$.*

Proof By Proposition 5.4, we know that for every $u \in \mathcal{H}_+$,

$$\lim_{t \to \infty} (P_t e^{-\langle u, \cdot \rangle})(x) = e^{-\int_0^\infty F(\psi(s,u))\, ds}, \qquad \forall x \in \mathcal{H}_+.$$

For $u \in \mathcal{H}_+$, define $\pi_u = e^{-\int_0^\infty F(\psi(s,u))\, ds} \mathbf{1}$, where $\mathbf{1}$ denotes the constant function one. We claim that the sequence $(P_t e^{-\langle u, \cdot \rangle})_{t \geq 0}$ converges in $\mathcal{B}_\rho(\mathcal{H}_{+,w})$ to the constant function $\pi_u \in \mathcal{B}_\rho(\mathcal{H}_{+,w})$. Indeed, we have

$$\|P_t e^{-\langle u, \cdot \rangle} - \pi_u\|_\rho = \sup_{x \in \mathcal{H}_+} \frac{|(e^{-\int_0^t F(\psi(s,u))ds - \langle \psi(t,u),x \rangle} - e^{-\int_0^\infty F(\psi(s,u))ds})|}{\rho(x)}$$

$$\leq \sup_{x \in \mathcal{H}_+} \frac{|\int_t^\infty F(\psi(s,u))\, ds - \langle \psi(t,u), x \rangle|}{\rho(x)}$$

$$\leq \int_t^\infty |F(\psi(s,u))|\, ds + \|\psi(t,u)\| \sup_{x \in \mathcal{H}_+} \frac{\|x\|}{\rho(x)},$$

where we have used $\rho(x) = 1 + \|x\|^2$. The first term converges to zero due to (5.3), while the second tends to zero by Lemma 5.2.

Let $\mathcal{D} := \mathrm{lin}(\{e^{-\langle u,\,\cdot\,\rangle} : u \in \mathcal{H}_+\})$ and define π as the linear extension of π_u onto $\mathcal{D}$. In particular, we have $\lim_{t\to\infty} P_t f = \pi(f)$ in $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ for every $f \in \mathcal{D}$. In view of Proposition 5.5, we know that $\sup_{t\geq 0} \|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}))} < \infty$ and therefore $|\pi(f)| \leq \sup_{t\geq 0} \|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}))} \|f\|_{\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})}$, i.e., π is bounded on $\mathcal{D}$. Since $\mathcal{D}$ is dense in $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$, see Cox et al. [17, Lemma 4.7], this means that there exists a unique extension of π to a continuous linear functional on $\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$, which we also denote by π. We have thus proved the existence of $\pi \in \mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}), \mathbb{R})$ and it only remains to show that $P_t f \to \pi(f)$ as $t \to \infty$ for all $f \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$. But this is an immediate consequence of an $\varepsilon/3$-argument using $\sup_{t\geq 0} \|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}))} < \infty$ and $\overline{\mathcal{D}} = \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$. Thus we obtain the assertion. $\qquad\square$

In the following result, we prove that the functional π is represented by a unique probability measure on $\mathcal{B}(\mathcal{H}_+)$.

Lemma 5.7 *Let π denote the continuous linear functional in Proposition 5.6. Then there exists a unique probability measure ν on $\mathcal{B}(\mathcal{H}_+)$ such that*

$$\pi(f) = \int_{\mathcal{H}_+} f(\xi)\nu(\mathrm{d}\xi) \qquad \text{for all } f \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}), \tag{5.10}$$

and ν is inner regular on $\mathcal{B}(\mathcal{H}_+)$ when $\mathcal{H}_+$ is equipped with the weak topology.

Proof By an application of the Riesz representation theorem in Cuchiero and Teichmann [22, Theorem 2.4], there exists a unique finite signed Radon measure ν on $\mathcal{B}(\mathcal{H}_+)$ such that (5.10) and

$$\int_{\mathcal{H}_+} (1 + \|x\|^2)|\nu|(\mathrm{d}\xi) = \|\pi\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}), \mathbb{R})} \tag{5.11}$$

hold. Here $|\nu|$ denotes the total variation measure of ν. Note that ν is a Radon measure with respect to the weak topology on $\mathcal{H}_+$, which implies the statement on the inner regularity. It remains to prove that ν is a probability measure. Note that since $\lim_{t\to\infty} P_t \mathbf{1}(x) = 1$, we have $\pi(\mathbf{1}) = 1$ and hence $\nu(\mathcal{H}_+) = 1$. Moreover, as $P_t f \geq 0$ for all nonnegative $f \in C_b(\mathcal{H}_{+,\mathrm{w}})$ and all $t \geq 0$, we have $\lim_{t\to\infty} P_t f(x) \geq 0$ for all $x \in \mathcal{H}_+$ and hence $\int_{\mathcal{H}_+} f(\xi)\nu(\mathrm{d}\xi) \geq 0$ for all nonnegative $f \in C_b(\mathcal{H}_{+,\mathrm{w}})$, which implies that the measure ν is also nonnegative and hence a probability measure on $\mathcal{B}(\mathcal{H}_+)$. $\qquad\square$

In the following, we identify the linear functional π with the measure ν given by Lemma 5.7 and write π instead of ν. Finally, we show that π is indeed the unique invariant measure of $(p_t(x, \cdot))_{t\geq 0}$.

Proposition 5.8 *Let (b, B, m, μ) be an admissible parameter set such that Assumption 3.1 is satisfied and denote the associated subcritical affine Markov process on $\mathcal{H}_+$ by $(X_t)_{t\geq 0}$ and its transition kernels by $(p_t(x, \cdot))_{t\geq 0}$. Then there exists a unique invariant measure π for $(p_t(x, \cdot))_{t\geq 0}$. Moreover, for every $x \in \mathcal{H}_+$, we have*

$$\lim_{t\to\infty} \int_{\mathcal{H}_+} f(\xi)p_t(x, \mathrm{d}\xi) = \int_{\mathcal{H}_+} f(\xi)\pi(\mathrm{d}\xi), \qquad \forall f \in C_b(\mathcal{H}_{+,\mathrm{w}}), \tag{5.12}$$

and the Laplace transform of π is given by (3.2).

Proof In Proposition 5.6 and the subsequent arguments, we have already shown the existence of the Borel measure π such that (5.12) holds. It is left to show that π is the unique invariant measure. We have

$$\int_{\mathcal{H}_+} e^{-\langle u, \xi \rangle} \left(\int_{\mathcal{H}_+} p_t(x, d\xi) \pi(dx) \right) = \int_{\mathcal{H}_+} \left(\int_{\mathcal{H}_+} e^{-\langle u, \xi \rangle} p_t(x, d\xi) \right) \pi(dx)$$

$$= e^{-\phi(t,u)} \int_{\mathcal{H}_+} e^{-\langle x, \psi(t,u) \rangle} \pi(dx).$$

Note that (5.1) gives $\psi(t + s, u) = \psi(t, \psi(s, u))$ and so for every $u \in \mathcal{H}_+$, we have

$$e^{-\phi(t,u)} \int_{\mathcal{H}_+} e^{-\langle x, \psi(t,u) \rangle} \pi(dx) = e^{-\phi(t,u)} e^{-\int_0^\infty F(\psi(s, \psi(t,u))) \, ds}$$

$$= e^{-\phi(t,u)} e^{-\int_0^\infty F(\psi(t+s,u)) ds}$$

$$= e^{-\phi(t,u)} e^{-\int_t^\infty F(\psi(s,u)) ds}$$

$$= e^{-\int_0^\infty F(\psi(s,u)) ds}$$

$$= \int_{\mathcal{H}_+} e^{-\langle x, u \rangle} \pi(dx).$$

This proves the invariance of π. To prove that π is the unique invariant measure, suppose $\pi' \in \mathcal{M}(\mathcal{H}_+)$ is invariant for $(p_t(x, \cdot))_{t \geq 0}$. Then for every $u \in \mathcal{H}_+$ and $t \geq 0$, we have

$$\int_{\mathcal{H}_+} e^{-\langle x, u \rangle} \pi'(dx) = \int_{\mathcal{H}_+ \setminus \{0\}} e^{-\langle u, \xi \rangle} \left(\int_{\mathcal{H}_+ \setminus \{0\}} p_t(x, d\xi) \pi'(dx) \right)$$

$$= \int_{\mathcal{H}_+ \setminus \{0\}} e^{-\phi(t,u) - \langle x, \psi(t,u) \rangle} \pi'(dx).$$

By letting $t \to \infty$, we find that

$$\int_{\mathcal{H}_+ \setminus \{0\}} e^{-\langle u, x \rangle} \pi'(dx) = \exp\left(-\int_0^\infty F(\psi(s, u)) \, ds \right).$$

As the Laplace transform is measure-determining for measures on $\mathcal{B}(\mathcal{H}_+)$, we get $\pi' = \pi$. $\square$

Remark 5.9 The convergence in (5.12) is the weak convergence of $p_t(x, \cdot)$ to π in the weak topology on $\mathcal{H}$ as $t \to \infty$. Even though the Borel-σ-algebras of $\mathcal{H}$ equipped with the norm topology and weak topology coincide, the weak convergence is different in general. We say that $p_t(x, \cdot) \to \pi$ as $t \to \infty$ *weakly in the weak topology* on $\mathcal{H}_+$ whenever $P_t f(x) \to \int_{\mathcal{H}_+} f(\xi) \pi(d\xi)$ for all $f \in C_b(\mathcal{H}_{+,w})$.

If the stronger assumption $P_t f(x) \to \int_{\mathcal{H}_+} f(\xi)\pi(\mathrm{d}\xi)$ for all $f \in C_b(\mathcal{H}_+)$ holds, we speak of the usual *weak convergence*, i.e., $p_t(x, \cdot) \Rightarrow \pi$ as $t \to \infty$. By Merkle [61, Theorems 1 and 2], we know that weak convergence in the weak topology together with

$$\lim_{N \to \infty} \sup_{t \geq 0} p_t(x, A_N) = 0 \qquad \text{for all } \epsilon > 0,$$

where $A_N := \{\sum_{i=N}^{\infty} \langle x, e_i \rangle^2 \geq \epsilon\}$ for $N \in \mathbb{N}$, implies $p_t(x, \cdot) \Rightarrow \pi$ as $t \to \infty$. Note that in our main result in Theorem 3.3, we assert weak convergence in the *strong topology*, which will be shown below.

5.3 Proof of Theorem 3.3

Proposition 5.8 ensures the existence and uniqueness of an invariant measure π of $(p_t(x, \cdot))_{t \geq 0}$ with Laplace transform (3.2). We also proved weak convergence of $(p_t(x, \cdot))_{t \geq 0}$ to π in the weak topology as $t \to \infty$. It remains to show the convergence rates in the Wasserstein distance of order p for $p \in [1, 2]$ as in (3.3). Then convergence in the Wasserstein distance of some order $p \in [1, \infty)$ implies weak convergence (in the strong topology) and convergence of the pth absolute moment; see Villani [66, Theorem 6.9]. This implies the last assertion of Theorem 3.3. In the remainder, we prove the convergence rate in (3.3) and (3.4).

Let $p \in [1, 2]$ and as before denote by $(q_t(x, \cdot))_{t \geq 0}$ the transition kernels of an affine process associated with the admissible parameter set $(0, B, 0, \mu)$. Let $t \geq 0$, $x \in \mathcal{H}_+$ and $G \in \mathcal{C}(\delta_x, \pi)$, i.e., G is a coupling with marginals δ_x and π. Note that

$$p_t(x, \mathrm{d}y) = \int_{\mathcal{H}_+} p_t(z, \mathrm{d}y)\delta_x(\mathrm{d}z) = \int_{\mathcal{H}_+ \times \mathcal{H}_+} p_t(z, \mathrm{d}y)\mathbf{1}(z')G(\mathrm{d}z, \mathrm{d}z'),$$

and by the invariance of π, we also have

$$\pi(\mathrm{d}y) = \int_{\mathcal{H}_+} p_t(z', \mathrm{d}y)\pi(\mathrm{d}z') = \int_{\mathcal{H}_+ \times \mathcal{H}_+} p_t(z', \mathrm{d}y)\mathbf{1}(z)G(\mathrm{d}z, \mathrm{d}z').$$

Thus by the convexity property in [66, Theorem 4.8] and since $W_p \leq W_2$ for $p \in [1, 2]$, we have

$$W_p\big(p_t(x, \cdot), \pi\big) = W_p\left(\int_{\mathcal{H}_+} p_t(z, \cdot)\delta_x(\mathrm{d}z), \int_{\mathcal{H}_+} p_t(y, \cdot)\pi(\mathrm{d}y)\right)$$

$$\leq \left(\int_{\mathcal{H}_+ \times \mathcal{H}_+} W_2\big(p_t(z, \cdot), p_t(y, \cdot)\big)^p G(\mathrm{d}z, \mathrm{d}y)\right)^{1/p}.$$

$$(5.13)$$

By Lemma 5.3, we have $p_t(z, \cdot) = q_t(z, \cdot) * p_t(0, \cdot)$ for every $t \geq 0$. Thus for $H \in \mathcal{C}(q_t(z, \cdot), q_t(y, \cdot))$, we obtain by Lemma B.1 below that

$$
\begin{aligned}
&W_2\big(p_t(z, \cdot), p_t(y, \cdot)\big)^p \\
&= W_2\big(q_t(z, \cdot) * p_t(0, \cdot), q_t(y, \cdot) * p_t(0, \cdot)\big)^p \\
&\leq W_2\big(q_t(z, \cdot), q_t(y, \cdot)\big)^p \\
&\leq \left(\int_{\mathcal{H}_+ \times \mathcal{H}_+} \|\tilde{x} - \tilde{y}\|^2 H(\mathrm{d}\tilde{x}, \mathrm{d}\tilde{y}) \right)^{\frac{p}{2}} \\
&\leq \left(2 \int_{\mathcal{H}_+ \times \mathcal{H}_+} (\|\tilde{x}\|^2 + \|\tilde{y}\|^2) H(\mathrm{d}\tilde{x}, \mathrm{d}\tilde{y}) \right)^{\frac{p}{2}} \\
&= \left(2 \int_{\mathcal{H}_+ \times \mathcal{H}_+} \|\tilde{x}\|^2 q_t(z, \mathrm{d}\tilde{x}) + 2 \int_{\mathcal{H}_+ \times \mathcal{H}_+} \|\tilde{y}\|^2 q_t(y, \mathrm{d}\tilde{y}) \right)^{\frac{p}{2}}. \quad (5.14)
\end{aligned}
$$

Now recall from (5.5) that

$$
\int_{\mathcal{H}_+ \times \mathcal{H}_+} \|\tilde{x}\|^2 q_t(z, \mathrm{d}\tilde{x}) \leq 2\|\mathrm{e}^{t\hat{B}^*} z\|^2 + \int_0^t \int_{\mathcal{H}_+ \backslash \{0\}} \|\mathrm{e}^{s\hat{B}^*} \xi\|^2 \Big\langle z, \mathrm{e}^{(t-s)\hat{B}} \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \Big\rangle \mathrm{d}s,
$$

while all the other terms vanish as $\hat{b} = 0$ and $m = 0$. By the same estimations as in (5.9), we conclude that

$$
\int_{\mathcal{H}_+ \times \mathcal{H}_+} \|\tilde{x}\|^2 q_t(z, \mathrm{d}\tilde{x}) \leq 2\|\mathrm{e}^{t\hat{B}^*} z\|^2 + \int_0^t \int_{\mathcal{H}_+ \backslash \{0\}} \|\mathrm{e}^{s\hat{B}^*} \xi\|^2 \Big\langle z, \mathrm{e}^{(t-s)\hat{B}} \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \Big\rangle \mathrm{d}s
$$

$$
\leq 2M^2 \mathrm{e}^{-2\delta t} \|z\|^2 + \frac{M^3}{\delta} \|\mu(\mathcal{H}_+ \backslash \{0\}) \| \mathrm{e}^{-\delta t} \|z\|.
$$

Inserting this back into (5.14) and using the subadditivity of $x \mapsto x^{p/2}$ for $p \in [1, 2]$, we obtain

$$
\begin{aligned}
W_2\big(p_t(z, \cdot), p_t(y, \cdot)\big)^p &\leq (C_1 \mathrm{e}^{-\delta t} \|z\|)^p + (C_2 \mathrm{e}^{-\delta/2} \|z\|^{1/2})^p \\
&\quad + (C_1 \mathrm{e}^{-\delta t} \|y\|)^p + (C_2 \mathrm{e}^{-\delta/2} \|y\|^{1/2})^p \quad (5.15)
\end{aligned}
$$

for $C_1 = 2M$ and $C_2 = 2^{1/2} M^{3/2} \delta^{-1/2} \|\mu(\mathcal{H}_+ \backslash \{0\})\|^{1/2}$. Now plugging (5.15) back into (5.13) and again by the subadditivity of $x \mapsto x^{1/p}$, we obtain the desired estimates in (3.3) and (3.4).

5.4 Proof of Proposition 3.5

For every $x \in \mathcal{H}_+$, let $(p_t(x, \cdot))_{t \geq 0}$ be the transition kernels associated to the admissible parameter set (b, B, m, μ) by Theorem 2.2. Moreover, let π be the unique invariant measure of $(p_t(x, \cdot))_{t \geq 0}$ (which is independent of $x \in \mathcal{H}_+$). From part (i)

in Theorem 3.3, we know that π is inner regular. Thus from Proposition A.5 below, we obtain the existence of a unique Markov process $(X_t^\pi)_{t \geq 0}$ such that for all $f \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$, we have $\mathbb{E}_\pi[f(X_t)] = \int_{\mathcal{H}_+} P_t f(x) \pi(\mathrm{d}x)$. Moreover, since π is the invariant measure, we have for each $t \geq 0$ that

$$\int_{\mathcal{H}_+} P_t f(x) \pi(\mathrm{d}x) = \int_{\mathcal{H}_+} \left(\int_{\mathcal{H}_+} f(\xi) p_t(x, \xi) \right) \pi(\mathrm{d}x) = \int_{\mathcal{H}_+} f(\xi) \pi(\mathrm{d}\xi),$$

which implies that for all $t \geq 0$, the random variable X_t^π has the distribution π.

5.5 Proof of Proposition 3.6

Let us denote the space of all Hilbert–Schmidt operators on $\mathcal{H}$ by $\mathcal{L}_2(\mathcal{H})$ and note that $(e_i \otimes e_j)_{i,j \in \mathbb{N}}$ is an orthonormal basis of $\mathcal{L}_2(\mathcal{H})$. For every $y \in \mathcal{H}$, the operator $y \otimes y \colon \mathcal{H} \to \mathcal{H}$ defined by $(y \otimes y)(x) = \langle x, y \rangle y$ for every $x \in \mathcal{H}_+$ is a Hilbert–Schmidt operator on $\mathcal{H}$ and we can write $y \otimes y = \sum_{i,j=1}^\infty \langle y, e_i \rangle \langle y, e_j \rangle (e_i \otimes e_j)$. Note that by (5.11), we have $\int_{\mathcal{H}_+} \rho(\xi) \pi(\mathrm{d}\xi) = \|\pi\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}}), \mathbb{R})} < \infty$ and hence the absolute second moment of π is finite. This implies that

$$\int_{\mathcal{H}_+} \|y \otimes y\|_{\mathcal{L}_2(\mathcal{H})} \pi(\mathrm{d}y) \leq \int_{\mathcal{H}_+} \mathrm{Tr}(y \otimes y) \pi(\mathrm{d}y) \leq \int_{\mathcal{H}_+} \|y\|^2 \pi(\mathrm{d}y) < \infty,$$

and hence the integral $\int_{\mathcal{H}_+} (y \otimes y) \pi(\mathrm{d}y)$ is well defined in the Bochner sense. Thus it remains to compute the first two moments of the invariant measure.

Note that for every $u \in \mathcal{H}$, the linear functional $\langle u, \cdot \rangle \colon \mathcal{H} \to \mathbb{R}$ satisfies the following two properties:

i) for every $R > 0$, we have $\langle u, \cdot \rangle \in C_b(K_\mathrm{w}^R)$, where the set

$$K_\mathrm{w}^R := \{x \in \mathcal{H}_+ \colon \|x\|^2 + 1 \leq R\}$$

is compact in $\mathcal{H}_+$ equipped with the weak topology;

ii) $\lim_{R \to \infty} \sup_{x \in \mathcal{H}_+ \setminus K_\mathrm{w}^R} |\langle u, x \rangle| (1 + \|x\|^2)^{-1} = 0$.

By Dörsek and Teichmann [24, Theorem 2.7], this implies that $\langle u, \cdot \rangle \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ for all $u \in \mathcal{H}$. By Proposition 5.6, we have $P_t f \to \pi(f)$ as $t \to \infty$ for all $f \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ and hence also $P_t \langle u, \cdot \rangle \to \int_{\mathcal{H}_+} \langle u, \xi \rangle \pi(\mathrm{d}\xi)$ as $t \to \infty$. Let $(e_i)_{i \in \mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$. Then by (2.10) for $u = e_i$ for $i \in \mathbb{N}$, we have

$$\lim_{t \to \infty} P_t \langle e_i, \cdot \rangle = \lim_{t \to \infty} \left(\int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} e_i \rangle \, \mathrm{d}s + \langle x, \mathrm{e}^{t\hat{B}} e_i \rangle \right) = \int_0^\infty \langle \hat{b}, \mathrm{e}^{s\hat{B}} e_i \rangle \, \mathrm{d}s,$$

and since $\xi = \sum_{i=1}^{\infty} \langle \xi, e_i \rangle e_i$, it follows that

$$
\lim_{t \to \infty} \int_{\mathcal{H}_+} \xi \, p_t(x, d\xi) = \lim_{t \to \infty} \sum_{i=1}^{\infty} \int_{\mathcal{H}_+} \langle \xi, e_i \rangle e_i \, p_t(x, d\xi)
$$

$$
= \sum_{i=1}^{\infty} \left(\int_0^{\infty} \langle \hat{b}, e^{s\hat{B}} e_i \rangle ds \right) e_i
$$

$$
= \int_0^{\infty} e^{s\hat{B}^*} \hat{b} \, ds,
$$

where we have used

$$
\lim_{t \to \infty} \sum_{i=1}^{\infty} \left(\int_0^t \langle \hat{b}, e^{s\hat{B}} e_i \rangle \, ds \right) e_i = \sum_{i=1}^{\infty} \left(\int_0^{\infty} \langle \hat{b}, e^{s\hat{B}} e_i \rangle \, ds \right) e_i
$$

which is justified if $\lim_{N \to \infty} \sup_{t \geq 0} \sum_{i=N}^{\infty} \| \int_0^t \langle e^{s\hat{B}^*} \hat{b}, e_i \rangle e_i \, ds \| = 0$. The latter follows from $\sup_{t \geq 0} \sum_{i=N}^{\infty} \| \int_0^t \langle e^{s\hat{B}^*} \hat{b}, e_i \rangle e_i \, ds \| \leq \int_0^{\infty} \sum_{i=1}^{\infty} |\langle e^{s\hat{B}^*} \hat{b}, e_i \rangle| \, ds$ and

$$
\int_0^{\infty} \sum_{i=1}^{\infty} |\langle e^{s\hat{B}^*} \hat{b}, e_i \rangle| \, ds \leq \int_0^{\infty} \| e^{s\hat{B}^*} \hat{b} \| \, ds \leq M \|\hat{b}\| \delta^{-1} < \infty.
$$

Recalling that $\hat{b} = b + \int_{\mathcal{H}_+ \cap \{\|\xi\| > 1\}} \xi m(d\xi)$ yields (3.5).

Next we prove the desired formula for the second moments of π. For $i, j \in \mathbb{N}$, we set $g^{i,j} := \langle \cdot, e_i \rangle \langle \cdot, e_j \rangle$. From (2.11) and analogous arguments as we used in Lemma 5.5 (to show that the integrals on the right-hand side of (5.16) below exist and are finite), we find that

$$
\lim_{t \to \infty} P_t g^{i,j}(x)
$$

$$
= \left(\int_0^{\infty} \langle \hat{b}, e^{s\hat{B}} e_i \rangle \, ds \right) \left(\int_0^{\infty} \langle \hat{b}, e^{s\hat{B}} e_j \rangle \, ds \right)
$$

$$
+ \int_0^{\infty} \int_{\mathcal{H}_+ \setminus \{0\}} \langle \xi, e^{s\hat{B}} e_i \rangle \langle \xi, e^{s\hat{B}} e_j \rangle m(d\xi) \, ds
$$

$$
+ \int_0^{\infty} \int_0^s \left\langle \hat{b}, e^{(s-u)\hat{B}} \int_{\mathcal{H}_+ \setminus \{0\}} \langle \xi, e^{u\hat{B}} e_i \rangle \langle \xi, e^{u\hat{B}} e_j \rangle \frac{\mu(d\xi)}{\|\xi\|^2} \right\rangle du \, ds \qquad (5.16)
$$

holds for all $i, j \in \mathbb{N}$. The second moment formula (3.6) then follows from this and $y \otimes y = \sum_{i,j=1}^{\infty} \langle y, e_i \rangle \langle y, e_j \rangle (e_i \otimes e_j)$ once we have shown that

$$
\lim_{t \to \infty} P_t g^{i,j}(x) = \int_{\mathcal{H}_+} g^{i,j}(y) \pi(dy), \qquad i, j \in \mathbb{N}. \qquad (5.17)
$$

Since the function $g^{i,j}$ does not belong to $\mathcal{B}_\rho(\mathcal{H}_{+,w})$, we cannot obtain (5.17) directly from Proposition 5.6. However, since we have $P_t \langle \cdot, e_i \rangle \langle \cdot, e_j \rangle(x) \leq P_t \rho(x) < \infty$

for all $t \geq 0$ and $x \in \mathcal{H}_+$, we see that the function is in the larger space $B_\rho(\mathcal{H}_{+,\mathrm{w}})$ and we deduce the assertion by an additional approximation argument. Namely, define $g_n^{i,j} := g^{i,j} \wedge n$ for $n \in \mathbb{N}$. Then $g_n^{i,j} \in \mathcal{B}_\rho(\mathcal{H}_{+,\mathrm{w}})$ and we find that

$$
\left| P_t g^{i,j}(x) - \int_{\mathcal{H}_+} g^{i,j}(x)\pi(\mathrm{d}x) \right| \leq |P_t g^{i,j}(x) - P_t g_n^{i,j}(x)|
$$

$$
+ \left| P_t g_n^{i,j}(x) - \int_{\mathcal{H}_+} g_n^{i,j}(x)\pi(\mathrm{d}x) \right|
$$

$$
+ \left| \int_{\mathcal{H}_+} g_n^{i,j}(x)\pi(\mathrm{d}x) - \int_{\mathcal{H}_+} g^{i,j}(x)\pi(\mathrm{d}x) \right|.
$$

Let $\varepsilon > 0$. Take $n \in \mathbb{N}$ large enough so that $| \int_{\mathcal{H}_+} (g_n^{i,j}(x) - g^{i,j}(x))\pi(\mathrm{d}x)| < \varepsilon$. Next note that

$$
\lim_{n\to\infty} \sup_{t\geq 0} |P_t g^{i,j}(x) - P_t g_n^{i,j}(x)| \leq \lim_{n\to\infty} \sup_{t\geq 0} \mathbb{E}[\|X_t\|^2 \mathbb{1}_{\{\|X_t\|^2 > n\}}] = 0,
$$

where the equality follows from the characterisation of convergence in the Wasserstein distance (see Villani [66, Sect. 6]). Hence we find n large enough such that $|P_t g^{i,j}(x) - P_t g_n^{i,j}(x)| < \varepsilon$ holds uniformly in $t \geq 0$. Finally, for this fixed choice of n, we may use Proposition 5.6 to choose t large enough so that we have

$$
\left| P_t g^{i,j}(x) - \int_{\mathcal{H}_+} g^{i,j}(x)\pi(\mathrm{d}x) \right| < \varepsilon.
$$

Combining all these estimates proves (5.17). This completes the proof of Proposition 3.6.

6 Conclusion and outlook

In this article, we studied the long-time behaviour of affine processes on $\mathcal{H}_+$. In particular, we proved the existence and uniqueness of an invariant measure, constructed the corresponding stationary affine process and provided explicit formulas for the first and second moment of the invariant measure. Moreover, we proved ergodicity of the affine processes and established explicit and dimension-free convergence rates for the convergence in the Wasserstein distance of order $p \in [1, 2]$ of the transition kernels to the invariant measure. From a theoretical viewpoint, this article provides the first systematic study of the long-time behaviour of affine processes in a Hilbert space setting, in particular for affine processes admitting state-dependent jumps. We believe that our techniques, e.g. the use of the generalised Feller framework, can be used effectively to study the long-time behaviour of affine processes in general Hilbert spaces.

From an application point of view, we used affine processes on $\mathcal{H}_+$ to model the instantaneous covariance process in infinite-dimensional stochastic covariance models. We defined Hilbert-valued affine stochastic covariance models in the *station-*

ary covariance regime by using the stationary affine process to model the instantaneous covariance. In this context, we defined a geometric affine stochastic covariance model for forward curve dynamics in commodity markets and studied the long-time behaviour of the implied forward volatility for large forward-start dates.

Appendix A: Generalised Feller semigroups

Let (Y, τ) be a completely regular Hausdorff topological space.

Definition A.1 A function $\rho \colon Y \to (0, \infty)$ such that for every $R > 0$, the set K_R defined by $K_R := \{x \in Y \colon \rho(x) \leq R\}$ is compact is called an *admissible weight function*. The pair (Y, ρ) is called *weighted space*.

Let $\rho \colon Y \to (0, \infty)$ be an admissible weight function and assume that $\rho(x) \geq 1$ for every $x \in \mathcal{H}_+$. For $f \colon Y \to \mathbb{R}$, we define $\|f\|_\rho \in [0, \infty]$ by

$$\|f\|_\rho := \sup_{x \in Y} \frac{|f(x)|}{\rho(x)}.$$

Note that $\|\cdot\|_\rho$ defines a norm on the vector space

$$B_\rho(Y) := \{f \colon Y \to \mathbb{R} \colon \|f\|_\rho < \infty\}$$

which renders $(B_\rho(Y), \|\cdot\|_\rho)$ a Banach space. Recall that $C_b(Y)$ denotes the space of bounded $\mathbb{R}$-valued τ-continuous functions on Y. As the admissible weight function ρ satisfies $\inf_{x \in Y} \rho(x) > 0$, we have $C_b(Y) \subseteq B_\rho(Y)$.

Definition A.2 We define $\mathcal{B}_\rho(Y)$ to be the closure of $C_b(Y)$ in $B_\rho(Y)$.

The following useful characterisation of $\mathcal{B}_\rho(Y)$ is proved in Dörsek and Teichmann [24, Theorem 2.7].

Theorem A.3 *Let (Y, ρ) be a weighted space. Then we have $f \in \mathcal{B}_\rho(Y)$ if and only if $f|_{K_R} \in C(K_R)$ for all $R > 0$ and*

$$\lim_{R \to \infty} \sup_{x \in Y \setminus K_R} \frac{|f(x)|}{\rho(x)} = 0.$$

Next we recall the definition of a generalised Feller semigroup as introduced in [24, Sect. 3].

Definition A.4 A family of bounded linear operators $(P_t)_{t \geq 0}$ in $\mathcal{L}(\mathcal{B}_\rho(Y))$ is called a *generalised Feller semigroup (on $\mathcal{B}_\rho(Y)$)* if
  1) $P_0 = I$, the identity on $\mathcal{B}_\rho(Y)$;
  2) $P_{t+s} = P_t P_s$ for all $t, s \geq 0$;
  3) $\lim_{t \to 0+} P_t f(x) = f(x)$ for all $f \in \mathcal{B}_\rho(Y)$ and $x \in Y$;

4) there exist constants $C \in \mathbb{R}$ and $\varepsilon > 0$ such that $\|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(Y))} \leq C$ for all $t \in [0, \varepsilon]$;

5) $(P_t)_{t \geq 0}$ is a positive semigroup, i.e., $P_t f \geq 0$ for all $t \geq 0$ and all $f \in \mathcal{B}_\rho(Y)$ satisfying $f \geq 0$.

We make use of the following adapted version of the Kolmogorov extension theorem presented in Cuchiero and Teichmann [22, Theorem 2.11].

Proposition A.5 *Let (Y, ρ) be a weighted space and $(P_t)_{t \geq 0}$ a generalised Feller semigroup on $\mathcal{B}_\rho(Y)$ with $P_t \mathbf{1} = \mathbf{1}$ for $t \geq 0$. Then for every $\nu \in \mathcal{M}(Y)$ which is inner regular, there exist a probability space $(\Omega, \mathcal{F}, \mathbb{P}_\nu)$ with a right-continuous filtration $(\mathcal{F}_t)_{t \geq 0}$ and a Markov process $(X_t)_{t \geq 0}$ valued in Y such that $\mathbb{P}_\nu[X_0 \in A] = \nu(A)$ for every $\mathcal{B}(Y)$ and*

$$\mathbb{E}_{\mathbb{P}_\nu}[f(X_t)] = \int_Y P_t f(\xi) \nu(d\xi), \qquad t \geq 0, f \in \mathcal{B}_\rho(Y).$$

Proof In [22, Theorem 2.11], this is proved for $\nu = \delta_x$ and $x \in Y$. In the proof of [22, Theorem 2.11], it is shown that the transition kernels $p_t(x, \cdot)$ for $x \in Y$ and $t \geq 0$ given through the relation

$$P_t f(x) = \int_Y f(\xi) p_t(x, d\xi),$$

form a consistent family according to Aliprantis and Border [2, Sect. 15.6]. It is then concluded from a Kolmogorov extension theorem given by [2, Theorem 15.23] that there exists a probability measure $\mathbb{P}_{\delta_x}$ such that the assertions of the proposition hold. We draw the same conclusion for any other probability measures ν in $\mathcal{M}(Y)$ satisfying

$$\nu(A) = \sup \{\mu(K) : K \subseteq A, K \in \mathcal{K} \cap \mathcal{B}(Y)\}, \tag{A.1}$$

where $\mathcal{K}$ is a compact class in Y. Note that every weighted space Y is a Hausdorff topological space and hence the family $\mathcal{K}$ of all compact sets of Y forms a compact class; see [2, Theorem 2.31]. We thus see that every inner regular probability measure $\nu \in \mathcal{M}(Y)$ satisfies (A.1), and so the assertion follows from this, Kolmogorov's extension theorem in [2, Theorem 15.23] and analogous arguments as in [22, Theorem 2.11]. □

Appendix B: A property of the Wasserstein distance

Lemma B.1 *Let W_2 be the Wasserstein distance on $\mathcal{H}_+$. Let μ, ν, ρ be Borel probability measures on $\mathcal{H}_+$. Then $W_2(\rho * \mu, \rho * \nu) \leq W_2(\mu, \nu)$.*

Proof Let G be any coupling of (μ, ν) and G' any coupling of (ρ, ρ). For all $f, g : \mathcal{H} \to \mathbb{R}_+$, we find that

$$\int_{\mathcal{H}_+ \times \mathcal{H}_+} \big(f(x) + g(y)\big)(G' * G)(\mathrm{d}x, \mathrm{d}y)$$

$$= \int_{\mathcal{H}_+ \times \mathcal{H}_+} \int_{\mathcal{H}_+ \times \mathcal{H}_+} \big(f(x + z) + g(y + z')\big) G'(\mathrm{d}z, \mathrm{d}z') G(\mathrm{d}x, \mathrm{d}y)$$

$$= \int_{\mathcal{H}_+ \times \mathcal{H}_+} f(x + z)\rho(\mathrm{d}z)\mu(\mathrm{d}x) + \int_{\mathcal{H}_+ \times \mathcal{H}_+} g(y + z')\rho(\mathrm{d}z')\nu(\mathrm{d}y),$$

which shows that $G' * G$ is a coupling of $(\rho * \mu, \rho * \nu)$. Hence

$$\big(W_2(\rho * \mu, \rho * \nu)\big)^2$$

$$\leq \int_{\mathcal{H}_+ \times \mathcal{H}_+} \|x - y\|^2 (G' * G)(\mathrm{d}x, \mathrm{d}y)$$

$$= \int_{\mathcal{H}_+ \times \mathcal{H}_+} \int_{\mathcal{H}_+ \times \mathcal{H}_+} \|(x + z) - (y + z')\|^2 G'(\mathrm{d}z, \mathrm{d}z') G(\mathrm{d}x, \mathrm{d}y)$$

$$= \int_{\mathcal{H}_+ \times \mathcal{H}_+} \int_{\mathcal{H}_+ \times \mathcal{H}_+} \big(\|x - y\|^2 + 2\langle x - y, z - z'\rangle + \|z - z'\|^2\big)$$

$$G'(\mathrm{d}z, \mathrm{d}z') G(\mathrm{d}x, \mathrm{d}y)$$

$$= \int_{\mathcal{H}_+ \times \mathcal{H}_+} \|x - y\|^2 G(\mathrm{d}x, \mathrm{d}y) + \int_{\mathcal{H}_+ \times \mathcal{H}_+} \|z - z'\|^2 G'(\mathrm{d}z, \mathrm{d}z'),$$

where the last equality follows from the fact that G' has the same marginals so that $\int_{\mathcal{H}_+ \times \mathcal{H}_+} \langle x - y, z - z'\rangle G'(\mathrm{d}z, \mathrm{d}z') = 0$. Letting now G' be the specific coupling determined by $G'(A \times B) = \rho(\{z \in \mathcal{H}_+ : z \in A \cap B\})$ with $A, B \in \mathcal{B}(\mathcal{H}_+)$ shows that $\int_{\mathcal{H}_+ \times \mathcal{H}_+} \|z - z'\|^2 G'(\mathrm{d}z, \mathrm{d}z') = 0$. Since G was arbitrary, the assertion is proved. □

Declarations

Competing Interests The authors declare no competing interests.

References

1. Alfonsi, A., Kebaier, A., Rey, C.: Maximum likelihood estimation for Wishart processes. Stoch. Process. Appl. **126**, 3243–3282 (2016)
2. Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis. A Hitchhiker's Guide, 3rd edn. Springer, Berlin (2006)
3. Barczy, M., Döring, L., Li, Z., Pap, G.: Stationarity and ergodicity for an affine two-factor model. Adv. Appl. Probab. **46**, 878–898 (2014)
4. Barndorff-Nielsen, O.E., Stelzer, R.: Positive-definite matrix processes of finite variation. Probab. Math. Stat. **27**, 3–43 (2007)
5. Bartle, R.G., Dunford, N., Schwartz, J.: Weak compactness and vector measures. Can. J. Math. **7**, 289–305 (1955)
6. Benth, F.E., Detering, N., Krühner, P.: Stochastic Volterra integral equations and a class of first-order stochastic partial differential equations. Stochastics **94**, 1054–1076 (2022)
7. Benth, F.E., Krühner, P.: Representation of infinite-dimensional forward price models in commodity markets. Commun. Math. Stat. **2**, 47–106 (2014)
8. Benth, F.E. Lord, G. Di Nunno, G. Petersson, A.: The heat modulated infinite dimensional Heston model and its numerical approximation (2022). Preprint, Available online at https://arxiv.org/abs/2206.10166
9. Benth, F.E., Rüdiger, B., Süss, A.: Ornstein–Uhlenbeck processes in Hilbert space with non-Gaussian stochastic volatility. Stoch. Process. Appl., **128**, 461–486 (2018)
10. Benth, F.E., Sgarra, C.: A Barndorff-Nielsen and Shephard model with leverage in Hilbert space for commodity forward markets. Finance Stoch. **28**, 1035–1076 (2024)
11. Benth, F.E., Simonsen, I.C.: The Heston stochastic volatility model in Hilbert space. Stoch. Anal. Appl. **36**, 733–750 (2018)
12. Black, F.: The pricing of commodity contracts. J. Financ. Econ. **3**, 167–179 (1976)
13. Carmona, R.A., Tehranchi, M.R.: Interest Rate Models: An Infinite Dimensional Stochastic Analysis Perspective. Springer, Berlin (2006)
14. Carr, P., Madan, D.B.: Option valuation using the fast Fourier transform. J. Comput. Finance **2**, 61–73 (1999)
15. Chojnowska-Michalik, A.: On processes of Ornstein–Uhlenbeck type in Hilbert space. Stochastics **21**, 251–286 (1987)
16. Clewlow, L., Strickland, C.: Energy Derivatives: Pricing and Risk Management. Lacima Publ. (2000)
17. Cox, S., Karbach, S., Khedher, A.: Affine pure-jump processes on positive Hilbert–Schmidt operators. Stoch. Process. Appl. **151**, 191–229 (2022)
18. Cox, S., Karbach, S., Khedher, A.: An infinite-dimensional affine stochastic volatility model. Math. Finance **32**, 878–906 (2022)
19. Crosby, J.: Pricing a class of exotic commodity options in a multi-factor jump-diffusion model. Quant. Finance **8**, 471–483 (2008)
20. Cuchiero, C.: Affine and Polynomial Processes. PhD Thesis, ETH, Zurich (2011). https://doi.org/10.3929/ethz-a-006715885. Available online at
21. Cuchiero, C., Filipović, D., Mayerhofer, E., Teichmann, J.: Affine processes on positive semidefinite matrices. Ann. Appl. Probab. **21**, 397–463 (2011)
22. Cuchiero, C., Teichmann, J.: Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. J. Evol. Equ. **20**, 1301–1348 (2020)
23. Dawson, D.A., Li, Z.: Skew convolution semigroups and affine Markov processes. Ann. Probab. **34**, 1103–1142 (2006)
24. Dörsek, P., Teichmann, J.: A semigroup point of view on splitting schemes for stochastic (partial) differential equations (2010). Preprint, Available online at https://arxiv.org/abs/1011.2651
25. Duffie, D., Filipović, D., Schachermayer, W.: Affine processes and applications in finance. Ann. Appl. Probab. **13**, 984–1053 (2003)
26. Engel, K.-J., Nagel, R.: One-Parameter Semigroups for Linear Evolution Equations. Springer, Berlin (2000)
27. Filipović, D.: Consistency Problems for Heath–Jarrow–Morton Interest Rate Models. Lecture Notes in Mathematics, vol. 1760. Springer, Berlin (2001)
28. Filipović, D., Mayerhofer, E.: Affine diffusion processes: theory and applications. In: Albrecher, H., et al. (eds.) Advanced Financial Modelling, pp. 125–164. de Gruyter, Berlin (2009)
29. Filipović, D., Tapppe, S., Teichmann, J.: Term structure models driven by Wiener processes and Poisson measures: existence and positivity. SIAM J. Financ. Math. **1**, 523–554 (2010)

30. Friesen, M.: Long-time behavior for subcritical measure-valued branching processes with immigration. Potential Anal. **59**, 705–730 (2023)
31. Friesen, M., Jin, P.: On the anisotropic stable JCIR process. ALEA Lat. Am. J. Probab. Math. Stat. **17**, 643–674 (2020)
32. Friesen, M., Jin, P., Kremer, J., Rüdiger, B.: Ergodicity of affine processes on the cone of symmetric positive semidefinite matrices. Adv. Appl. Probab. **52**, 825–854 (2020)
33. Friesen, M., Jin, P., Kremer, J., Rüdiger, B.: Regularity of transition densities and ergodicity for affine jump-diffusion processes. Math. Nachr. **296**, 1117–1134 (2023)
34. Friesen, M., Jin, P., Rüdiger, B.: Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes. Ann. Appl. Probab. **30**, 2165–2195 (2020)
35. Glasserman, P., Kim, K.-K.: Moment explosions and stationary distributions in affine diffusion models. Math. Finance **20**, 1–33 (2010)
36. Gouriéroux, C., Sufana, R.: Derivative pricing with Wishart multivariate stochastic volatility. J. Bus. Econ. Stat. **28**, 438–451 (2010)
37. Hubalek, F., Keller-Ressel, M., Sgarra, C.: Geometric Asian option pricing in general affine stochastic volatility models with jumps. Quant. Finance **17**, 873–888 (2017)
38. Jacquier, A., Pannier, A., Spiliopoulos, K.: On the large-time behaviour of affine Volterra processes (2022). Preprint, Available online at https://arxiv.org/abs/2204.05270
39. Jacquier, A., Roome, P.: The small-maturity Heston forward smile. SIAM J. Financ. Math. **4**, 831–856 (2013)
40. Jacquier, A., Roome, P.: Asymptotics of forward implied volatility. SIAM J. Financ. Math. **6**, 307–351 (2015)
41. Jin, P., Kremer, J., Rüdiger, B.: Exponential ergodicity of an affine two-factor model based on the α-root process. Adv. Appl. Probab. **49**, 1144–1169 (2017)
42. Jin, P., Kremer, J., Rüdiger, B.: Moments and ergodicity of the jump-diffusion CIR process. Stochastics **91**, 974–997 (2019)
43. Jin, P., Kremer, J., Rüdiger, B.: Existence of limiting distribution for affine processes. J. Math. Anal. Appl. **486**, Article No. 123912-1–31 (2020)
44. Kallsen, J., Muhle-Karbe, J.: Exponentially affine martingales, affine measure changes and exponential moments of affine processes. Stoch. Process. Appl. **120**, 163–181 (2010)
45. Kallsen, J., Muhle-Karbe, J., Voß, M.: Pricing options on variance in affine stochastic volatility models. Math. Finance **21**, 627–641 (2011)
46. Karbach, S.: Stochastic Covariance Models in Hilbert Spaces with Jumps. PhD Thesis, University of Amsterdam (2022) https://hdl.handle.net/11245.1/094afe6c-2f6b-4ee6-9ba4-e8bb7a55c868. Available online at
47. Karbach, S.: Finite-rank approximation of affine processes on positive Hilbert–Schmidt operators (2023). Preprint, Available online at https://arxiv.org/abs/2301.06992
48. Keller-Ressel, M.: Affine Processes: Theory and Applications in Finance. PhD Thesis, TU Wien (2008) https://repositum.tuwien.at/handle/20.500.12708/13431. Available online at
49. Keller-Ressel, M.: Moment explosions and long-term behavior of affine stochastic volatility models. Math. Finance **21**, 73–98 (2011)
50. Keller-Ressel, M., Mayerhofer, E.: Exponential moments of affine processes. Ann. Appl. Probab. **25**, 714–752 (2015)
51. Keller-Ressel, M., Mijatović, A.: On the limit distributions of continuous-state branching processes with immigration. Stoch. Process. Appl. **122**, 2329–2345 (2012)
52. Koekebakker, S., Ollmar, F.: Forward curve dynamics in the Nordic electricity market. Manag. Finance **31**, 73–94 (2005)
53. Kruse, S., Nögel, U.: On the pricing of forward starting options in Heston's model on stochastic volatility. Finance Stoch. **9**, 233–250 (2005)
54. Leippold, M., Trojani, F.: Asset pricing with matrix jump diffusions (2008) https://doi.org/10.2139/ssrn.1274482. Working paper, Available online at
55. Lemmert, R., Volkmann, P.: On the positivity of semigroups of operators. Comment. Math. Univ. Carol. **39**, 483–489 (1998)
56. Lewis, D.R.: Integration with respect to vector measures. Pac. J. Math. **33**, 157–165 (1970)
57. Li, Z.: Ergodicities and exponential ergodicities of Dawson–Watanabe type processes. Theory Probab. Appl. **66**, 276–298 (2021)
58. Li, Z., Ma, C.: Asymptotic properties of estimators in a stable Cox–Ingersoll–Ross model. Stoch. Process. Appl. **125**, 3196–3233 (2015)

59. Mayerhofer, E.: Affine processes on positive semidefinite $d \times d$ matrices have jumps of finite variation in dimension $d > 1$. Stoch. Process. Appl. **122**, 3445–3459 (2012)
60. Mayerhofer, E., Stelzer, R., Vestweber, J.: Geometric ergodicity of affine processes on cones. Stoch. Process. Appl. **130**, 4141–4173 (2020)
61. Merkle, M.J.: On weak convergence of measures on Hilbert spaces. J. Multivar. Anal. **29**, 252–259 (1989)
62. Pigorsch, C., Stelzer, R.: On the definition, stationary distribution and second order structure of positive semidefinite Ornstein–Uhlenbeck type processes. Bernoulli **15**, 754–773 (2009)
63. Rosenblum, M.: On the operator equation $BX - XA = Q$. Duke Math. J. **23**, 263–269 (1956)
64. Schmidt, T., Tappe, S., Yu, W.: Infinite dimensional affine processes. Stoch. Process. Appl. **130**, 7131–7169 (2020)
65. Teng, L., Ehrhardt, M., Günther, M.: Modelling stochastic correlation. J. Math. Ind. **6**, Article No. 2-1–18 (2016)
66. Villani, C.: Optimal Transport. Old and New. Springer, Berlin (2009)
67. Wang, J.: On the exponential ergodicity of Lévy-driven Ornstein–Uhlenbeck processes. J. Appl. Probab. **49**, 990–1004 (2012)
68. Yu, W.: Infinite Dimensional Affine Term Structure Models under Incomplete Information. PhD Thesis, Universität Freiburg (2017) Available online at https://d-nb.info/1150643404/34