UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Stochastic covariance models in Hilbert spaces with jumps

Karbach, S.

Publication date
2022

Document Version
Final published version

Link to publication

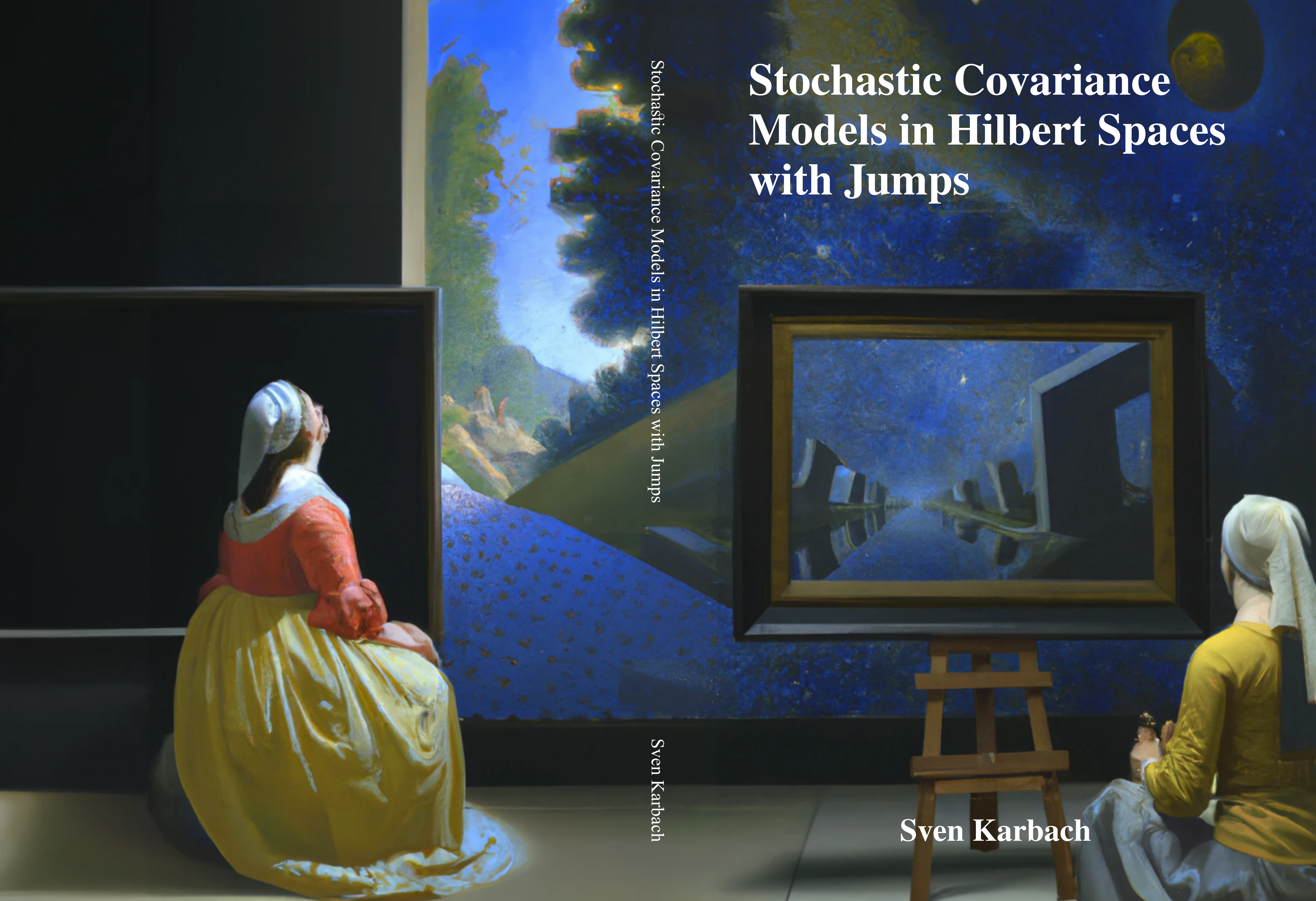
Citation for published version (APA):
Karbach, S. (2022). *Stochastic covariance models in Hilbert spaces with jumps*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Apr 2023



Stochastic Covariance Models in Hilbert Spaces with Jumps

Sven Karbach

Stochastic Covariance Models in Hilbert Spaces with Jumps
Sven Karbach

Stochastic covariance models in Hilbert spaces with jumps

Sven Karbach

The research for publication of this doctoral thesis received financial assistance from *The Dutch Research Council (NWO)* (GrantNo: C.2327.0099).

Stochastic covariance models in Hilbert spaces with jumps

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 7 december 2022, te 11.00 uur

door Sven Karbach

geboren te Wuppertal

Promotiecommissie

Promotor:	prof. dr. P.J.C. Spreij	Universiteit van Amsterdam
Copromotores:	dr. S.G. Cox	Universiteit van Amsterdam
	dr. A. Khedher	Universiteit van Amsterdam
Overige leden:	prof. dr. C. Cuchiero	University of Vienna
	prof. dr. F. Espen Benth	University of Oslo
	prof. dr. M.R.H. Mandjes	Universiteit van Amsterdam
	prof. dr. B. Rüdiger	University of Wuppertal
	prof. dr. ir. M.H. Vellekoop	Universiteit van Amsterdam
	dr. ir. E.M.M. Winands	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Contents

CHAPTER 1

Introduction

This thesis is devoted to the theoretical aspects of stochastic covariance models with jumps taking values in finite- or infinite-dimensional Hilbert spaces. More precisely, we provide the mathematical foundations for two specific classes of stochastic processes with jumps taking values in cones of positive operators and we demonstrate their potential to model the *instantaneous stochastic covariance process* in stochastic covariance models. In particular, we study applications of the two model classes to forward curve and multivariate asset price modeling.

The thesis is divided into two parts, each of which consists of separate chapters based on the authors publications [38, 39], submitted preprints [67, 17] and working papers (Chapter 4 and 6). In the first part of the thesis, we introduce and study various aspects of *affine pure-jump processes* taking values in the cone of positive self-adjoint Hilbert-Schmidt operators and extend the class of *affine stochastic covariance models* from a multivariate to an infinite-dimensional Hilbert space setting. In the second part we, deal with *positive multivariate continuous-time autoregressive moving-average* (positive MCARMA) processes and study the second order moment structure of a class of multivariate stochastic covariance models based on positive semi-definite MCARMA processes.
We provide a substantive introduction and an outline of our main contributions at the beginning of each of the two parts of this thesis. Therefore, in the remainder of this general introduction, we restrict ourselves to outlining the general ideas behind stochastic covariance modeling and discussing the main application examples for the two stochastic covariance models considered in this thesis.

A generic stochastic covariance model Broadly speaking, a stochastic covariance model is used to model a stochastic process exhibiting an instantaneous covariance process that itself is also stochastic. In particular, a stochastic covariance model consists of two components: A model for the quantity of interest, e.g. the (logarithmic) price vector process of some correlated financial assets, see for example [77, 73, 9, 10, 23, 21, 104], and a model for its instantaneous covariance process. More specifically, the stochastic covariance models that we are concerned with fall into the following generic setup (or slight generalization of it): We let $(Y_t)_{t\geq 0}$ be a stochastic process with values in some finite- or infinite-dimensional separable Hilbert space H given by a stochastic differential equation of the form

$$\begin{cases} \mathrm{d}Y_t = \big(A(Y_t) + G(X_t)\big)\,\mathrm{d}t + X_t^{1/2}\,\mathrm{d}W_t, \quad t > 0, \\ Y_0 = y, \end{cases} \tag{1.1}$$

where we assume that A is a continuous linear operator on H, G is continuous and maps X_t (for every $t \geq 0$) affine linearly into the space H and $(W_t)_{t\geq 0}$ denotes a H-cylindrical Brownian motion. Moreover, we let $(X_t)_{t\geq 0}$ be a positive operator-valued stochastic process such that the operator square-root process $(X_t^{1/2})_{t\geq 0}$ is integrable with respect to $(W_t)_{t\geq 0}$. We note that the model (1.1) is specified such that the stochastic quadratic-covariation process of $(Y_t)_{t\geq 0}$ (see [117]) is given by

$$\langle\langle Y, Y\rangle\rangle_t = \int_0^t X_s\,\mathrm{d}s, \quad t \geq 0, \tag{1.2}$$

which means that the covariance structure of the process Y is completely determined by the process X. The joint process (Y, X) is what we call the *stochastic covariance model given by* (1.1). In particular, we call X the *instantaneous covariance process of the model* (1.1), which is justified by the relation (1.2).

The main objective in stochastic covariance modeling is the appropriate specification of the instantaneous covariance process such that the stochastic covariance model is both *flexible* and *tractable*. We say that the model is flexible if the theoretic covariance structure implied by the model matches a wide variety of observed realized (cross)-covariance structures of the underlying process and displays the stylized facts of observed data. In practice, the flexibility of the model must be met with a reasonable tractability, as for most applications stochastic covariance models are supposed to be feasible for, e.g. conducting statistical inference, simulations or, more finance specific, option pricing.
In this thesis we demonstrate that affine pure-jump processes on positive Hilbert-Schmidt operators and MCARMA processes on positive semi-definite matrices are both flexible and tractable classes for modeling the instantaneous covariance process in an infinite-dimensional, respectively, multivariate setting. Next, we introduce the main applications for the two stochastic covariance models, where conversely to their appearance in this thesis we, begin with the multivariate case.

Stochastic covariance models for multivariate asset price dynamics Most stochastic covariance models for multivariate asset price dynamics in the literature [10, 104, 126, 73] are of the form (1.1) with $H = \mathbb{R}_d$, $A = 0$ and $(W_t)_{t\geq 0}$ a standard Brownian motion on $\mathbb{R}_d$. For the operator G a common modeling option is to assume $G(X) = \alpha + X\beta$, where $\alpha \in \mathbb{R}_d$ is a constant drift term and $\beta \in \mathbb{R}_d$ denotes the risk-premium. Classical choices for the instantaneous covariance in these multivariate asset price models are either matrix-valued pure-jump models, such as positive semi-definite Ornstein-Uhlenbeck (OU) processes or superposition of these, so called supOU processes, see e.g. [10, 126], pure diffusion-based models such as Wishart processes, see [31, 73], or mixtures of both, e.g. affine jump-diffusions on positive semi-definite matrices, see [104, 42]. Instantaneous covariance processes with jumps provide good models for the financial time series in stock, energy or fixed-income markets, as it is illustrated, e.g., in [57, 24, 37, 104]. We refer in particular to [104] where the authors discussed convincing empirical evidence for (state dependent) jumps in the instantaneous covariance process of multivariate stochastic covariance models.
In the second part of the thesis, we propose to model the instantaneous covariance process in multivariate stochastic covariance models by positive semi-definite MCARMA processes driven by Lévy subordinators. This model class extends the popular multivariate Barndorff–Nielsen-Shepard (BNS) stochastic volatility model (first order MCARMA) to higher-order MCARMA based models that have the potential to model certain short-memory features that are often observed in realized covariance processes and are not captured by the classical BNS model.

Stochastic covariance models for forward curve dynamics In infinite dimensions, stochastic covariance models of (roughly) the form (1.1) have been studied in [21, 23, 22]. These models are applied in the context of forward curve modeling in commodity or fixed-income markets formulated in the Heath-Jarrow-Morton-Musiela (HJMM) framework, see e.g. [60, 32, 18]. In this setting, the Hilbert space H is taken from a class of weighted Sobolev spaces that match the economic reasoning about forward curves, see [60]. Moreover, the operator A under the HJMM formulation of the forward curve dynamics turns out to be unbounded; it is given by the first derivative, i.e. $A = \frac{\partial}{\partial x}$. The models in [21, 22] can be considered as operator-valued extensions of the multivariate BNS model. Indeed, the instantaneous covariance is given by a Lévy driven OU process taking values in the cone of positive Hilbert-Schmidt operators, which can be viewed as the natural infinite-dimensional analog of the positive semi-definite matrices.
In the first part of the thesis, we extend the operator-valued BNS model towards *affine stochastic covariance models* in Hilbert spaces, where the instantaneous covariance is modeled by an affine process on positive Hilbert-Schmidt operators admitting state-dependent jump intensities. We view this class as the infinite-dimensional version of the multivariate affine stochastic covariance model in [42].

Part I

Infinite-Dimensional Affine Stochastic Covariance Models

Introduction to Part I: Infinite-Dimensional Affine Stochastic Covariance Models

In order to model infinite-dimensional instantaneous covariance processes, we introduce the following setting: Let $(H, \langle\cdot,\cdot\rangle_H)$ be a separable Hilbert space and denote by $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ the Hilbert space of all self-adjoint Hilbert-Schmidt operators on H equipped with the trace inner-product $\langle x, y\rangle \coloneqq \mathrm{Tr}(xy)$ for $x, y \in \mathcal{H}$. We denote the cone of all positive self-adjoint Hilbert-Schmidt operators by $\mathcal{H}^+$ and let $(X, (\mathbb{P}_x)_{x\in\mathcal{H}^+})$ be a time-homogeneous Markov process with values in the cone $\mathcal{H}^+$, where $\mathbb{P}_x$ denotes the law of X given $X_0 = x \in \mathcal{H}^+$. As mentioned in the introduction, in this Hilbert space-valued setting we focus on instantaneous covariance processes $(X_t)_{t\geq 0}$ that are commonly known as *affine processes*, see e.g. [52, 42, 43]. This roughly speaking means that the Laplace transform of X_t, for all $t \geq 0$, is of an exponential affine form in the initial value $X_0 = x \in \mathcal{H}^+$, i.e. we have

$$\mathbb{E}_{\mathbb{P}_x}\left[\mathrm{e}^{-\langle X_t,u\rangle}\right] = \mathrm{e}^{-\phi(t,u)-\langle x,\psi(t,u)\rangle}, \qquad t \geq 0,\, u \in \mathcal{H}^+, \tag{1.3}$$

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 solutions to a pair of *generalized Riccati differential equations.* The appeal of affine processes lays in their good *tractability* entailed by the *affine transform formula* (1.3). Indeed, for affine processes computations of the Fourier-Laplace transform reduce to mere evaluations of the functions ϕ and ψ, provided that both functions are (approximately) known. The affine class is also recognized for its *flexibility* as, depending on the state-space, it admits various modeling options including processes with drift depending affine linearly on the state of the process, diffusion components and jumps governed by a jump-intensity measure depending affine linearly on the state.

In finite dimensions affine processes and their applications were studied by various authors on state spaces including the canonical state space $\mathbb{R}^d_+ \times \mathbb{R}^n$ $(d, n \in \mathbb{N})$, see e.g. [52, 93, 49, 62, 95, 90], and the cone $\mathbb{S}^+_d$ of positive semi-definite $d \times d$-matrices, see [42]. More recently, it became increasingly popular to study their infinite-dimensional versions, see e.g. [45, 134, 74]. Affine processes on $\mathcal{H}^+$, as introduced and studied in this thesis, can be considered as the natural infinite-dimensional analog of affine processes on $\mathbb{S}^+_d$. The relevance of affine processes on the cone $\mathbb{S}^+_d$, in particular for applications in multivariate stochastic covariance modeling is widely recognized, see [42, 104, 10, 73].

In the first part of this thesis we are concerned with the infinite-dimensional extension of the ideas in [42, 104], i.e. we let H be an *infinite-dimensional Hilbert space* and propose to model the instantaneous covariance process in stochastic covariance models on H by an affine process on $\mathcal{H}^+$.

Indeed, in the forthcoming five chapters we provide the mathematical foundations for infinite-dimensional affine stochastic covariance models with jumps and demonstrate the capability of the affine class on $\mathcal{H}^+$ to model the instantaneous covariance process in this infinite-dimensional environment. In the following paragraphs we give a brief outline of the first part of the thesis and highlight our main contributions to the general understanding of infinite-dimensional operator-valued affine processes and their applications to stochastic covariance modeling in Hilbert spaces.

Existence of affine pure-jump processes on positive Hilbert-Schmidt operators

In Chapter 2 we prove the existence of a broad class of affine pure-jump processes with values in the cone of positive self-adjoint Hilbert-Schmidt operators specified by a set of *admissible parameters*. The parameters are such that an associated pair of *operator-valued generalized Riccati equations* modulated by this parameter set admits a unique global solution $(\phi(\cdot,u),\psi(\cdot,u))$ on the convex cone $\mathbb{R}^+\times\mathcal{H}^+$. Our main objective is to prove that for all such solutions $(\phi(\cdot,u),\psi(\cdot,u))$ there exists an associated Markov process $(X,(\mathbb{P}_x)_{x\in\mathcal{H}^+})$ satisfying the *affine transform formula* (1.3) for this choice of ϕ and ψ. We formulate our main existence result (Theorem 2.8) in an abbreviated form as follows:

Theorem 1.1. *Let (b,B,m,μ) be a tuple consisting of a vector $b\in\mathcal{H}$, a bounded linear operator $B\in\mathcal{L}(\mathcal{H})$, a measure m on the Borel-σ-algebra $\mathcal{B}(\mathcal{H}^+\setminus\{0\})$ and a $\mathcal{H}$-valued measure μ on $\mathcal{B}(\mathcal{H}^+\setminus\{0\})$, satisfying the* admissibility assumptions *posed in Definition 2.3 below. Then there exists an affine Markov process $(X,(\mathbb{P}_x)_{x\in\mathcal{H}^+})$ with values in $\mathcal{H}^+$, such that the functions ϕ and ψ in* (1.3) *are the unique solutions to the so called* generalized Riccati equations*:*

$$\begin{cases} \dfrac{\partial\phi(t,u)}{\partial t}=F(\psi(t,u)), \qquad \phi(0,u)=0\,, & \text{(1.4a)}\\ \dfrac{\partial\psi(t,u)}{\partial t}=R(\psi(t,u)), \qquad \psi(0,u)=u\,, & \text{(1.4b)}\end{cases}$$

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

$$F(u)=\langle b,u\rangle-\int_{\mathcal{H}^+\setminus\{0\}}\big(\mathrm{e}^{-\langle\xi,u\rangle}-1+\mathbf{1}_{\|\xi\|\leq 1}(\xi)\langle\xi,u\rangle\big)\,m(\mathrm{d}\xi), \tag{1.5a}$$

$$R(u)=B^*(u)-\int_{\mathcal{H}^+\setminus\{0\}}\big(\mathrm{e}^{-\langle\xi,u\rangle}-1+\mathbf{1}_{\|\xi\|\leq 1}(\xi)\langle\xi,u\rangle\big)\,\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\,. \tag{1.5b}$$

Without providing the exact parameter conditions, the implicit claim here is, that the class of affine processes $(X, (\mathbb{P}_x)_{x\in\mathcal{H}^+})$ given by Theorem 1.1 is *flexible*. More precisely, we show that the affine class under consideration admits novel modeling options on positive Hilbert-Schmidt operators, such as processes with drift term $b + B(X_t)$, depending affine linearly on the state of the process, together with jumps of possibly infinite-variation governed by a state-dependent jump-intensity measure of the form $\nu(X_t, \mathrm{d}\xi) \coloneqq m(\mathrm{d}\xi) + \|\xi\|^{-2}\langle X_t, \mu(\mathrm{d}\xi)\rangle$ (see Proposition 3.5).

Our main contribution is the inclusion of state-dependent jump-intensities in the dynamics of X, which is reflected by the model parameter μ. Indeed, notice that for $\mu = 0$ the affine process X given by Theorem 1.1 is simply an OU-type process on $\mathcal{H}^+$ driven by an $\mathcal{H}^+$-valued Lévy process with characteristics $(b, 0, m)$, which was introduced and studied in [21]. Conceptually, we view the inclusion of state-dependent jump-intensities to an OU-type model as a shift from a linear infinite-dimensional SDE (the SDE for the OU process) to a non-linear infinite-dimensional SDE driven by state-dependent jump-noise of possibly infinite-variation (the SDE for the affine process). Solving these equations associated with affine processes is, in general, an intricate problem, especially so in this infinite-dimensional non-locally compact setting.

In Chapter 2 we solve the SDE associated with the class of affine pure-jump processes in a stochastically weak sense by using the theory of *generalized Feller processes*. The biggest challenges we face is that like many infinite-dimensional cones, the cone of positive self-adjoint Hilbert-Schmidt operators has empty interior and is non-locally compact. One consequence is that one cannot employ classical localization arguments as, e.g. in [42], to establish the existence of the desired affine processes (and solutions to the generalized Riccati equations). Another consequence is that it is difficult to incorporate a diffusion term. Indeed, it remains an open question whether and under what conditions infinite-dimensional affine processes on positive Hilbert-Schmidt operators allow for a diffusion term.

Affine stochastic covariance models in infinite-dimensions

Our main motivation for studying affine processes on $\mathcal{H}^+$ is their potential to model the instantaneous covariance process in infinite-dimensional stochastic covariance models. This is motivated by the finite-dimensional case studied in [73, 34, 104, 125]. In infinite-dimensions, the stochastic covariance model we propose, generalizes the *operator-valued Barndorff–Nielsen-Shepard model* in [21] to a model admitting state-dependent jump intensities in the instantaneous covariance process, while maintaining the tractable *affine* property. Proving well-posedness and establishing the affine property of the stochastic covariance model with affine pure-jump instantaneous covariance process is our main concern in Chapter 3.

More specifically, we demonstrate that stochastic covariance models of the form in (1.6) below, with instantaneous covariance process modeled by an affine process on $\mathcal{H}^+$ given by Theorem 1.1 are well-defined and the tractable affine form of the Laplace-transform of the instantaneous covariance process inherits to the characteristic function of the joint-model. More precisely, we state our main result (Theorem 3.14) on the existence of affine stochastic covariance models as follows:

Theorem 1.2. *Let X be an affine process on $\mathcal{H}^+$ associated with an admissible parameter set (b, B, m, μ) given by Theorem 1.1 and assume that X has càdlàg paths. Moreover, let $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ be the generator of a strongly continuous semigroup on H, $D \in \mathcal{H}^+$, $(W_t)_{t\geq 0}$ a cylindrical Brownian motion on H and let $(Y_t)_{t\geq 0}$ be the solution of*

$$\begin{cases} \mathrm{d}Y_t = \mathcal{A}Y_t + D^{1/2}X_t^{1/2}\,\mathrm{d}W_t, & t > 0, \\ Y_0 = y. \end{cases} \tag{1.6}$$

Then for every initial value $(Y_0, X_0) = (y, x) \in H \times \mathcal{H}^+$ and every $u = (u_1, u_2) \in \mathrm{i}H \times \mathcal{H}^+$ we have

$$\mathbb{E}\left[\mathrm{e}^{\langle Y_t, u_1\rangle_H - \langle X_t, u_2\rangle}\right] = \mathrm{e}^{-\Phi(t,u) + \langle y, \psi_1(t,u)\rangle_H - \langle x, \psi_2(t,u)\rangle}, \quad t \geq 0\,, \tag{1.7}$$

where $(\Phi(\cdot, u), (\psi_1(\cdot, u), \psi_2(\cdot, u)))$ is the unique (mild) solution of the equations

$$\begin{cases} \dfrac{\partial \Phi(t,u)}{\partial t} = F(\psi_2(t,u)), & \Phi(0,u) = 0, \\ \psi_1(t,u) = u_1 - \mathrm{i}\mathcal{A}^*\left(\mathrm{i}\displaystyle\int_0^t \psi_1(s,u)\,\mathrm{d}s\right), & \psi_1(0,u) = u_1, \\ \dfrac{\partial \psi_2(t,u)}{\partial t} = R\big(\psi_2(t,u)\big) - \dfrac{1}{2}\big(D^{1/2}\psi_1(t,u)\big)^{\otimes 2}, & \psi_2(0,u) = u_2, \end{cases}$$

where the functions F and R are as in (1.5).

Having established the general existence and affine property of the *affine stochastic covariance model* (Y, X) in Theorem 1.2, we present various examples of stochastic covariance models included in this class (see Section 3.4). First, we show that the operator-valued Barndorff–Nielsen-Shepard model from [21] is indeed a special case of our model. But more importantly, our proposed model class contains models where the instantaneous covariance process X has state-dependent jump-intensity of possibly infinite variation. This, although no diffusion part is admissible in our model class, has the potential to resemble some diffusive-like behavior in the instantaneous covariance component. In particular, our model class has the potential to model *volatility clustering* in infinite-dimensional stochastic covariance models, see Section 3.4.

Pricing options on forwards modeled in infinite-dimensional affine stochastic covariance models

In Chapter 4 we introduce a *geometric affine stochastic covariance model* for forward curve dynamics formulated in the HJMM framework, see [60, 32, 18, 19]. More precisely, we propose to model the dynamics of the logarithmic forward price curves in, e.g., commodity markets by an affine stochastic covariance model (Y, X) given by Theorem 1.2 with additional drift term $G(X_t)$ in the Y dynamics. In this setting, the underlying space H stems from a class of Hilbert spaces H_w, proposed in [60], containing viable forward curves (see also Section 4.3). Moreover, in the HJMM framework one has $\mathcal{A} = \frac{\partial}{\partial x}$, G satisfying a certain *drift condition* (see Lemma 4.6), $D \in \mathcal{L}_2(H_w)$ being self-adjoint and positive and $(W_t)_{t\geq 0}$ a Brownian motion on H_w. Denoting by $f_t(x)$ the price of the forward contract at time $t \geq 0$ with time-to-maturity $x \geq 0$, we propose to model the *arbitrage-free* forward price as

$$f_t(x) \coloneqq \exp\big(\delta_x(Y_t)\big) \quad \text{for } x, t \geq 0,$$

where δ_x denotes the point evaluation at $x \geq 0$. Within this setting we price a *European call option* written on a forward contract $F(T, \hat{T}) = f_T(\hat{T} - T)$ $0 \leq T \leq \hat{T}$ with pay-off function at time T given by

$$\max\big(F(T, \hat{T}) - K, 0\big).$$

In particular, in Proposition 4.7 we present quasi-explicit option price formulas in terms of the solutions to an associated generalized Riccati equation and the Fourier transform of the pay-off function. Our derivations show that the virtues of the affine class for option-pricing via Fourier-methods, which is well recognized in the finite-dimensional setting, see [33, 53, 52], remain valid in infinite-dimensions.

Long-time behavior of affine processes on positive Hilbert-Schmidt operators and the stationary covariance regime

In Chapter 5 we study the long-time behavior of affine processes on $\mathcal{H}^+$. The long-time behavior of instantaneous covariance processes plays an important role in the calibration of stochastic covariance models, see [1, 105].
Note that for an *affine* process $(X_t)_{t\geq 0}$ on $\mathcal{H}^+$ with transition-kernels $(p_t(x, \cdot))_{t\geq 0}$, we can write the affine transform formula (1.3) as

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

In Chapter 5, we then study the existence of a unique *invariant measure* π of $(p_t(x, \cdot))_{t\geq 0}$, the existence of stationary affine processes and ergodicity.

An invariant measure π on $\mathcal{H}^+$ of X with transition-kernels $(p_t(x,\cdot))_{t\geq 0}$ satisfies

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

Heuristically, if the transition kernels $(p_t(x,\cdot))_{t\geq 0}$ converge weakly to a probability measure π on $\mathcal{H}^+$, then π is the invariant measure of X. Indeed, we prove the existence of unique limit distribution of $(p_t(x,\cdot))_{t\geq 0}$ as t tends to infinity and, moreover, we show the following assertions (Theorem 5.3):

Theorem 1.3. *Let $(X_t)_{t\geq 0}$ be an affine process on $\mathcal{H}^+$ with transition kernels $(p_t(x,\cdot))_{t\geq 0}$ and assume that all eigenvalues of the linear effective drift of X given by* (5.1) *have strictly negative real-parts. Then the following holds 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 $x\in\mathcal{H}^+$ the sequence $(p_t(x,\cdot))_{t\geq 0}$ converges exponentially fast to the invariant measure π in the Wasserstein distance of order two as $t\to\infty$.

iii) There exists a Markov process X^π such that the transition kernels of X are given by $(p_t(x,\cdot))_{t\geq 0}$ and the distribution of X^π_t is equal to π for all $t\geq 0$.

In addition to the assertions of Theorem 1.3, we also provide explicit rates for the convergence of $(p_t(x,\cdot))_{t\geq 0}$ to π in Wasserstein distance of order $q\in[1,2]$ in terms of the second moment of the invariant measure and the spectral bound of the effective drift (see Theorem 5.3 and Proposition 5.6).

The existence of stationary affine processes on $\mathcal{H}^+$ allows us to introduce the *stationary covariance regime* for infinite-dimensional affine stochastic covariance models known from the univariate case in [94]. In the stationary covariance regime, we replace the affine instantaneous covariance process X with its stationary version given by Theorem 1.3. In this context, we establish an intimate connection between the implied forward volatility smile for large forward-start dates in the (non-stationary) geometric affine stochastic covariance model for forward curve dynamics and the implied volatility of a plain vanilla option on forwards modeled under the stationary covariance regime (Section 5.4.2 and Proposition 5.12).

Finite-rank approximations of affine stochastic covariance models in infinite-dimensions

We motivated the use of $\mathcal{H}^+$-valued affine processes as a model for the instantaneous covariance process in stochastic covariance models by their tractable affine structure, i.e. that the Fourier-Laplace transform is quasi-explicitly given in terms of the solution to the generalized Riccati equation (1.4).

This is particularly relevant when numerical approximations of the SDE describing the affine process X converge slowly and/or are difficult to implement, e.g. because the SDE contains a square-root term or admits state-dependent jump-intensity. However, the tractability of the affine class relies on the assumption that the generalized Riccati equations can be (approximately) solved, which is not immediately clear, especially not in this infinite-dimensional setting.

Therefore, we study in Chapter 6 a feasible approximation scheme for affine processes on positive Hilbert-Schmidt operators. More precisely, we reduce the operator-valued generalized Riccati equations (1.4) to finite-dimensional (essentially matrix-valued) equations using Galerkin-type approximations, and provide explicit upper bounds for the error

$$\sup_{t\in[0,T]} \|\psi_d(t,\mathbf{P}_d(u)) - \psi(t,u)\|,$$

where for $d \in \mathbb{N}$, we let $\mathbf{P}_d \colon \mathcal{H} \to \mathcal{H}_d$ be an appropriately chosen projection onto a finite-dimensional subspace $\mathcal{H}_d$ (in particular such that $\mathbf{P}_d(\mathcal{H}^+)|_{H_d} \simeq \mathcal{S}_+^d$ for some d-dimensional subspace H_d of H) and for all $u \in \mathcal{H}^+$ let $\psi_d(\cdot,\mathbf{P}_d(u))$ denotes the associated $d \times d$-dimensional Galerkin approximation of $\psi(\cdot,u)$.

In addition to that, we construct a sequence of positive finite-rank operator valued affine processes $(X^d)_{d\in\mathbb{N}}$ associated with the Galerkin approximations $(\phi_d(\cdot,\mathbf{P}_d(u)),\psi_d(\cdot,\mathbf{P}_d(u))_{d\in\mathbb{N}}$ and we prove its weak convergence in the Skorohod space to the affine process X associated with ϕ and ψ as the rank d tends to infinity. This approach also solves the open question on the existence of càdlàg paths, see Theorem 1.2, and provides an alternative proof for the existence of affine pure-jump processes on $\mathcal{H}^+$ by exploiting the connections to their matrix-valued versions.

CHAPTER 2

Affine Pure-Jump Processes on Positive Hilbert-Schmidt Operators

Abstract of the chapter In this chapter we prove the existence of a broad class of affine pure-jump Markov processes with values in the cone of positive self-adjoint Hilbert-Schmidt operators, defined on an infinite-dimensional separable Hilbert space. This class of affine processes is the natural infinite-dimensional analog of affine (pure-jump) processes on positive semi-definite matrices. As for their matrix-valued versions, the processes we construct allow for a drift depending affine linearly on the state, as well as jumps governed by a jump-intensity measure that depends affine linearly on the state. The fact that the cone of positive self-adjoint Hilbert-Schmidt operators has empty interior and is not locally compact calls for a new approach to proving existence: Instead of using standard localization techniques and solving the martingale problem, we employ the theory of *generalized Feller semigroups*, which was introduced in [51] and further developed in [45]. Our approach requires a second moment condition on the jump measures involved, consequently, we obtain convenient explicit formulas for the first and second moments of the affine processes.

This chapter is based on [38]:
Cox, S., Karbach, S., and Khedher, A.
Affine pure-jump processes on positive Hilbert–Schmidt operators.
Stochastic Processes and their Applications 151 (2022), 191–229.

2.1 Introduction

In this chapter, we are concerned with proving the existence of a broad class of affine Markov processes with values in the cone of positive self-adjoint Hilbert-Schmidt operators. More specifically, the processes under consideration have the following model parameters: a constant drift vector b, a linear drift term B, a constant jump measure m, and a state-dependent jump measure μ. To prove the existence of class of affine processes on $\mathcal{H}^+$, we first study the associated generalized Riccati equations (1.4) and prove the existence of a unique global solution $(\phi(\cdot,u),\psi(\cdot,u))$ on the cone $\mathbb{R}^+\times\mathcal{H}^+$. In order to prove the existence of a unique associated affine process X, our main objective is to show that the operators $(P_t)_{t\geq 0}$ defined as $P_t\exp(-\langle\cdot,u\rangle)(x)=\exp(-\phi(t,u)-\langle\cdot,\psi(t,u)\rangle)(x)$ for $x\in\mathcal{H}^+$ give rise to a proper Markovian semigroup (on some reasonably large space of functions) that can be associated with the affine process X.
We are facing two main challenges in proving the existence of this class of processes: First, the state-space $\mathcal{H}^+$ is not locally compact and consequently tools known from the classical Feller theory are not applicable. To compensate for that, we use the *generalized Feller theory* introduced in [51]. This approach copes with the non-locally compact setting, but computations appear to be more intricate compared to the classical setting. The second challenge we face is that like many infinite-dimensional cones, the cone of positive self-adjoint Hilbert-Schmidt operators has empty interior. One consequence is that one cannot employ classical localization arguments to establish existence of the desired processes; we take a different approach outlined below. Another consequence is that it is difficult to incorporate a diffusion term. Indeed, as pointed out in the introduction, it remains an open question whether and under what conditions, infinite-dimensional affine processes on positive Hilbert-Schmidt operators allow for a diffusion term.

Our new approach to proving existence involves approximating the transition semigroup associated with the desired Markov process by simpler transition semigroups corresponding to affine finite-activity jump processes. We then exploit the aforementioned generalized Feller theory, in particular two approximation results from [45] as well as a version of Kolmogorov's extension theorem to show that the limiting semigroup gives rise to a generalized Feller process. Note that the idea of proving the existence of affine processes with jumps of infinite activity (in our case even infinite-variation) through an approximation with simpler affine processes was already used on, e.g. convex sets in finite dimensions, where it is known that affine processes are (classical) Feller processes (see [52] and [42]). However, our approach is somewhat different, and a considerable amount of effort goes into verifying that the approximating generalized Feller semigroups satisfy all necessary conditions for convergence. In particular, a subtle analysis of the regularity of ϕ and ψ is conducted and uniform growth bounds for the approximations derived.

2.1.1 Layout of the chapter

In Section 2.2 we provide the definition of *admissible parameter sets* and present the comprehensive version of our main result (Theorem 2.8) on the existence of affine pure-jump processes on the cone of positive self-adjoint Hilbert-Schmidt operators. Moreover, we specify the exact form of the *weak generator* of these affine Markov processes on the linear span of the Fourier basis elements in terms of the introduced admissible parameter set. A brief outline of the proof of Theorem 2.8 is presented in Section 2.2 and the full proof is left to Section 2.4. In Section 2.3 we prove existence and uniqueness of the solution to the generalized Riccati equations (1.4) and moreover study its regularity with respect to the initial value. In Section 2.4 we are making use of the results in Section 2.3 and some intricate approximation techniques for generalized Feller semigroups to complete the proof of Theorem 2.8.

2.1.2 Notation

We set $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \{0, 1, \dots\}$. For a vector space X and $U \subseteq X$ we denote the linear span of U by $\operatorname{lin}(U)$. For (X, τ) a topological space and $S \subseteq X$ we let $\mathcal{B}(S)$ denote the *Borel-σ-algebra* generated by the relative topology on S. Let $(H, \langle \cdot, \cdot \rangle_H)$ be a Hilbert space. Then we denote by $C(S, H)$ the space of H-valued functions on S that are continuous with respect to the relative topology and we denote by $C_b(S, H)$ the space of bounded H-valued continuous functions on S. This is a Banach space when endowed with the supremum norm $\|\cdot\|_{C(S)}$. Notice that when $H = \mathbb{R}$, we typically omit H in the notation: $C(S) := C(S, \mathbb{R})$. Let $\mathcal{L}(X)$ denote the space of *bounded linear operators* from a Banach space X to X. This is a Banach space when equipped with the operator norm $\| \cdot \|_{\mathcal{L}(X)}$. If $\mathcal{G}$ is a linear operator on a Banach space X, we denote its *domain* by $\operatorname{dom}(\mathcal{G})$ and denote by I the identity in $\mathcal{L}(X)$. We denote unbounded operators by a calligraphic font and bounded ones by the standard font, e.g., $\mathcal{G}$ versus G. Let $\mathcal{L}^{(2)}(H \times H, H)$ denote the space of continuous bilinear forms from $H \times H$ to H. The *adjoint* of an operator $A \colon H \to H$ is denoted by A^*. An operator $A \in \mathcal{L}(H)$ is *positive* if $\langle Ax, x \rangle_H \geq 0$ for all $x \in H$. We let $\mathcal{L}_2(H)$denote the space of *Hilbert-Schmidt operators* from H to H, this is a Hilbert space when endowed with the inner product

$$\langle A, B \rangle_{\mathcal{L}_2(H)} \coloneqq \operatorname{Tr}(B^* A) = \sum_{n=1}^{\infty} \langle A e_n, B e_n \rangle_H,$$

where Tr denotes the trace of an operator, $(e_n)_{n \in \mathbb{N}}$ is an orthonormal basis for H and $\langle \cdot, \cdot \rangle_{\mathcal{L}_2(H)}$ is independent of the choice of the orthonormal basis (see, e.g., [146, Section VI.6]).

A nonempty subset K of a vector space is called a *wedge* if $K + K \subseteq K$ and $\alpha K \subseteq K$ for all $\alpha \geq 0$, if moreover $K \cap (-K) = \{0\}$ then we call K a *cone*. A cone K in a vector space X induces a partial ordering: we write $x \leq_K y$ if $y - x \in K$ (and $x \geq_K y$ if $x - y \in K$). If $K \subset H$ is a wedge, we define the *dual* of K by

$$K^* = \{x \in H \colon \langle x, y\rangle_H \geq 0 \text{ for all } y \in K\}, \tag{2.1}$$

and we say that K is *self-dual* if $K = K^*$. Note that if K is self-dual then $0 \leq_K x \leq_K y$ implies $\|x\|_H^2 \leq \langle x, y\rangle_H \leq \|x\|_H \|y\|_H$, i.e. $0 \leq_K x \leq_K y$ implies $\|x\|_H \leq \|y\|_H$ (in other words, K is *monotonic*). We say that a cone K is *regular* if for all $y, x_1, x_2, \ldots \in K$ satisfying $x_1 \leq_K x_2 \leq_K \ldots \leq_K y$ there exists an $x \in H$ such that $\lim_{n\to\infty} \|x_n - x\|_H = 0$. A cone K is said to have *generating dual* if $B^* = K^* - K^*$. It is true that K has generating dual if and only if K is *normal*, i.e. $0 \leq_K x \leq_K y$ for $y \in K$, implies $\|x\| \leq \lambda \|y\|$ where $\lambda > 0$, see e.g. [91]. In finite dimensions, self-dual normal cones have non-empty interior. However, in infinite dimensions, the property $H = K - K$ does in general not imply that K has non-empty interior, see [100]. Let $(S, \mathcal{S})$ be a measurable space and $U \subseteq H$. A mapping $\mu \colon \mathcal{S} \to U$ is called a *U-valued measure* (on S) if it is weakly countably additive, i.e., if for every pairwise disjoint sequence $U_1, U_2 \ldots \in \mathcal{S}$ satisfying $\cup_{n\in\mathbb{N}} U_n = U$ it holds that $\langle \mu(U)x, y\rangle_H = \sum_{k\in\mathbb{N}} \langle \mu(U_k)x, y\rangle_H$ for all $x, y \in H$. We know from the work of Pettis [124] that if $\mu \colon \mathcal{F} \to H$ is weakly σ-additive, then it is also strongly σ-additive. For a H-valued measure μ and $h \in H$ we define the signed measure $\langle \mu, h\rangle \colon \mathcal{F} \to \mathbb{R}$ by $\langle \mu, h\rangle(A) = \langle \mu(A), h\rangle_H$, $A \in \mathcal{F}$. Throughout this thesis we are required to integrate with respect to vector-valued measures, for the readers convenience we added a section about it to Appendix A.1.

2.1.3 Setting

Throughout the first part of this thesis we assume that $(H, \langle \cdot, \cdot\rangle_H)$ is an infinite-dimensional, separable and real Hilbert space. For notational brevity we reserve $\langle \cdot, \cdot\rangle$ to denote the inner product on $\mathcal{L}_2(H)$, and $\| \cdot \|$ for the norm induced by $\langle \cdot, \cdot\rangle$. In addition, we define $\mathcal{H}$ to be the space of all self-adjoint Hilbert-Schmidt operators on H and $\mathcal{H}^+$ to be the cone of all positive operators in $\mathcal{H}$, i.e.

$$\mathcal{H} \coloneqq \{A \in \mathcal{L}_2(H) \colon A = A^*\} \quad \text{and}$$
$$\mathcal{H}^+ \coloneqq \{A \in \mathcal{H} \colon \langle Ah, h\rangle_H \geq 0 \text{ for all } h \in H\}.$$

Note that $\mathcal{H}$ is a closed subspace of $\mathcal{L}_2(H)$, and that $\mathcal{H}^+$ is a self-dual cone in $\mathcal{H}$ (indeed, $(\mathcal{H}^+)^* \subseteq \mathcal{H}^+$ by the spectral theorem for compact operators, and the reverse inclusion is trivial). Consequently, $\mathcal{H}$ is monotonic. Moreover, $\mathcal{H}^+$ is regular (see, e.g., [92, Theorem 1]), we have $\mathcal{H} = \mathcal{H}^+ - \mathcal{H}^+$ and we note that $\mathcal{H}^+$ has empty interior. We define the *truncation function* $\chi : \mathcal{H} \to \mathcal{H}$ by $\chi(\xi) \coloneqq \xi \mathbf{1}_{\{\|\xi\| \leq 1\}}$ and fix it throughout this chapter.

2.2 Existence of affine pure-jump processes

In this section we give a detailed definition of affine processes on the state space $\mathcal{H}^+$, introduce the notion of *admissible parameter sets* and compare our admissible parameter conditions to the matrix-valued case, which is done in Remark 2.4. Given an admissible parameter set we deduce first properties of the two functions in (1.5), i.e. the right-hand side functions of the generalized Riccati equations. At the end of this section we state our main result of this chapter in Theorem 2.8, which guarantees the existence of affine Markov processes on $\mathcal{H}^+$ associated with a given admissible parameter set and specifies the form of the weak generator on the linear span of the Fourier-basis elements. However, we relegate the proof to Section 2.4.3 and only give a brief outline of it at the end of this section.
We consider a time-homogeneous Markov process X with state space $\mathcal{H}^+$ and transition semigroup $(P_t)_{t\geq 0}$ acting on functions $f \in C_b(\mathcal{H}^+)$ as

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

where $p_t(x,\cdot)$, $t \geq 0$, $x \in \mathcal{H}^+$, denotes the transition kernel of X. Moreover for $x \in \mathcal{H}^+$, we denote the law of X given $X_0 = x$ by $\mathbb{P}_x$.

Definition 2.1. The Markov process $(X, (\mathbb{P}_x)_{x\in\mathcal{H}^+})$ is called *affine* if its Laplace transform has exponential-affine dependence on the initial state, i.e., if

$$P_t\, \mathrm{e}^{-\langle x,u\rangle} = \int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,\xi\rangle}\, p_t(x, \mathrm{d}\xi) = \mathrm{e}^{-\phi(t,u)-\langle x,\psi(t,u)\rangle}\,, \tag{2.2}$$

for all $t \geq 0$, and for all $u, x \in \mathcal{H}^+$, for some functions $\phi\colon \mathbb{R}_+ \times \mathcal{H}^+ \to \mathbb{R}_+$ and $\psi\colon \mathbb{R}_+ \times \mathcal{H}^+ \to \mathcal{H}^+$.

We follow the approach in [42] and consider the Laplace transform instead of the characteristic function which is justified by the non-negativity of X.
Note, that we do not require stochastic continuity of the affine process here, as in this chapter we are not aiming to provide a complete characterization of affine processes. As discussed in the introduction, our existence result requires an analysis of the corresponding generalized Riccati equations. In particular, a direct consequence of our approach (see Theorem 2.8 below) is that the processes we consider are *regular* in the sense of [42, Def. 2.2]. We recall this concept for the reader's convenience:

Definition 2.2. We call the affine process *regular*, whenever the functions

$$\frac{\partial\phi(t,u)}{\partial t}\Big|_{t=0+} \qquad \text{and} \qquad \frac{\partial\psi(t,u)}{\partial t}\Big|_{t=0+},$$

exist and are continuous at $u = 0$.

As we will see, the established class of affine processes satisfy an even stronger regularity condition, see Section 2.3.2. In finite dimensions stochastically continuous affine processes are always regular (see [97]), however, there exist finite-dimensional affine processes that are not stochastically continuous. Note that in infinite dimensions the regularity condition is somewhat more restrictive, as it implies, e.g., that the operator B in Definition 2.3 must be bounded. In this thesis we treat only the regular case and leave the unregular case to future work. In order to identify *pure-jump* affine processes, we introduce an *admissible parameter set* in the following definition.

Definition 2.3. 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) $\int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^2 \, m(\mathrm{d}\xi) < \infty$ and

   (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

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

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

$$M(x, \mathrm{d}\xi) := \frac{\langle x, \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2}, \tag{2.3}$$

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

$$\int_{\mathcal{H}^+\setminus\{0\}} \langle \chi(\xi), u\rangle M(x, \mathrm{d}\xi) < \infty; \tag{2.4}$$

iv) an operator $B \in \mathcal{L}(\mathcal{H})$ with adjoint B^* satisfying

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

for all $x, u \in \mathcal{H}^+$ with $\langle u, x\rangle = 0$.

We think of b as the **constant drift vector**, B the **linear term in the drift**, m the **constant jump measure**, and μ the **state-dependent jump measure**.

Remark 2.4 (Comparison to the finite-dimensional case)**.** Definition 2.3 above is analogous to the definition of an admissible parameter set for $\mathbb{R}^d_+$-valued processes see [52, Def. 2.6]) and the case of positive semi-definite and symmetric matrices, see [42, Def. 2.3]. However, as mentioned in the introduction, we do not consider any diffusion terms in this thesis. A more subtle difference is that we require second moment conditions on the measures $m(\mathrm{d}\xi)$ and $\|\xi\|^{-2}\mu(\mathrm{d}\xi)$, whereas no moment conditions are needed in the finite-dimensional setting. These second moment conditions are a consequence of our *generalized Feller approach*, for which we take the weight function $\rho = \|\cdot\|^2+1$. See Remark 2.39 for a detailed discussion regarding the necessity of these moment conditions to our approach.

In what follows we will frequently use the following observation:

$$\begin{aligned}&\forall \xi, u \in \mathcal{H}^+ :\\ -\min(\langle \xi, u\rangle, 1)\mathbf{1}_{\{\|\xi\|>1\}} &\leq \mathrm{e}^{-\langle \xi, u\rangle} -1 + \langle \chi(\xi), u\rangle \\ &\leq \tfrac{1}{2}|\langle \xi, u\rangle|^2\mathbf{1}_{\{\|\xi\|\leq 1\}} \leq \tfrac{1}{2}\|\xi\|^2\|u\|^2\mathbf{1}_{\{\|\xi\|\leq 1\}}.\end{aligned} \tag{2.5}$$

Given admissible parameters (b, B, m, μ), we recall $F\colon \mathcal{H}^+ \to \mathbb{R}$ and $R\colon \mathcal{H}^+ \to \mathcal{H}$ from (1.5), respectively, given by

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

$$R(u) = B^*(u) - \int_{\mathcal{H}^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, u\rangle} - 1 + \langle \chi(\xi), u\rangle\right) \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\,. \tag{2.6b}$$

Note that the admissibility conditions, Corollary A.7 and equation (2.5) ensure that F and R are well-defined. We also have that F and R are continuous and grow at most quadratically:

Lemma 2.5. *Let* (b, B, m, μ) *be an admissible parameter set conform Definition 2.3 and let* F *and* R *be as in* (2.6)*. Then* F *and* R *are continuous on* $\mathcal{H}^+$*.*

Proof. This follows immediately from (2.5) and Theorem A.8. □

Lemma 2.6. *Let* (b, B, m, μ) *be an admissible parameter set conform Definition 2.3 and let* F *and* R *be given by* (2.6)*. Then for all* $u \in \mathcal{H}^+$ *we have*

$$|F(u)| \leq \Big(\|b\| + \int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^2 m(\mathrm{d}\xi)\Big)(1 + \|u\|^2)\,, \tag{2.7a}$$

$$\|R(u)\| \leq \Big(\|B^*\|_{\mathcal{L}(\mathcal{H})} + \|\mu(\mathcal{H}^+ \setminus \{0\})\|\Big)(1 + \|u\|^2)\,. \tag{2.7b}$$

Proof. This follows immediately from the admissibility conditions, (2.5), (A.7), and (A.4). □

Inspired by the finite-dimensional theory, we consider a system of ordinary differential equations associated with the admissible parameter set (b, B, m, μ) given by the system (1.4). The equations are commonly known as the *generalized Riccati equations* which is due to the typically quadratic growth of F and R. To keep the chapter reasonably self-contained we recall the equations from (1.4). Using the formulas for F and R in (2.6) we define

$$\begin{cases} \dfrac{\partial\phi(t,u)}{\partial t} = F(\psi(t,u))\,, \quad t>0; \qquad \phi(0,u)=0\,, & \text{(2.8a)} \\ \dfrac{\partial\psi(t,u)}{\partial t} = R(\psi(t,u))\,, \quad t>0; \qquad \psi(0,u)=u\,. & \text{(2.8b)} \end{cases}$$

Definition 2.7. Let $u \in \mathcal{H}^+$. We say that $(\phi(\cdot,u),\psi(\cdot,u))\colon [0,\infty) \to \mathbb{R}\times\mathcal{H}$ is a *solution to* (2.8) if $(\phi(\cdot,u),\psi(\cdot,u))$ is continuously differentiable, takes values in $\mathbb{R}^+ \times \mathcal{H}^+$, and satisfies the equations (2.8a)-(2.8b).

For a transition semigroup $(P_t)_{t\geq 0}$ defined on bounded measurable functions on $\mathcal{H}^+$ we recall the notion of a *weak generator* $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ of $(P_t)_{t\geq 0}$ (see [123, Definition 9.36]), i.e. $f \in C_b(\mathcal{H}^+)$ belongs to $\mathrm{dom}(\mathcal{A})$, whenever the limit $\mathcal{A}f(x) := \lim_{t\to 0+} t^{-1}\big(P_t f(x) - f(x)\big)$, exists for every $x \in \mathcal{H}^+$; $\mathcal{A}f \in C_b(\mathcal{H}^+)$ and

$$P_t f(x) = f(x) + \int_0^t P_s \mathcal{G} f(x)\,\mathrm{d}s, \quad x \in \mathcal{H}^+.$$

The following theorem is our main result on the existence of affine pure-jump processes on the cone of positive self-adjoint Hilbert-Schmidt operators.

Theorem 2.8. *Let (b, B, m, μ) be an admissible parameter set as in Definition 2.3 and set $\nu(x,\mathrm{d}\xi) := m(\mathrm{d}\xi) + \|\xi\|^{-2}\langle \mu(\mathrm{d}\xi), x\rangle$. Then there exist constants $M, \omega \in [1,\infty)$ and a time-homogeneous $\mathcal{H}^+$-valued Markov process X with transition semigroup $(P_t)_{t\geq 0}$ such that*

$$\mathbb{E}\left[\|X_t\|^2 |\, X_0 = x\right] \leq M e^{\omega t}(\|x\|^2+1) \tag{2.9}$$

and

$$P_t\left(\mathrm{e}^{-\langle\cdot,u\rangle}\right)(x) = \mathrm{e}^{-\phi(t,u)-\langle x,\psi(t,u)\rangle}, \quad t\geq 0,\, u\in\mathcal{H}^+, \tag{2.10}$$

where $(\phi(\cdot,u),\psi(\cdot,u))$ is the unique solution to the associated generalized Riccati equations (2.8a)-(2.8b)*. Moreover, let $(\mathcal{G}, \mathrm{dom}(\mathcal{G}))$ be the weak generator of $(P_t)_{t\geq 0}$, then it holds that $\mathrm{lin}\left\{\mathrm{e}^{-\langle\cdot,u\rangle} : u\in\mathcal{H}^+\right\} \subseteq \mathrm{dom}(\mathcal{G})$ and for every function $f \in \mathrm{lin}\left\{\mathrm{e}^{-\langle\cdot,u\rangle} : u\in\mathcal{H}^+\right\}$ we have:*

$$\mathcal{G}f(x) = \langle b + B(x), f'(x)\rangle + \int_{\mathcal{H}^+\setminus\{0\}} \left(f(x+\xi) - f(x) - \langle \chi(\xi), f'(x)\rangle\right)\nu(x,\mathrm{d}\xi). \tag{2.11}$$

Outline of the proof. The proof is based on the approximation procedure that we conduct in detail in Section 2.4.2, where we work in the realm of generalized Feller semigroups, see Section 2.4.2 for more details. Here we limit ourselves to give a brief outline of the proof that shall provide a rough guidance for the upcoming sections and condensing the main ideas therein. The detailed proof is then given in Section 2.4.3.
Inspired by [45], we approximate the Kolmogorov type operator $\mathcal{G}$ in (2.11) by operators $(\mathcal{G}^{(k)})_{k\in\mathbb{N}}$ corresponding to processes of pure-jump type with finite activity, i.e. for every $k \in \mathbb{N}$ we replace the constant jump measure $m(\mathrm{d}\xi)$ in formula (2.11) by $\mathbf{1}_{\{\xi\geq 1/k\}}m(\mathrm{d}\xi)$ and the linear jump measures $\mu(\mathrm{d}\xi)$ by $\mathbf{1}_{\{\xi\geq 1/k\}}\mu(\mathrm{d}\xi)$. The approximation operators $\mathcal{G}^{(k)}$ generate strongly continuous semigroups $(P_t^{(k)})_{t\geq 0}$ on a space of functions, being weakly continuous with subquadratic growth, see Proposition 2.34.
Having established the existence of affine processes of pure-jump type associated with the strongly continuous semigroups $(P_t^{(k)})_{t\geq 0}$, we next apply a Trotter-Kato type result from [45] to obtain the limiting semigroup $(P_t)_{t\geq 0}$, see Proposition 2.37. To this end we first need to establish growth bounds on $(P_t^{(k)})_{t\geq 0}$, that are uniform in k, see Proposition 2.36. This requires understanding the associated generalized Riccati equations (2.8).
We provide global existence and uniqueness results in Section 2.3. The crucial importance of the associated generalized Riccati equations is that they substitute for the Kolmogorov equations, hence semigroup theoretic arguments involving the Kolmogorov-type operators, respectively the abstract Cauchy problem can be reduced to arguments from the theory of ordinary differential equations (ODEs). Lastly, we apply a version of Kolmogorov's extension theorem (see Theorem 2.26) to the limiting semigroup $(P_t)_{t\geq 0}$, which then yields the existence of an underlying Markovian process. This process associated via the semigroup to the operator $(\mathcal{G}, \mathrm{dom}(\mathcal{G}))$ is the desired affine process identified by the set (b, B, m, μ).

Remark 2.9. The equation for $\psi(\cdot, u)$ in the generalized Riccati equation (2.8b) is a non-linear differential equation on the cone of positive self-adjoint Hilbert-Schmidt operators. This type of infinite-dimensional differential equations has been of interest in the literature as they also show up e.g. in optimal control problems and stochastic filtering theory [46, 69, 111]. Hence, several articles deal with the problem of numerical tractability of this type of equations. See, e.g. [128] where Galerkin approximation and convergence theory was developed for operator-valued Riccati differential equations formulated in the space of Hilbert-Schmidt operators and [54] where the author studied a backward Euler approximation scheme and convergence results for this type of equations. In the subsequent Chapter 6, we investigate the Galerkin approximation further and draw a connection to matrix-valued affine processes.

2.3 Analysis of the generalized Riccati equations

In this section we investigate the generalized Riccati equations given by (2.8). More precisely, in Section 2.3.1 we introduce Lipschitz continuous approximations of the mappings R and F in (2.6) and use these approximations to show existence and uniqueness of a solution to (2.8). In Section 2.3.2 we establish regularity properties of R and F and use this to show that the solution map depends in a differentiable way on its initial value.

2.3.1 Solving the generalized Riccati equations (2.8)

The goal of this subsection is to prove the existence of a unique solution to the generalized Riccati equations given an admissible parameter set (b, B, m, μ). A common approach in the finite-dimensional case, e.g. in the case of the cone of positive semi-definite and symmetric matrices, is to use a localization argument exploiting the fact that the function R is analytic on the interior of the cone. Note, however, that in general R fails to be Lipschitz continuous on the boundary of the cone. The cone of positive self-adjoint Hilbert-Schmidt operators has an empty interior, a property that is shared by many cones in infinite dimensions. This has the consequence that localization arguments for solving equations (2.8) on the interior of $\mathbb{R}^+ \times \mathcal{H}^+$ are not valid anymore. Instead, for every $k \in \mathbb{N}$ we introduce approximations $F^{(k)}$ of F in equation (2.13a) and $R^{(k)}$ of R in equation (2.13b), which involve only finite-activity jump-measures, see (2.12) below. These approximations are Lipschitz continuous on $\mathcal{H}^+$ and in Proposition 2.16 we show that the solution to the generalized Riccati equations associated with $(b, B, m^{(k)}, \mu^{(k)})$ converges to the (unique) solution of (2.8).
We begin by introducing the approximating functions for F and R: For $k \in \mathbb{N}$ we set

$$m^{(k)}(\mathrm{d}\xi) \coloneqq \mathbf{1}_{\{\|\xi\|>1/k\}} m(\mathrm{d}\xi) \quad \text{and} \quad \mu^{(k)}(\mathrm{d}\xi) \coloneqq \mathbf{1}_{\{\|\xi\|>1/k\}} \mu(\mathrm{d}\xi), \tag{2.12}$$

and we introduce the functions $F^{(k)} \colon \mathcal{H}^+ \to \mathbb{R}$ and $R^{(k)} \colon \mathcal{H}^+ \to \mathcal{H}$ defined respectively as follows

$$F^{(k)}(u) = \langle b, u \rangle - \int_{\mathcal{H}^+ \setminus \{0\}} \big(\mathrm{e}^{-\langle \xi, u \rangle} - 1 + \langle \chi(\xi), u \rangle \big) m^{(k)}(\mathrm{d}\xi)\,, \tag{2.13a}$$

$$R^{(k)}(u) = B^*(u) - \int_{\mathcal{H}^+ \setminus \{0\}} \big(\mathrm{e}^{-\langle \xi, u \rangle} - 1 + \langle \chi(\xi), u \rangle \big) \frac{\mu^{(k)}(\mathrm{d}\xi)}{\|\xi\|^2}\,. \tag{2.13b}$$

We denote the generalized Riccati equations associated to $(b, B, m^{(k)}, \mu^{(k)})$ by:

$$\begin{cases} \dfrac{\partial \phi^{(k)}(t,u)}{\partial t} = F^{(k)}(\psi^{(k)}(t,u))\,, \quad t>0; \qquad \phi^{(k)}(0,u) = 0\,, & \text{(2.14a)} \\ \dfrac{\partial \psi^{(k)}(t,u)}{\partial t} = R^{(k)}(\psi^{(k)}(t,u))\,, \quad t>0; \qquad \psi^{(k)}(0,u) = u\,. & \text{(2.14b)} \end{cases}$$

The notion of quasi-monotonicity will be needed to guarantee that the solution to (2.14) stays in $\mathbb{R}^+ \times \mathcal{H}^+$.

Definition 2.10. Let $(V, \|\cdot\|_V)$ be a Hilbert space and let $K \subset V$ be a self-dual cone. In addition, let $D \subseteq V$ and let $f\colon D \to V$, then f is called *quasi-monotone with respect to* K if for all $v_1, v_2 \in D$ satisfying $v_1 \leq_K v_2$ and for all $u \in K$ satisfying $\langle v_2 - v_1, u\rangle = 0$ we have

$$\langle f(v_2) - f(v_1), u\rangle \geq 0.$$

Intuitively, quasi-monotone functions are pointing 'inwards' at the boundary points, which ensures that solutions stay in a cone (see Section A.2). For details on quasi-monotone functions on Banach spaces and their connection to differential equations see [50, Section 5.3].
The following lemma states that the admissibility of parameters implies that $R^{(k)}$, $k \in \mathbb{N}$, is quasi-monotone with respect to $\mathcal{H}^+$. The proof is analogous to the proof of [42, Lemma 5.1], we present an abridged version.

Lemma 2.11. *Let B and μ satisfy the conditions in Definition* 2.3 iii) *and Definition* 2.3 iv)*. Then for all $k \in \mathbb{N}$ the function $R^{(k)}$ given by* (2.13b) *is quasi-monotone with respect to $\mathcal{H}^+$.*

Proof. The admissibility condition Definition 2.3 iv) (which makes sense thanks to Definition 2.3 iii) and the monotonicity of the exponential function imply the quasi-monotonicity of $R^{(k)}$. □

By removing the small jumps and since m and μ have finite first moment, we obtain Lipschitz continuous mappings on $\mathcal{H}^+$:

Lemma 2.12. *Let B and μ satisfy the conditions in Definition* 2.3 iii) *and Definition* 2.3 iv)*. Let $k \in \mathbb{N}$ and $R^{(k)}$ given by* (2.13b)*. Then for all $u, v \in \mathcal{H}^+$ we have*

$$\|R^{(k)}(u) - R^{(k)}(v)\| \leq \left(\|B\|_{\mathcal{L}(\mathcal{H})} + 2k\|\mu(\mathcal{H}^+ \setminus \{0\})\|\right)\|u - v\|. \tag{2.15}$$

Proof. Observe that for all $u, v, \xi \in \mathcal{H}^+$ we have

$$\left|\mathrm{e}^{-\langle \xi, u\rangle} - \mathrm{e}^{-\langle \xi, v\rangle}\right| \leq \|\xi\|\|u - v\|.$$

Thus, (A.4) and (A.7) imply that

$$\begin{aligned}\|R^{(k)}(u) - R^{(k)}(v)\| &\leq \|B^*(u-v)\| + \left\| \int_{\mathcal{H}^+\setminus\{0\}\cap\{\frac{1}{k}<\|\xi\|<1\}} \langle \xi, u-v\rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \right\| \\ &\quad + \left\| \int_{\mathcal{H}^+\setminus\{0\}\cap\{\|\xi\|>\frac{1}{k}\}} (\mathrm{e}^{-\langle \xi,u\rangle} - \mathrm{e}^{-\langle \xi,v\rangle}) \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \right\| \\ &\leq \left(\|B\|_{\mathcal{L}(\mathcal{H})} + 2k\|\mu(\mathcal{H}^+\setminus\{0\})\|\right)\|u-v\|.\end{aligned}$$

□

Note that R is typically not Lipschitz continuous on the whole $\mathcal{H}^+$:

Remark 2.13. Note that

$$\left|e^{-\langle \xi,u\rangle} - e^{-\langle \xi,v\rangle} + \langle \xi, u-v\rangle\right| \leq \left|\int_{\langle \xi,u\rangle}^{\langle \xi,v\rangle} s \ \mathrm{d}s\right| \leq \|\xi\|^2(\|u\|\vee\|v\|)\|u-v\|, \quad (2.16)$$

for all $\xi, u, v \in \mathcal{H}^+$. This implies that R is in general Lipschitz continuous only on bounded sets in $\mathcal{H}^+$.

By Lemmas 2.11 and 2.12 we have that $R^{(k)}$ is Lipschitz continuous on $\mathcal{H}^+$ and quasi-monotone with respect to $\mathcal{H}^+$. Hence classical infinite dimensional ODE theory guarantees the existence of a global solution to the equations (2.14):

Proposition 2.14. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3 and let $R^{(k)}$, $k \in \mathbb{N}$, be given by equation* (2.13b)*. Then for every $k \in \mathbb{N}$ and $u \in \mathcal{H}^+$ there exists a solution $(\phi^{(k)}(\cdot,u), \psi^{(k)}(\cdot,u))$ to* (2.14)*. Moreover,*

$$\psi^{(k)}(t,u) \leq_{\mathcal{H}^+} \psi^{(k)}(t,v), \quad \forall u,v \in \mathcal{H}^+ \text{ satisfying } u \leq_{\mathcal{H}^+} v, \quad (2.17)$$

for all $t \geq 0$ and

$$\|\psi^{(k)}(t,u) - \psi^{(k)}(t,v)\| \leq \exp\left(\left(\|B\|_{\mathcal{L}(\mathcal{H})} + 2k\|\mu(\mathcal{H}^+\setminus\{0\})\|\right)t\right)\|u-v\|, \quad (2.18)$$

for all $t \geq 0$ and $u, v \in \mathcal{H}^+$.

Proof. Let $k \in \mathbb{N}$. By Lemma 2.12 the function $R^{(k)}$ is Lipschitz continuous on $\mathcal{H}^+$, by (2.15) with $v = 0$ the function $R^{(k)}$ satisfies the linear growth condition $\|R^{(k)}(u)\| \leq \left(\|B\|_{\mathcal{L}(\mathcal{H})} + 2k\|\mu(\mathcal{H}^+\setminus\{0\})\|\right)\|u\|$ and by Lemma 2.11 $R^{(k)}$ is quasi-monotone with respect to $\mathcal{H}^+$, thus by [113, VI.3. Theorem 3.1 and Proposition 3.2] there exists a unique global solution $\psi^{(k)}(\cdot,u)\colon [0,\infty) \to \mathcal{H}^+$ to the second equation of (2.14). Now, setting $\phi^{(k)}(t,u) = \int_0^t F^{(k)}(\psi^{(k)}(s,u))\,\mathrm{d}s$, for all $t \geq 0$, we obtain by continuity of $F^{(k)}$ and $\psi^{(k)}(\cdot,u)$ a solution of (2.14), denoted by $(\phi^{(k)}(\cdot,u), \psi^{(k)}(\cdot,u))$, satisfying the inequality (2.17).

Finally, observe that Lemma 2.12 implies that

$$\begin{aligned}
&\frac{\partial}{\partial t}\|\psi^{(k)}(t,u)-\psi^{(k)}(t,v)\|^2 \\
&= 2\left\langle \psi^{(k)}(t,u)-\psi^{(k)}(t,v), R^{(k)}(\psi^{(k)}(t,u)) - R^{(k)}(\psi^{(k)}(t,v))\right\rangle \\
&\leq 2\left(\|B\|_{\mathcal{L}(\mathcal{H})} + 2k\|\mu(\mathcal{H}^+\setminus\{0\})\|\right)\|\psi^{(k)}(t,u)-\psi^{(k)}(t,v)\|^2.
\end{aligned}$$

This and Gronwall's lemma implies the second inequality (2.18). □

Proposition (2.16) below guarantees the existence of a unique solution to the original generalized Riccati equations (2.8) on the interval $[0,\infty)$. First, we prove the following lemma:

Lemma 2.15. *Let B and μ satisfy the conditions in Definition* 2.3 iii) *and Definition* 2.3 iv)*, let $R^{(k)}$ and R be respectively given by equation* (2.13b) *and* (2.6b)*. Then for every $M>0$ we have*

$$\lim_{k\to\infty}\sup_{u\in\mathcal{H}^+:\|u\|\leq M}\|R^{(k)}(u)-R(u)\| = 0\,.$$

Proof. It follows immediately from (A.7) and (2.5) that

$$\|R^{(k)}(u)-R(u)\| \leq \|\mu(\{\xi\in\mathcal{H}^+ : \|\xi\|\leq \tfrac{1}{k}\})\|\|u\|^2. \tag{2.19}$$

The assertion follows from the above and the continuity of μ, see (A.2). □

Proposition 2.16. *Let (b,B,m,μ) be an admissible parameter set conform Definition 2.3. Then for every $u\in\mathcal{H}^+$ there exists a unique solution $(\phi(\cdot,u),\psi(\cdot,u))$ of the equations* (2.8)*. Moreover,*

$$\psi(t,u)\leq_{\mathcal{H}^+}\psi^{(k)}(t,u) \quad \forall\, k\in\mathbb{N},\ t\geq 0 \text{ and } u\in\mathcal{H}^+,$$

and $\psi(t,u)=\lim_{k\to\infty}\psi^{(k)}(t,u)$ for all $t\geq 0$ and $u\in\mathcal{H}^+$, as well as

$$\psi(t,u)\leq_{\mathcal{H}^+}\psi(t,v), \quad \forall\, t\geq 0 \text{ and } u,v\in\mathcal{H}^+ \text{ with } u\leq_{\mathcal{H}^+} v, \tag{2.20}$$

and

$$\|\psi(t,u)\|\leq \exp\left(\left(\|B\|_{\mathcal{L}(\mathcal{H})}+2\|\mu(\mathcal{H}^+\setminus\{0\})\|\right)t\right)\|u\|, \quad \forall\, t\geq 0,\ u\in\mathcal{H}^+. \tag{2.21}$$

Finally, for all $M,T\geq 0$ there exists a $K(M,T)\geq 0$ such that for all $u,v\in\mathcal{H}^+$ satisfying $\|u\|,\|v\|\leq M$ and all $t\in[0,T]$ it holds that

$$\|\psi(t,u)-\psi(t,v)\|\leq K(M,T)\|u-v\|. \tag{2.22}$$

Proof. First of all note that uniqueness of a solution follows from the fact that R is Lipschitz continuous on bounded sets of $\mathcal{H}^+$, see Remark 2.13. Observe that by (A.5), (2.5), and (2.13b) we have, for all $u \in \mathcal{H}^+$ and $k \in \mathbb{N}$,

$$\begin{aligned} R^{(k)}(u) - R^{(k+1)}(u) &= \int_{\mathcal{H}^+ \cap \{\frac{1}{k+1} < \|\xi\| \leq \frac{1}{k}\}} \left(\mathrm{e}^{-\langle u, \xi \rangle} - 1 + \langle \xi, u \rangle \right) \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \\ &\geq_{\mathcal{H}^+} 0. \end{aligned} \tag{2.23}$$

Now, fix $u \in \mathcal{H}^+$. By Proposition 2.14 we know that there exists a unique global solution $\psi^{(k)}(\cdot, u)$ to equation (2.14) for every $k \in \mathbb{N}$. This combined with (2.23) implies that for all $k \in \mathbb{N}$ and $t \geq 0$ we have

$$\begin{aligned} \frac{\partial \psi^{(k+1)}}{\partial t}(t,u) - R^{(k+1)}(\psi^{(k+1)}(t,u)) &= \frac{\partial \psi^{(k)}}{\partial t}(t,u) - R^{(k)}(\psi^{(k)}(t,u)) \\ &\leq_{\mathcal{H}^+} \frac{\partial \psi^{(k)}}{\partial t}(t,u) - R^{(k+1)}(\psi^{(k)}(t,u)). \end{aligned}$$

It follows from Lemma 2.12 and Theorem A.9 with $K = \mathcal{H}^+$, $F = R^{(k+1)}$, $f = \psi^{(k+1)}(\cdot, u)$ and $g = \psi^{(k)}(\cdot, u)$ that

$$\psi^{(k+1)}(t,u) \leq_{\mathcal{H}^+} \psi^{(k)}(t,u)\,, \quad t \geq 0. \tag{2.24}$$

As moreover $\psi^{(k)}(t,u) \geq_{\mathcal{H}^+} 0$ for all $t \geq 0$ and $k \in \mathbb{N}$, the regularity of the cone $\mathcal{H}^+$ implies that for all $t \geq 0$ there exists a $\psi(t,u) \in \mathcal{H}^+$ such that

$$\psi(t,u) = \lim_{k \to \infty} \psi^{(k)}(t,u). \tag{2.25}$$

Note that by (2.24), the monotonicity of $\mathcal{H}^+$, and the continuity of $\psi^{(1)}(\cdot, u)$ we have, for all $T > 0$,

$$\sup_{k \in \mathbb{N}, s \in [0,T]} \|\psi^{(k)}(s,u)\| \leq \sup_{s \in [0,T]} \|\psi^{(1)}(s,u)\| < \infty\,. \tag{2.26}$$

It follows from this, (2.25), the dominated convergence theorem, and Lemmas 2.15 and 2.6 that for all $t \geq 0$ we have

$$\begin{aligned} \psi(t,u) &= \lim_{k \to \infty} \psi^{(k)}(t,u) \\ &= u + \lim_{k \to \infty} \int_0^t R^{(k)}(\psi^{(k)}(s,u))\,\mathrm{d}s \\ &= u + \lim_{k \to \infty} \int_0^t \Big(R^{(k)}(\psi^{(k)}(s,u)) - R(\psi^{(k)}(s,u)) \Big)\,\mathrm{d}s \\ &\quad + \lim_{k \to \infty} \int_0^t R(\psi^{(k)}(s,u))\,\mathrm{d}s \\ &= u + \int_0^t R(\psi(s,u))\,\mathrm{d}s. \end{aligned}$$

The equation above combined with Lemma 2.6 implies that the map $\psi(\cdot, u)$ is continuous, whence Lemma 2.5 and the fundamental theorem of calculus imply that $\psi(\cdot, u) \in C^1([0, \infty), \mathcal{H})$ and

$$\frac{\partial \psi}{\partial t}(t, u) = R(\psi(t, u)), \quad t \geq 0; \qquad \psi(0, u) = u. \tag{2.27}$$

Moreover, the continuity of F and of $\psi(\cdot, u)$ ensures that by setting

$$\phi(t, u) = \int_0^t F(\psi(s, u))\,\mathrm{d}s, \quad t \geq 0, \tag{2.28}$$

we obtain that $(\phi(\cdot, u), \psi(\cdot, u))$ is a solution to (2.8a)-(2.8b).
Next, note that (2.20) follows from (2.17) and (2.25). Moreover, (2.21) follows from (2.18) with $k = 1$, (2.24), (2.25), and the fact that $\psi^{(1)}(t, 0) \equiv 0$. Finally, (2.22) follows from the Lipschitz continuity of R on bounded sets (see Remark 2.13), (2.21), and the same reasoning as we used to obtain (2.18). □

2.3.2 Regularity with respect to the initial value of the solution

Having established the existence of a unique solution to (2.8), we now turn to the regularity of the solution with respect to its initial value. To this end we first must introduce a fitting concept of differentiability:

Definition 2.17. Let X and Y be Banach spaces and $D \subseteq X$ a convex subset. We say that a function $f\colon D \subseteq X \to Y$ has a *one-sided derivative* at $x \in D$ in the direction $v \in X$, whenever $x + \lambda v \in D$ for all λ sufficiently small and the limit

$$\lim_{\lambda \to 0+} \frac{f(x + \lambda v) - f(x)}{\lambda},$$

exists in Y. We denote this limit by $\mathrm{d}_+ f(x)(v)$. We define the second one-sided derivative in $x \in D$ in direction $(v, w) \in X \times X$ as

$$\lim_{\lambda \to 0+} \frac{\mathrm{d}_+ f(x + \lambda w)(v) - \mathrm{d}_+ f(x)(v)}{\lambda},$$

whenever $x + \lambda w \in D$ and $\mathrm{d}_+ f(x + \lambda w)(v)$ exists for all λ sufficiently small and moreover the limit exists in Y. We denote the second one-sided derivative of f at x in directions (v, w) by $\mathrm{d}_+^2 f(x)(v, w)$.

Lemma 2.18. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3 and let F and R be given by* (2.6a)-(2.6b). *For $u \in \mathcal{H}^+$ define $\mathrm{d}R(u) \in \mathcal{L}(\mathcal{H})$ by*

$$\mathrm{d}R(u)v = B^*(v) + \int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, v\rangle \,\mathrm{e}^{-\langle \xi, u\rangle} - \langle \chi(\xi), v\rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}, \quad v \in \mathcal{H}, \tag{2.29}$$

and $\mathrm{d}F(u) \in \mathcal{L}(H, \mathbb{R})$ by

$$\mathrm{d}F(u)v = \langle b, v\rangle + \int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, v\rangle \,\mathrm{e}^{-\langle \xi, u\rangle} - \langle \chi(\xi), v\rangle m(\mathrm{d}\xi), \quad v \in \mathcal{H}. \tag{2.30}$$

Moreover, define $\mathrm{d}^2R(u) \in \mathcal{L}^{(2)}(\mathcal{H} \times \mathcal{H}, \mathcal{H})$ by

$$\mathrm{d}^2R(u)(v, w) = -\int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, v\rangle\langle \xi, w\rangle \,\mathrm{e}^{-\langle \xi, u\rangle} \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}, \quad v, w \in \mathcal{H}, \tag{2.31}$$

and $\mathrm{d}^2F(u) \in \mathcal{L}^{(2)}(\mathcal{H} \times \mathcal{H}, \mathbb{R})$ by

$$\mathrm{d}^2F(u)(v, w) = -\int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, v\rangle\langle \xi, w\rangle \,\mathrm{e}^{-\langle \xi, u\rangle} m(\mathrm{d}\xi), \quad v, w \in \mathcal{H}. \tag{2.32}$$

Then the operator $\mathrm{d}R(u)$ is quasi-monotone for all $u \in \mathcal{H}^+$, and for all $u_0, u_1 \in \mathcal{H}^+$ and $v, w \in \mathcal{H}$ we have

$$\|\,\mathrm{d}R(u_0)(v)\| \leq \|B^*\|_{\mathcal{L}(\mathcal{H})}\|v\| + \left\|\mu(\mathcal{H}^+ \setminus \{0\})\right\| (1 + \|u_0\|)\|v\|, \tag{2.33}$$

$$\|\,\mathrm{d}R(u_0)(v) - \mathrm{d}R(u_1)(v)\| \leq \left\|\mu(\mathcal{H}^+ \setminus \{0\})\right\| \|u_0 - u_1\|\|v\|, \tag{2.34}$$

$$\|\,\mathrm{d}^2R(u_0)(v, w)\| \leq \left\|\mu(\mathcal{H}^+ \setminus \{0\})\right\| \|v\|\|w\|, \tag{2.35}$$

and $u \mapsto \mathrm{d}^2R(u)(v, w)$ is continuous. Moreover, F and R are two-times one-sided differentiable in u in the direction (v, w) for all $u, v, w \in \mathcal{H}^+$, and for all $u, v, w \in \mathcal{H}^+$ we have:

$$\mathrm{d}_+R(u)(v) = \mathrm{d}R(u)v, \tag{2.36}$$

$$\mathrm{d}^2_+R(u)(v, w) = \mathrm{d}^2R(u)(v, w), \tag{2.37}$$

$$\mathrm{d}_+F(u)(v) = \mathrm{d}F(u)v, \tag{2.38}$$

$$\mathrm{d}^2_+F(u)(v, w) = \mathrm{d}F(u)(v, w). \tag{2.39}$$

Proof. The quasi-monotonicity of $\mathrm{d}R$ follows directly from the admissibility assumption. For all $u, \xi \in \mathcal{H}^+$ and all $v \in \mathcal{H}$ we have

$$\left|\langle \xi, v\rangle \,\mathrm{e}^{-\langle \xi, u\rangle} - \langle \chi(\xi), v\rangle\right| \leq \|\xi\|\|v\|(\mathbf{1}_{\{\|\xi\|>1\}} + \|\xi\|\|u\|\mathbf{1}_{\{\|\xi\|\leq 1\}}),$$

which together with (2.29) yields (2.33).

Estimate (2.34) is obtained similarly, estimate (2.35) is immediate from the definition, and the continuity of $u \mapsto \mathrm{d}^2 R(u)(v,w)$ follows from the dominated convergence theorem (Theorem A.8).
We next confirm the asserted differentiability of the map $u \mapsto R(u)$. Let $u, v \in \mathcal{H}^+$ then

$$\begin{aligned}
\mathrm{d}_+ R(u)(v) &= \lim_{\lambda \to 0+} \frac{R(u+\lambda v) - R(u)}{\lambda} \\
&= B^*(v) - \lim_{\lambda \to 0+} \int_{\mathcal{H}^+ \setminus \{0\}} \frac{\mathrm{e}^{-\langle \xi, u+\lambda v\rangle} - \mathrm{e}^{-\langle \xi, u\rangle}}{\lambda} + \langle \chi(\xi), v\rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \\
&= B^*(v) - \int_{\mathcal{H}^+ \setminus \{0\}} \lim_{\lambda \to 0+} \frac{\mathrm{e}^{-\langle \xi, u+\lambda v\rangle} - \mathrm{e}^{-\langle \xi, u\rangle}}{\lambda} + \langle \chi(\xi), v\rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \qquad (2.40) \\
&= B^*(v) + \int_{\mathcal{H}^+ \setminus \{0\}} \langle \xi, v\rangle\, \mathrm{e}^{-\langle \xi, u\rangle} - \langle \chi(\xi), v\rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}.
\end{aligned}$$

where the interchange of the integral and the limit in equation (2.40) is justified, since $\lambda \mapsto \mathrm{e}^{-\langle u+\lambda v, \xi\rangle}$ is a convex mapping, hence its differential quotient is non-decreasing in λ and non-negative, and thus we can apply the monotone convergence theorem to obtain that the one-sided derivative of R exists in u in the direction v and (2.36) holds. An analogous derivation for F leads to equation (2.38).
The proof that the *second* one-sided directional derivative of both F and R exist and that (2.37)–(2.39) hold is again analogous. Note in particular that for the existence of the second derivatives we use that the measures $m(\mathrm{d}\xi)$ and $\|\xi\|^{-2}\mu(\mathrm{d}\xi)$ have finite second moments. ☐

In Proposition 2.20 below we show that the solution $(\phi(\cdot,u), \psi(\cdot,u))$ of (2.8) is such that the mappings $u \mapsto \psi(t,u)$ and $u \mapsto \phi(t,u)$ are twice one-sided differentiable in the origin in all directions. The techniques to prove this might be well-known, however, as we are dealing with a non-standard concept of differentiability we still give the detailed proof. We first need the following lemma which is a consequence of the fundamental theorem of calculus:

Lemma 2.19. *Let X, Y be Banach spaces, let $F\colon D \subset X \to Y$, let $x, y \in D$ and assume that the one-sided derivative of F in z exists in the direction $y-x$ for all $z \in \{x + s(y-x) \colon s \in [0,1]\}$ and that the mapping*

$$[0,1] \ni s \mapsto \mathrm{d}_+ F(x + s(y-x))(y-x) \in Y \qquad (2.41)$$

is continuous. Then $F(y) - F(x) = \int_0^1 \mathrm{d}_+ F(x+s(y-x))(y-x)\,\mathrm{d}s$.

Proof. The continuity of $[0,1] \ni s \mapsto \mathrm{d}_+F(x+s(y-x))(y-x) \in Y$ and the fundamental theorem of calculus imply that the right derivative of the mapping

$$[0,1] \ni t \mapsto \left(F(x+t(y-x)) - F(x) - \int_0^t \mathrm{d}_+F(x+s(y-x))(y-x)\,\mathrm{d}s \right) \in Y$$

equals zero. As any function with right derivative equal to zero is constant, this leads to the desired assertion. □

Proposition 2.20. *Let (b,B,m,μ) be an admissible parameter set conform Definition 2.3, for every $u \in \mathcal{H}^+$ let $(\phi(\cdot,u),\psi(\cdot,u))$ be the solution to* (2.8a)-(2.8b)*, and let $\mathrm{d}R$, $\mathrm{d}F$, d^2R, and d^2F be defined by* (2.29)–(2.32)*. Then the maps $u \mapsto \psi(t,u)$ and $u \mapsto \phi(t,u)$ are twice one-sided differentiable in 0 in all directions $(v,w) \in \mathcal{H}^+ \times \mathcal{H}^+$. Moreover, $\mathrm{d}_+\psi(t,0)(v), \mathrm{d}_+^2\psi(t,0)(v,w) \in \mathcal{H}^+$ for all $v,w \in \mathcal{H}^+$ and the mappings $t \mapsto \mathrm{d}_+\phi(t,0)(v)$ and $t \mapsto \mathrm{d}_+\psi(t,0)(v)$ solves the following pair of differential equations:*

$$\frac{\partial}{\partial t}\mathrm{d}_+\phi(t,0)(v) = \mathrm{d}F(0)\big(\mathrm{d}_+\psi(t,0)(v)\big), \quad t \geq 0; \quad \mathrm{d}_+\phi(0,0)(v) = 0, \tag{2.42a}$$

$$\frac{\partial}{\partial t}\mathrm{d}_+\psi(t,0)(v) = \mathrm{d}R(0)\big(\mathrm{d}_+\psi(t,0)(v)\big), \quad t \geq 0; \quad \mathrm{d}_+\psi(0,0)(v) = v. \tag{2.42b}$$

Moreover, the mappings $t \mapsto \mathrm{d}_+^2\psi(t,0)(v,w)$ and $t \mapsto \mathrm{d}_+^2\phi(t,0)(v,w)$ solve the following pair of differential equations:

$$\begin{aligned}\frac{\partial}{\partial t}\mathrm{d}_+^2\phi(t,0)(v,w) &= \mathrm{d}^2F(0)(\mathrm{d}_+\psi(t,0)(v),\mathrm{d}_+\psi(t,0)(w)) \\ &\quad + \mathrm{d}F(0)\big(\mathrm{d}_+^2\psi(t,0)(v,w)\big), \quad t \geq 0; \quad \mathrm{d}_+^2\phi(0,0)(v,w) = 0,\end{aligned} \tag{2.43a}$$

$$\begin{aligned}\frac{\partial}{\partial t}\mathrm{d}_+^2\psi(t,0)(v,w) &= \mathrm{d}^2R(0)\big(\mathrm{d}_+\psi(t,0)(v),\mathrm{d}_+\psi(t,0)(w)\big) \\ &\quad + \mathrm{d}R(0)(\mathrm{d}_+^2\psi(t,0)(v,w)), \quad t \geq 0; \quad \mathrm{d}_+^2\psi(0,0)(v,w) = 0.\end{aligned} \tag{2.43b}$$

Proof. Note that in order to prove that the *second* directional derivative in 0 of a mapping exists, we need that its *first* directional derivative exists in $u \in \mathcal{H}^+$ for all $u \in \mathcal{H}^+$ sufficiently small. Hence, we begin by proving that the first derivative of $u \mapsto \psi(t,u)$ exists in u in the direction v for all $u,v \in \mathcal{H}^+$ and all $t \in [0,\infty)$. To this end we fix $u,v \in \mathcal{H}^+$.
Recall the definition of the operators $\mathrm{d}R(u) \in \mathcal{L}(\mathcal{H})$ and $\mathrm{d}^2R(u) \in \mathcal{L}^{(2)}(\mathcal{H}\times\mathcal{H},\mathcal{H})$ from (2.29) and (2.31). Define the operator $C_\theta(t) \in \mathcal{L}(\mathcal{H})$, $\theta, t \in [0,\infty)$, by

$$C_\theta(t)w = \int_0^1 \mathrm{d}R\left(\psi(t,u) + s(\psi(t,u+\theta v) - \psi(t,u))\right) w \,\mathrm{d}s. \tag{2.44}$$

Note that the integral in (2.44) is well-defined as the integrand is continuous in s by (2.34) and bounded by (2.21) and (2.33). Lemma 2.19, equation (2.36), the fact that $(1-s)\psi(t,u)+s\psi(t,u+\theta v)\in\mathcal{H}^+$ for all $s\in[0,1]$, $t\in[0,\infty)$, and the fact that $\psi(t,u+\theta v)\geq_{\mathcal{H}^+}\psi(t,u)$ for all $t\in[0,\infty)$ by (2.20) imply that

$$C_\theta(t)(\psi(t,u+\theta v)-\psi(t,u)) = R(\psi(t,u+\theta v))-R(\psi(t,u)),\quad \theta,t\in[0,\infty).$$

This and equation (2.8b) imply

$$\frac{\partial}{\partial t}(\psi(t,u+\theta v)-\psi(t,u)) = C_\theta(t)(\psi(t,u+\theta v)-\psi(t,u)),\quad \theta,t\in[0,\infty).$$

It follows that

$$\psi(t,u+\theta v)-\psi(t,u)=\theta\exp\left(\int_0^t C_\theta(s)\,\mathrm{d}s\right)v,\quad \theta,t\in[0,\infty),$$

where we note that $\int_0^t C_\theta(s)\,\mathrm{d}s$ is well-defined in $\mathcal{L}(\mathcal{H})$ as the $\mathcal{L}(\mathcal{H})$-valued integrand is continuous in s by (2.34) and bounded due to (2.33). This implies that for all $\theta\in(0,\infty)$ we have

$$\begin{aligned}&\left\|\frac{\psi(t,u+\theta v)-\psi(t,u)}{\theta}-\exp\left(\int_0^t C_0(s)\,\mathrm{d}s\right)v\right\|\\&=\left\|\left(\exp\left(\int_0^t C_\theta(s)\,\mathrm{d}s\right)-\exp\left(\int_0^t C_0(s)\,\mathrm{d}s\right)\right)v\right\|.\end{aligned}\tag{2.45}$$

Using the identity $\|\,\mathrm{e}^A-\mathrm{e}^B\|_{\mathcal{L}(\mathcal{H})}\leq\|A-B\|_{\mathcal{L}(\mathcal{H})}\,\mathrm{e}^{\|A\|_{\mathcal{L}(\mathcal{H})}\vee\|B\|_{\mathcal{L}(\mathcal{H})}}$, $A,B\in\mathcal{L}(\mathcal{H})$, we obtain from (2.44), (2.45), (2.21), (2.22), and (2.34) that the one-sided derivative $\mathrm{d}_+\psi(t,u)(v)$ exists. Moreover, the fact that $C_0(t)v=\mathrm{d}R(\psi(t,u))v$ implies that $t\mapsto\mathrm{d}_+\psi(t,u)(v)$ is the solution to the following ODE

$$\frac{\partial}{\partial t}\,\mathrm{d}_+\psi(t,u)(v)=\mathrm{d}R(\psi(t,u))\big(\mathrm{d}_+\psi(t,u)(v)\big),\quad t\geq 0;\quad \mathrm{d}_+\psi(0,u)(v)=v.\tag{2.46}$$

This, together with the quasi-monotonicity of $\mathrm{d}R(\psi(t,u))$ (see Lemma 2.18) and Theorem A.9 implies that $\mathrm{d}_+\psi(t,u)(v)\in\mathcal{H}^+$. Regarding the derivative of ϕ, note that estimates analogous to (2.33) and (2.34) hold for $\mathrm{d}F$, which, in combination with the fact that $\mathrm{d}_+\psi(t,u)(v)\in\mathcal{H}^+$, (2.8), (2.38), and Lemma 2.19 implies that for all $\theta\in(0,\infty)$ and all $t\in[0,\infty)$ we have

$$\begin{aligned}&\frac{\phi(t,u+\theta v)-\phi(t,u)}{\theta}\\&=\int_0^t\int_0^1\mathrm{d}F(\psi(s,u)+r(\psi(s,u+\theta v)-\psi(s,u)))\,\mathrm{d}r\,\frac{\psi(s,u+\theta v)-\psi(s,u)}{\theta}\,\mathrm{d}s.\end{aligned}$$

This in combination with (2.22) and (2.20) implies that the dominated convergence theorem can be applied to obtain that $\mathrm{d}_+\phi(t,u)$ exists for all t and satisfies

$$\frac{\partial}{\partial t}\,\mathrm{d}_+\phi(t,u)(v) = \mathrm{d}F(\psi(t,u))\big(\,\mathrm{d}_+\psi(t,u)(v)\big), \quad t \geq 0; \quad \mathrm{d}_+\phi(0,u)(v) = 0. \tag{2.47}$$

This proves in particular that $u \mapsto (\phi(t,u), \psi(t,u))$ is differentiable in 0 in the direction $v \in \mathcal{H}^+$ for all $v \in \mathcal{H}^+$ and that the corresponding derivatives solve the ODEs (2.42a) and (2.42b).
We now turn to the second derivative in 0. To this end, fix $v, w \in \mathcal{H}^+$ and observe that Lemma 2.19, the boundedness and continuity of d^2R, see Lemma 2.18, together with (2.37) and the fact that $\psi(t,\theta v), \mathrm{d}_+\psi(t,\theta v) \in \mathcal{H}^+$ for all $\theta \in [0,\infty)$ imply that

$$\begin{aligned}\frac{\partial}{\partial t}\big((\mathrm{d}_+\psi(t,\theta v) - \mathrm{d}_+\psi(t,0))(w)\big) &= \int_0^1 \mathrm{d}^2R(s\psi(t,\theta v))(\mathrm{d}_+\psi(t,\theta v)(w), \psi(t,\theta v))\,\mathrm{d}s \\ &\quad + \mathrm{d}R(0)\,(\mathrm{d}_+\psi(t,\theta v)(w) - \mathrm{d}_+\psi(t,0)(w))\end{aligned}$$

for all $\theta \in [0,\infty), t \in [0,\infty)$. As $\mathrm{d}_+\psi(0,\theta v)(w) - \mathrm{d}_+\psi(0,0)(w) = 0$ this implies

$$\begin{aligned}&\frac{\mathrm{d}_+\psi(t,\theta v)(w) - \mathrm{d}_+\psi(t,0)(w)}{\theta} \\ &= \int_0^t \mathrm{e}^{(t-r)\,\mathrm{d}R(0)} \int_0^1 \mathrm{d}^2R(s\psi(r,\theta v))\left(\mathrm{d}_+\psi(r,\theta v)(w), \frac{\psi(r,\theta v)}{\theta}\right)\mathrm{d}s\ \mathrm{d}r\end{aligned} \tag{2.48}$$

for all $\theta \in (0,\infty), t \in [0,\infty)$. Note that (2.22), (2.34), and (2.46) imply that $\lim_{\theta\to 0+} \mathrm{d}_+\psi(t,\theta v)(w) = \mathrm{d}_+\psi(t,0)(w)$. Moreover, we have already established that $\lim_{\theta\to 0+} \frac{\psi(t,\theta v)}{\theta} = \mathrm{d}_+\psi(t,0)(v)$. Combining these observations with equations (2.21), (2.35), and (2.48) implies that $\mathrm{d}_+^2\psi(t,0)(v,w)$ exists and satisfies equation (2.43a). We leave it to the reader to now verify that also $\mathrm{d}_+^2\phi(t,u)(v,w)$ exists and that $\mathrm{d}_+^2\phi(t,u)(v,w)$ satisfies (2.43b). □

For $u = 0$ we derive explicit formulas for the solutions to the pairs of differential equations in (2.42b) and (2.43b) of Proposition 2.20, as those will be needed for proving Lemma 2.35 in the approximating case and for Proposition 2.38 below.
First, note that

$$\mathrm{d}_+R(0)(v) = B^*(v) + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \langle \xi, v\rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}.$$

Recall the definition of $\mathrm{d}R(0)$ from (2.29). The solution of equation (2.42b) is then given by

$$\mathrm{d}_+\psi(t,0)(v) = \mathrm{e}^{t\,\mathrm{d}R(0)}v. \tag{2.49}$$

By inserting formula (2.49) into equation (2.43b) (note that $e^{t\,\mathrm{d}R(0)}v \in \mathcal{H}^+$) and solving this inhomogeneous linear equation we obtain

$$\mathrm{d}_+^2\psi(t,0)(v,w) = \int_0^t \mathrm{e}^{(t-s)\,\mathrm{d}R(0)}\,\mathrm{d}^2R(0)(\mathrm{e}^{s\,\mathrm{d}R(0)}v, \mathrm{e}^{s\,\mathrm{d}R(0)}w)\,\mathrm{d}s. \tag{2.50}$$

Remark 2.21. Note that one can also prove that $u \mapsto \psi(t,u)$ and $u \mapsto \phi(t,u)$ are twice one-sided differentiable in u for every $u \in \mathcal{H}^+$, in every direction $(v,w) \in \mathcal{H}^+ \times \mathcal{H}^+$. We do not need this, but what we *needed* in the proof of Proposition 2.20 was the existence of the first derivative in $u \in \mathcal{H}^+$ for sufficiently small u in order to obtain the second derivative.

2.4 Proof: Existence of affine pure-jump processes

In this section we use the well-posedness and regularity results of the generalized Riccati equations (2.8) from Section 2.3 to show the existence of an affine process in $\mathcal{H}^+$ associated to a given admissible parameter set (b, B, m, μ) conform Definition 2.3. Due to the lack of local compactness of the underlying state space, standard Feller theory cannot be employed in our context and we use the theory of generalized Feller processes as introduced in [51]. The existence proof is based on the approximation procedure roughly sketched at the end of Section 2.2. In this section we rigorously build up this approximation procedure in the *generalized Feller* setting. Essentially, we approximate the transition semigroup $(P_t)_{t\geq 0}$, that can be associated to an affine process on $\mathcal{H}^+$ with infinite-activity jump behavior, by simpler transition semigroups corresponding to affine finite-activity jump processes. The considered semigroups are strongly continuous semigroups on a certain Banach space of real functions being weakly-continuous on compact sets and having at most quadratic growth in the tails. In Section 2.4.2 we apply approximation results from the theory of strongly continuous semigroups adapted to the generalized Feller setting by [45].
Note that a completely different approach to proving the existence of affine pure-jump processes on $\mathcal{H}^+$ is taken in Chapter 6, which is based on finite-dimensional approximations of this class of affine processes. Both approaches together give valuable insights on the structure and capability of affine processes on positive Hilbert-Schmidt operators.

2.4.1 Preliminaries: generalized Feller semigroups

We recall the concept of generalized Feller semigroups introduced in [51] and further developed in [45].
Throughout this section let (Y, τ) be a complete regular Hausdorff space.

Definition 2.22. A function $\rho\colon Y \to (0, \infty)$ such that for every $R > 0$ the set $K_R := \{x \in Y : \rho(x) \leq R\}$ 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. For $f\colon Y \to \mathbb{R}$ we define $\|f\|_\rho \in [0, \infty]$ by

$$\|f\|_\rho := \sup_{x \in Y} \frac{|f(x)|}{\rho(x)}. \tag{2.51}$$

$\|f\|_\rho := \sup_{x\in Y} \frac{|f(x)|}{\rho(x)}$ Note that $\|\cdot\|_\rho$ is 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 any admissible weight function satisfies $\inf_{x\in Y} \rho(x) > 0$, we have that $C_b(Y) \subseteq B_\rho(Y)$.

Definition 2.23. We define $\mathcal{B}_\rho(Y)$ to be the closure of $C_b(Y)$ in $B_\rho(Y)$.

The following useful characterization of $\mathcal{B}_\rho(Y)$ is proven in [51, Theorem 2.7]:

Theorem 2.24. *Let (Y, ρ) be a weighted space. Then $f \in \mathcal{B}_\rho(Y)$ if and only if $f|_{K_R} \in C(K_R)$ for all $R > 0$ and*

$$\lim_{R\to\infty} \sup_{x\in Y\setminus K_R} \frac{|f(x)|}{\rho(x)} = 0\,. \tag{2.52}$$

We can now present the definition of a generalized Feller semigroup, as introduced in [51, Section 3].

Definition 2.25. A family of bounded linear operators $(P_t)_{t\geq 0}$ in $\mathcal{L}(\mathcal{B}_\rho(Y))$ is called a *generalized Feller semigroup (on $\mathcal{B}_\rho(Y)$)*, if

i) $P_0 = \mathsf{I}$, the identity on $\mathcal{B}_\rho(Y)$,

ii) $P_{t+s} = P_t P_s$ for all $t, s \geq 0$,

iii) $\lim\limits_{t\to 0+} P_t f(x) = f(x)$ for all $f \in \mathcal{B}_\rho(Y)$ and $x \in Y$,

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

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

By [51, Theorem 3.2] any generalized Feller semigroup is strongly continuous. Moreover, generalized Feller semigroups allow for a Kolmogorov theorem, see [45, Theorem 2.11] for a proof:

Theorem 2.26. *Let $(P_t)_{t\geq 0}$ be a generalized Feller semigroup on $\mathcal{B}_\rho(Y)$ satisfying $P_t\mathbf{1} = \mathbf{1}$ for all $t \geq 0$. Then there exists a filtered measurable space $(\Omega, (\mathcal{F}_t)_{t\geq 0})$ with a right-continuous filtration and a family of functions $X_t \colon \Omega \to Y$, $t \geq 0$, such that X_t is $\mathcal{F}_t$ measurable for all $t \geq 0$ and for any initial value $x \in Y$ there exists a probability measure $\mathbb{P}_x$ such that*

$$\mathbb{E}_{\mathbb{P}_x}[f(X_t)] = P_t f(x) \tag{2.53}$$

for every $t \geq 0$ and every $f \in \mathcal{B}_\rho(Y)$. Moreover, for all $x \in Y$ the process $(X_t)_{t\geq 0}$ is a time-homogeneous $\mathbb{P}_x$-Markov process, i.e., for all $x \in Y$, $0 \leq s < t$, $f \in \mathcal{B}_\rho(Y)$ we have

$$\mathbb{E}_{\mathbb{P}_x}[f(X_t) \,|\, \mathcal{F}_s] = P_{t-s} f(X_s), \tag{2.54}$$

almost surely with respect to $\mathbb{P}_x$.

Let $(P_t)_{t\geq 0}$ be a generalized Feller semigroup satisfying $P_t\mathbf{1} = \mathbf{1}$ for all $t \geq 0$. The process $(X_t)_{t\geq 0}$, the existence of which is guaranteed by Theorem 2.26, is called a *generalized Feller process* with initial value x with respect to the measure $\mathbb{P}_x$. From now on we write $\mathbb{E}_x$ for expectations with respect to the probability measure $\mathbb{P}_x$.

Remark 2.27. Let $(P_t)_{t\geq 0}$ be a generalized Feller semigroup and let $x \in Y$, then by a Riesz representation-type result (see [45, Theorem 2.4 and Remark 2.8]) $P_t\rho(x) \in \mathbb{R}$ can be defined by the integral of ρ with respect to the measure representing the linear functional $f \mapsto P_t f(x)$, $f \in \mathcal{B}_\rho(Y)$. Moreover, as there exist $M > 1$, $\omega \in \mathbb{R}$ such that $|P_t f(x)| \leq M \exp(\omega t)\rho(x)\|f\|_\rho$ for all $f \in \mathcal{B}_\rho(Y)$, we obtain

$$P_t\rho \leq M \exp(\omega t)\rho \tag{2.55}$$

for $t \geq 0$. If moreover $(P_t)_{t\geq 0}$ is associated to a Markov process $(X_t)_{t\geq 0}$ such that equation (2.53) holds, we obtain:

$$\mathbb{E}_x[\rho(X_t)] = P_t\rho(x) \leq M \exp(\omega t)\rho.$$

This can be seen by equation (2.55) and a monotone convergence argument by choosing for every $n \in \mathbb{N}$ the approximations $\rho_n = \sum_{i=1}^n \langle \cdot, e_i\rangle^2 \wedge n \in \mathcal{B}_\rho(Y)$, where $(e_i)_{i\in\mathbb{N}}$ is an ONB of $\mathcal{H}$, then $\rho_n \to \rho$ in pointwise as $n \to \infty$ and $\rho_n \leq \rho_{n+1}$ for all $n \in \mathbb{N}$.

2.4.2 Approximation of the semigroups of affine processes

We equip the Hilbert space $\mathcal{H}$ with its weak topology $\sigma(\mathcal{H},\mathcal{H}')$ (which, by the Riesz representation theorem, is the weak-$*$-topology). Note that as $\mathcal{H}^+$ is self-dual and it is closed in $(\mathcal{H},\sigma(\mathcal{H},\mathcal{H}'))$. For brevity of notation we let $\mathcal{H}^+_{\mathrm{w}}$ denote the complete regular Hausdorff space $(\mathcal{H}^+,\sigma(\mathcal{H},\mathcal{H}')_{\mathcal{H}^+})$, where $\sigma(\mathcal{H},\mathcal{H}')_{\mathcal{H}^+}$ denotes the relative topology $\sigma(\mathcal{H},\mathcal{H}')$ on $\mathcal{H}^+$. Note that the Borel σ-algebra $\mathcal{B}(\mathcal{H}^+)$ coincides with the Borel σ-algebra $\mathcal{B}(\mathcal{H}^+_{\mathrm{w}})$. In addition, we define $\rho\colon \mathcal{H}^+\to\mathbb{R}$ by

$$\rho(x) \coloneqq 1+\|x\|^2\,,\quad x\in\mathcal{H}^+\,, \tag{2.56}$$

and observe that ρ is an admissible weight function on $\mathcal{H}^+_{\mathrm{w}}$ by the Banach-Alaoglu theorem, i.e., $(\mathcal{H}^+_{\mathrm{w}},\rho)$ is a weighted space. Note that for every $R>0$, the pre-image $\{x\in\mathcal{H}^+:\ \rho(x)\leq R\}$ is compact in $\mathcal{H}^+$ equipped with the norm topology, if and only if $\mathcal{H}$ is finite-dimensional. Since we assumed throughout the chapter that $\mathcal{H}$ is infinite-dimensional, we see that ρ is not an admissible weight function in the stronger norm topology.

The linear span of the set of Fourier basis elements $\left\{\mathrm{e}^{-\langle\cdot,u\rangle}:\ u\in\mathcal{H}^+\right\}$ is denoted by

$$\mathcal{D}\coloneqq \operatorname{lin}\left(\left\{\mathrm{e}^{-\langle\cdot,u\rangle}:\ u\in\mathcal{H}^+\right\}\right)\,. \tag{2.57}$$

The relevance of this set lies in the following lemma.

Lemma 2.28. *The set $\mathcal{D}$ is dense in $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$.*

Proof. It suffices to prove that for every $\varepsilon>0$ and every $f\in C_b(\mathcal{H}^+_{\mathrm{w}})$ there exists an $f_\varepsilon\in\mathcal{D}$ such that $\|f-f_\varepsilon\|_\rho<\varepsilon$. To this end, observe that for every $\varepsilon>0$ and every $f\in C_b(\mathcal{H}^+_{\mathrm{w}})$ there exists an $R>0$ such that $\sup_{x\in\mathcal{H}^+,\|x\|>R}\frac{f(x)}{\rho(x)}<\frac{\varepsilon}{2}$, and apply Stone-Weierstrass to $C(\mathcal{H}^+_{\mathrm{w}}\cap\{x\in\mathcal{H}^+\colon \|x\|\leq R\})$. □

Corollary 2.29. *The space $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ is separable.*

Proof. Let U be a countable dense set in $(\mathcal{H}^+,\|\cdot\|)$ (recall from Section 2.1.3 that $\mathcal{H}^+$ is separable). Then by Lemma 2.28 we see that the set given by $\big\{\sum_{j=1}^{n} q_j\,\mathrm{e}^{-\langle\cdot,u_j\rangle}:\ n\in\mathbb{N},\, q_j\in\mathbb{Q},\, u_j\in U\big\}$ is dense in $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$. □

Throughout the remainder of this section let (b,B,m,μ) be an admissible parameter set, see Definition 2.3. For every $k\in\mathbb{N}$, recall $m^{(k)}$ and $\mu^{(k)}$ from (2.12) and define $\tilde{B}^{(k)}\in\mathcal{L}(\mathcal{H})$ and $\tilde{b}^{(k)}\in\mathcal{H}^+$ by

$$\begin{aligned}\tilde{B}^{(k)}(x) &\coloneqq B(x)-\int_{\mathcal{H}^+\cap\{0<\|\xi\|\leq 1\}}\xi\,\frac{\langle\mu^{(k)}(\mathrm{d}\xi),x\rangle}{\|\xi\|^2}\,,\quad x\in\mathcal{H}^+\,,\\ \tilde{b}^{(k)} &\coloneqq b-\int_{\mathcal{H}^+\cap\{0<\|\xi\|\leq 1\}}\xi\,m^{(k)}(\mathrm{d}\xi)\,.\end{aligned}$$

Note that the fact that $B \in \mathcal{L}(\mathcal{H})$ and that μ is an $\mathcal{H}^+$-valued measure, as well as (A.7) and (2.12) ensure that $\tilde{B}^{(k)} \in \mathcal{L}(\mathcal{H})$ is well-defined. Moreover, Definition 2.3 i) and Definition 2.3 ii) ensure that $\tilde{b}^{(k)} \in \mathcal{H}^+$ is well-defined. For $x \in \mathcal{H}^+$ and $k \in \mathbb{N}$ we consider the following deterministic equation in differential form:

$$\begin{cases} \mathrm{d}\mathbf{x}_t^{(x,k)} &= \big(\tilde{b}^{(k)} + \tilde{B}^{(k)}(\mathbf{x}_t^{(x,k)})\big)\,\mathrm{d}t, \qquad t \geq 0, \\ \mathbf{x}_0^{(x,k)} &= x. \end{cases} \tag{2.58}$$

Standard infinite-dimensional ODE theory ensures that for all $x \in \mathcal{H}^+$ and $k \in \mathbb{N}$ the unique classical solution to (2.58) is given by

$$\mathbf{x}_t^{(x,k)} := \mathrm{e}^{t\tilde{B}^{(k)}}\,x + \int_0^t \mathrm{e}^{(t-s)\tilde{B}^{(k)}}\,\tilde{b}^{(k)}\,\mathrm{d}s\,, \quad t \geq 0\,. \tag{2.59}$$

The following lemma provides some properties of $\mathbf{x}^{(x,k)}$, $x \in \mathcal{H}^+$, $k \in \mathbb{N}$.

Lemma 2.30. *Let (b, B, m, μ) be an admissible parameter set as in Definition 2.3. For $x \in \mathcal{H}^+$ and $k \in \mathbb{N}$ let $\mathbf{x}^{(x,k)}$ be given by* (2.59)*. Then*

$$0 \leq_{\mathcal{H}^+} \mathbf{x}_t^{(x,k+1)} \leq_{\mathcal{H}^+} \mathbf{x}_t^{(x,k)} \tag{2.60}$$

for all $k \in \mathbb{N}$, $x \in \mathcal{H}^+$, and $t \geq 0$.

Proof. It follows immediately from Definition 2.3 iv) that $\mathcal{H} \ni x \mapsto \tilde{b}^{(k)} + \tilde{B}^{(k)}(x) \in \mathcal{H}$ is quasi-monotone with respect to $\mathcal{H}^+$. As $\tilde{b}^{(k)} \in \mathcal{H}^+$, Theorem A.9 with $K = \mathcal{H}^+$, $F(\cdot) = \tilde{b}^{(k)} + \tilde{B}^{(k)}(\cdot)$, $f \equiv 0$, and $g(\cdot) = \mathbf{x}_{\cdot}^{(x,k)}$ ensures that $\mathbf{x}_t^{(x,k)} \in \mathcal{H}^+$ for all $t \geq 0$, $x \in \mathcal{H}^+$, $k \in \mathbb{N}$.
Moreover, for all $k \in \mathbb{N}$ and $x \in \mathcal{H}^+$ we have

$$\tilde{b}^{(k)} + \tilde{B}^{(k)}(x) - \Big(\tilde{b}^{(k+1)} + \tilde{B}^{(k+1)}(x)\Big) \geq_{\mathcal{H}^+} 0\,.$$

This implies that for every $x \in \mathcal{H}^+$, $k \in \mathbb{N}$, and $t \geq 0$ we have

$$\begin{aligned} \tfrac{\partial\,\mathbf{x}_t^{(x,k+1)}}{\partial t} - \Big(\tilde{b}^{(k+1)} + \tilde{B}^{(k+1)}(\mathbf{x}_t^{(x,k+1)})\Big) &= \tfrac{\partial\,\mathbf{x}_t^{(x,k)}}{\partial t} - \Big(\tilde{b}^{(k)} + \tilde{B}^{(k)}(\mathbf{x}_t^{(x,k)})\Big) \\ &\leq_{\mathcal{H}^+} \tfrac{\partial\,\mathbf{x}_t^{(x,k)}}{\partial t} - \Big(\tilde{b}^{(k+1)} + \tilde{B}^{(k+1)}(\mathbf{x}_t^{(x,k)})\Big)\,. \end{aligned}$$

Again applying Theorem A.9 with $K = \mathcal{H}^+$, $F(\cdot) = \tilde{b}^{(k)} + \tilde{B}^{(k)}(\cdot)$, $f(t) = \mathbf{x}_t^{(x,k+1)}$ and $g(t) = \mathbf{x}_t^{(x,k)}$, $t \geq 0$, implies that $\mathbf{x}_t^{(x,k+1)} \leq_{\mathcal{H}^+} \mathbf{x}_t^{(x,k)}$ for all $t \geq 0$.

□

For $k \in \mathbb{N}$, $t \geq 0$ and $f \in \mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$ define $P_t^{(\det,k)} f \colon \mathcal{H}^+ \to \mathbb{R}$ by

$$(P_t^{(\det,k)} f)(x) \coloneqq f(\mathbf{x}_t^{(x,k)}), \quad x \in \mathcal{H}^+. \tag{2.61}$$

Lemma 2.31. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3. Let $k \in \mathbb{N}$, $t \geq 0$, $f \in C_b(\mathcal{H}_{\mathrm{w}}^+)$ and let $P_t^{(\det,k)} f \colon \mathcal{H}^+ \to \mathbb{R}$ be defined by* (2.61)*. In addition, let*

$$M \coloneqq \max\{1 + 2\|\tilde{B}^{(1)}\|_{\mathcal{L}(\mathcal{H})}^{-2} \|\tilde{b}^{(1)}\|^2, 2\}, \tag{2.62}$$

$$\omega \coloneqq 2\|\tilde{B}^{(1)}\|_{\mathcal{L}(\mathcal{H})}. \tag{2.63}$$

Then, $P_t^{(\det,k)} f \in C_b(\mathcal{H}_{\mathrm{w}}^+)$,

$$\|P_t^{(\det,k)} f\|_\rho \leq M \,\mathrm{e}^{\omega t} \,\|f\|_\rho, \tag{2.64}$$

and

$$\|P_t^{(\det,k)} f\|_{\sqrt{\rho}} \leq \sqrt{M} \,\mathrm{e}^{\omega t/2} \,\|f\|_{\sqrt{\rho}}. \tag{2.65}$$

Proof. For every $t \geq 0$ the operator $\mathrm{e}^{t\tilde{B}^{(k)}}$ is strong-to-strong continuous, hence it is also weak-to-weak continuous, and thus $P_t^{(\det,k)} f \in C_b(\mathcal{H}_{\mathrm{w}}^+)$. Next, note that Lemma 2.30 implies that

$$\begin{aligned}
\tfrac{1+\|\mathbf{x}_t^{(x,k)}\|^2}{1+\|x\|^2} &\leq \tfrac{1+\|\mathbf{x}_t^{(x,1)}\|^2}{1+\|x\|^2} \\
&\leq \tfrac{1+2\,\mathrm{e}^{2t\|\tilde{B}^{(1)}\|_{\mathcal{L}(\mathcal{H})}}(\|\tilde{B}^{(1)}\|_{\mathcal{L}(\mathcal{H})}^{-2}\|\tilde{b}^{(1)}\|^2+\|x\|^2)}{1+\|x\|^2} \\
&\leq M\,\mathrm{e}^{\omega t}
\end{aligned}$$

for all $x \in \mathcal{H}^+$. Using the above estimate and (2.61) we obtain

$$\begin{aligned}
\|P_t^{(\det,k)} f\|_\rho = \sup_{x\in\mathcal{H}^+} \tfrac{(P_t^{(\det,k)} f)(x)}{1+\|x\|^2} = \sup_{x\in\mathcal{H}^+} \tfrac{f(\mathbf{x}_t^{(x,k)})}{1+\|x\|^2} \leq \|f\|_\rho \sup_{x\in\mathcal{H}^+} \tfrac{1+\|\mathbf{x}_t^{(x,k)}\|^2}{1+\|x\|^2} \\
\leq M\,\mathrm{e}^{\omega t}\,\|f\|_\rho.
\end{aligned}$$

Similarly,

$$\|P_t^{(\det,k)} f\|_{\sqrt{\rho}} = \sup_{x\in\mathcal{H}^+} \tfrac{f(\mathbf{x}_t^{(x,k)})}{\sqrt{1+\|x\|^2}} \leq \|f\|_{\sqrt{\rho}} \sup_{x\in\mathcal{H}^+} \tfrac{\sqrt{1+\|\mathbf{x}_t^{(x,k)}\|^2}}{\sqrt{1+\|x\|^2}} \leq \sqrt{M}\,\mathrm{e}^{\omega t/2}\,\|f\|_{\sqrt{\rho}}.$$

□

Recall that if $(A, \mathrm{dom}(A))$ is the generator of a strongly continuous semigroup $S = (S_t)_{t\geq 0}$ on a Banach space X, then a subspace $D \subseteq \mathrm{dom}(A)$ is a *core* for A if D is dense in $\mathrm{dom}(A)$ for the graph norm $\|\cdot\|_{\mathrm{dom}(A)} = \|\cdot\|_X + \|A\|_X$ (see [55, Chapter II, Def. 1.6]). By [55, Chapter II, Prop. 1.7] any subspace $D \subseteq \mathrm{dom}(A)$ that is dense in X and invariant under S is a core.

Lemma 2.32. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3. For all $k \in \mathbb{N}$, $t \geq 0$, $f \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ let $P^{(\det,k)}_t f\colon \mathcal{H}^+ \to \mathbb{R}$ be defined by (2.61). Then $(P^{(\det,k)}_t)_{t\geq 0}$ is a generalized Feller semigroup on both $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ and $\mathcal{B}_{\sqrt{\rho}}(\mathcal{H}^+_{\mathrm{w}})$ for all $k \in \mathbb{N}$. Moreover, $\mathcal{D}$ is a core for the generator $\mathcal{G}^{(k)}_{\det}$ of $(P^{(\det,k)}_t)_{t\geq 0}$ on $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ and for all $f \in \mathcal{D}$ we have*

$$(\mathcal{G}^{(k)}_{\det} f)(x) = \langle \tilde{b}^{(k)} + \tilde{B}^{(k)}(x), f'(x)\rangle, \quad x \in \mathcal{H}^+. \tag{2.66}$$

Proof. Let $k \in \mathbb{N}$. It follows from Lemma 2.31 that $(P^{(\det,k)}_t)_{t\geq 0}$ is a family of bounded linear operators on both $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ and $\mathcal{B}_{\sqrt{\rho}}(\mathcal{H}^+_{\mathrm{w}})$. Moreover, properties i), ii), and v) in Definition 2.25 are trivially satisfied. Property iv) follows from Lemma 2.31. Finally, property iii) follows from Theorem 2.24 and the fact that $\lim_{t\to 0^+} \|\mathbf{x}^{(x,k)}_t - x\| = 0$.
It is easily verified that $\mathcal{D}$ is a subspace of $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ that is invariant under the semigroup $(P^{(\det,k)}_t)_{t\geq 0}$. We know from Lemma 2.28 that $\mathcal{D}$ is dense in $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$, thus by [55, Chapter II, Prop. 1.7] it remains to prove that $\mathcal{D} \subseteq \mathrm{dom}(\mathcal{G}^{(k)}_{\det})$ and that (2.66) holds. To this end, let $u \in \mathcal{H}^+$ and consider $f(\cdot) = \mathrm{e}^{-\langle u,\cdot\rangle} \in \mathcal{D}$. For f of this latter form, we define $f'(x) := -\mathrm{e}^{-\langle u,x\rangle}u$, for $u, x \in \mathcal{H}^+$ and $f''(x)$ to be the bounded linear map on $\mathcal{H}^+$ defined for $u, x \in \mathcal{H}^+$ by $f''(x)(v) := \mathrm{e}^{-\langle u,x\rangle}u\langle u, v\rangle$, $v \in \mathcal{H}^+$. Now, observe that for $\tilde{B}(x) := \tilde{B}^{(k)}(x) + \tilde{b}^{(k)}$, we have

$$\begin{aligned}
&\frac{(P^{(\det,k)}_t f)(x) - f(x)}{t} - \langle f'(x), \tilde{B}(x)\rangle \\
&= \int_0^1 \left\langle f'(s(\mathbf{x}^{(x,k)}_t - x) + x), \frac{\mathbf{x}^{(x,k)}_t - x}{t} - \tilde{B}(x) \right\rangle \mathrm{d}s \\
&\quad + \int_0^1 \int_0^1 \left\langle f'' \left(us\big(\mathbf{x}^{(x,k)}_t - x\big) + x \right) \left(s\big(\mathbf{x}^{(x,k)}_t - x\big) \right), \tilde{B}(x) \right\rangle \mathrm{d}u\, \mathrm{d}s,
\end{aligned} \tag{2.67}$$

where we used Lemma 2.19 twice. Note that this lemma is applicable as the one-sided derivatives of f, considered as a function on $\mathcal{H}^+$, exist.

Observe that

$$\begin{aligned}
&\lim_{t\to 0+}\sup_{x\in\mathcal{H}^+}\frac{\left|\frac{1}{t}\left(\mathbf{x}_t^{(x,k)}-x\right)-\left(\tilde{B}^{(k)}x+\tilde{b}^{(k)}\right)\right|}{\sqrt{\rho(x)}}\\
&\leq \lim_{t\to 0+}\sup_{x\in\mathcal{H}^+}\frac{\|\tilde{B}^{(k)}\|_{\mathcal{L}(\mathcal{H})}\|e^{t\tilde{B}^{(k)}}-\mathrm{I}\|_{\mathcal{L}(\mathcal{H})}\|x\|+\frac{1}{t}\|\tilde{b}^{(k)}\|\int_0^t\|e^{(t-s)\tilde{B}^{(k)}}-\mathrm{I}\|_{\mathcal{L}(\mathcal{H})}\,\mathrm{d}s}{\sqrt{1+\|x\|^2}}=0.
\end{aligned} \tag{2.68}$$

Moreover, we have

$$\lim_{t\to 0+}\sup_{x\in\mathcal{H}^+}\frac{|\mathbf{x}_t^{(x,k)}-x|}{\sqrt{\rho(x)}}\leq \lim_{t\to 0+}\sup_{x\in\mathcal{H}^+}\frac{\|e^{t\tilde{B}^{(k)}}-\mathrm{I}\|_{\mathcal{L}(\mathcal{H})}\|x\|+\int_0^t\|e^{(t-s)\tilde{B}^{(k)}}\tilde{b}^{(k)}\|\,\mathrm{d}s}{\sqrt{\rho(x)}}=0. \tag{2.69}$$

Since $\sup_{x\in\mathcal{H}^+}|\rho(x)|^{-\frac{1}{2}}\|f'(x)\|<\infty$ and $\sup_{x\in\mathcal{H}^+}\|f''(x)\|_{\mathcal{L}(\mathcal{H})}<\infty$, it follows from equations (2.67), (2.68), and (2.69) that

$$\lim_{t\to 0+}\left\|\frac{(P_t^{(\mathrm{det},k)}f)(x)-f(x)}{t}-\langle f'(x),\tilde{B}^{(k)}(x)+\tilde{b}^{(k)}\rangle\right\|_\rho=0. \tag{2.70}$$

This, the linearity of $\mathcal{G}_{\mathrm{det}}^{(k)}$ and the fact that $\mathcal{D}$ is invariant for $P_t^{(\mathrm{det},k)}$ (and thus $P_t^{(\mathrm{det},k)}f\in\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$ whenever $f\in\mathcal{D}$) implies that $\mathcal{D}\subseteq\mathrm{dom}(\mathcal{G}_{\mathrm{det}}^{(k)})$ and that (2.66) holds. □

We now introduce the family of measures $\nu^{(k)}\colon\mathcal{H}^+\times\mathcal{B}(\mathcal{H}^+\setminus\{0\})\to[0,\infty)$ for every $x\in\mathcal{H}^+$ given by

$$\nu^{(k)}(x,\mathrm{d}\xi)=m^{(k)}(\mathrm{d}\xi)+\frac{\langle\mu^{(k)}(\mathrm{d}\xi),x\rangle}{\|\xi\|^2} \tag{2.71}$$

and define the operator $\mathcal{G}_{\mathrm{jump}}^{(k)}\colon\ \mathrm{dom}(\mathcal{G}_{\mathrm{jump}}^{(k)})\subseteq\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)\to\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$ by

$$\begin{aligned}
&\mathrm{dom}(\mathcal{G}_{\mathrm{jump}}^{(k)})\\
&=\left\{f\in\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)\colon\left(x\mapsto\int_{\mathcal{H}^+\setminus\{0\}}(f(\xi+x)-f(x))\,\nu^{(k)}(x,\mathrm{d}\xi)\right)\in\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)\right\}
\end{aligned} \tag{2.72}$$

and for $f\in\mathrm{dom}(\mathcal{G}_{\mathrm{jump}}^{(k)})$:

$$\mathcal{G}_{\mathrm{jump}}^{(k)}f(x)\coloneqq\int_{\mathcal{H}^+\setminus\{0\}}(f(\xi+x)-f(x))\,\nu^{(k)}(x,\mathrm{d}\xi),\quad x\in\mathcal{H}^+. \tag{2.73}$$

Note that for all $k \in \mathbb{N}$ the measure $\nu^{(k)}(x, \mathrm{d}\xi)$ is finite, i.e. $\nu^{(k)}(x, \mathcal{H}^+ \setminus \{0\}) < \infty$ for all $x \in \mathcal{H}^+$, but it is an affine function in x and hence unbounded in the first component. For that reason $\mathcal{G}^{(k)}_{\text{jump}} f$ may not be in $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ for all $f \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$. However, the following lemma ensures that $C_b(\mathcal{H}^+_{\mathrm{w}}) \subseteq \mathrm{dom}(\mathcal{G}^{(k)}_{\text{jump}})$:

Lemma 2.33. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3. Let $k \in \mathbb{N}$, and let $\mathcal{G}^{(k)}_{\text{jump}}$ be as defined in* (2.72) *and* (2.73)*. Then $C_b(\mathcal{H}^+_{\mathrm{w}}) \subseteq \mathrm{dom}(\mathcal{G}^{(k)}_{\text{jump}})$.*

Proof. Let $f \in C_b(\mathcal{H}^+_{\mathrm{w}})$ and let $g_f \colon \mathcal{H}^+ \to \mathbb{R}$ be defined by

$$g_f(x) = \int_{\mathcal{H}^+ \setminus \{0\}} f(x+\xi) \tfrac{\langle \mu^{(k)}(\mathrm{d}\xi), x \rangle}{\|\xi\|^2} \tag{2.74}$$

We will prove that $g_f \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ using Theorem 2.24. All other terms in the definition of $\mathcal{G}^{(k)}_{\text{jump}} f$ can be dealt with in a similar (simpler) way.
To see that g_f is continuous on $K_R := \{\rho \leq R\}$ for all $R > 0$ it suffices to show that g_f is sequentially continuous on K_R for every $R > 0$ as the weak topology restricted to K_R is metrizable. Fix $R > 0$ and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in K_R converging (weakly) to an $x \in K_R$. By the dominated convergence theorem (Theorem A.8) and the fact that $\sup_{n \in \mathbb{N}} \|x_n\| \leq \sqrt{R}$ we obtain

$$\begin{aligned} \lim_{n\to\infty} |g_f(x_n) - g_f(x)| &\leq \lim_{n\to\infty} \left\| \int_{\mathcal{H}^+ \setminus \{0\}} (f(x_n+\xi) - f(x+\xi)) \tfrac{\mu^{(k)}(\mathrm{d}\xi)}{\|\xi\|^2} \right\| \|x_n\| \\ &\quad + \lim_{n\to\infty} \left| \int_{\mathcal{H}^+ \setminus \{0\}} f(x+\xi) \tfrac{\langle \mu^{(k)}, x_n - x \rangle (\mathrm{d}\xi)}{\|\xi\|^2} \right| = 0. \end{aligned}$$

Finally, observe that

$$\lim_{R\to\infty} \sup_{x \in \mathcal{H}^+ : \rho(x) \geq R} |\rho(x)|^{-1} |g_f(x)| = 0$$

as f is bounded and $\int_{\mathcal{H}^+ \setminus \{0\}} \frac{\mu^{(k)}(\mathrm{d}\xi)}{\|\xi\|^2} \in \mathcal{H}$ (recall (A.7)). By Theorem 2.24 this ensures that $g_f \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$, which completes the proof of the lemma. □

In the next proposition we achieve an important intermediate stage, that allows us to conclude the existence of generalized Feller processes in $\mathcal{H}^+$ admitting for bounded drifts and finite-activity jump behavior, as well as satisfying the exponential affine formula (2.2).

Proposition 2.34. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3. Let $k \in \mathbb{N}$, and let $(\phi^{(k)}(\cdot, u), \psi^{(k)}(\cdot, u))$ be the unique solution to* (2.14) *(cf. Proposition 2.14). Let $\mathcal{D} \subseteq \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ be given by* (2.57) *and $\mathcal{G}^{(k)}_{\mathrm{det}}$ and $\mathcal{G}^{(k)}_{\mathrm{jump}}$ be as defined in* (2.66)*, respectively* (2.73)*. Consider the operator $\mathcal{G}^{(k)}_{\mathrm{det}} + \mathcal{G}^{(k)}_{\mathrm{jump}} \colon \mathrm{dom}(\mathcal{G}^{(k)}_{\mathrm{det}}) \cap \mathrm{dom}(\mathcal{G}^{(k)}_{\mathrm{jump}}) \subseteq \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}}) \to \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$. Then $\mathcal{D} \subseteq \mathrm{dom}(\mathcal{G}^{(k)}_{\mathrm{det}}) \cap \mathrm{dom}(\mathcal{G}^{(k)}_{\mathrm{jump}})$. Moreover, there exists a generalized Feller semigroup $(P^{(k)}_t)_{t\geq 0}$ with generator $(\mathcal{G}^{(k)}, \mathrm{dom}(\mathcal{G}^{(k)}))$ such that*

i) $\mathcal{D} \subseteq \mathrm{dom}(\mathcal{G}^{(k)})$,

ii) $\mathcal{G}^{(k)} f = (\mathcal{G}^{(k)}_{\mathrm{det}} + \mathcal{G}^{(k)}_{\mathrm{jump}}) f$ *for all* $f \in \mathcal{D}$,

iii) $P^{(k)}_t \mathbf{1} = \mathbf{1}$ *for all* $t \geq 0$*, and*

iv) for all $u, x \in \mathcal{H}^+$, $t \geq 0$ *we have*

$$\left(P^{(k)}_t \, \mathrm{e}^{-\langle \cdot, u\rangle}\right)(x) = \mathrm{e}^{-\phi^{(k)}(t,u) - \langle x, \psi^{(k)}(t,u)\rangle}. \tag{2.75}$$

Proof of Proposition 2.34. Roughly speaking, we can ensure the existence of a generalized Feller semigroup $P^{(k)}_t$ satisfying ii) in Proposition 2.34 by verifying that all conditions of [45, Proposition 3.3] are satisfied. However, the assertions of [45, Proposition 3.3] do not immediately give us i), iii), and iv). In order to obtain these statements we need to dig into the proof of [45, Proposition 3.3], which makes this proof somewhat technical and tricky. To enhance the readability, we split the proof in to several parts.

Step 1: Verifying the assumptions of [45, Proposition 3.3]. We consider, in the notation of that Proposition, $(X, \rho) = (\mathcal{H}^+_{\mathrm{w}}, \rho)$, $A = \mathcal{G}^{(k)}_{\mathrm{det}}$, ω as in (2.63), $M_1 = M$ where M is as in (2.62), $\mu(x, E) = \nu^{(k)}(x, E - x \cap \mathcal{H}^+)$ (recall the definition of $\nu^{(k)}$ from (2.71); here $E - x \coloneqq \{y \in \mathcal{H} \colon y + x \in E\}$), and $B = \mathcal{G}^{(k)}_{\mathrm{jump}}$. By Lemma 2.32, $\mathcal{G}^{(k)}_{\mathrm{det}}$ is the generator of a generalized Feller semigroup $(P^{(\mathrm{det},k)}_t)_{t\geq 0}$ of transport type on both $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ and $\mathcal{B}_{\sqrt{\rho}}(\mathcal{H}^+_{\mathrm{w}})$. In particular, by [51, Theorem 3.2], $(P^{(\mathrm{det},k)}_t)_{t\geq 0}$ defines a strongly continuous semigroup on both $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ and $\mathcal{B}_{\sqrt{\rho}}(\mathcal{H}^+_{\mathrm{w}})$, i.e., it automatically holds that the domain of $\mathcal{G}^{(k)}_{\mathrm{det}}$ is dense and that $P^{(\mathrm{det},k)}_t)_{t\geq 0}$ allows for exponential bounds (see Lemma 2.31 for explicit bounds). Lemma 2.33 implies that $\mathcal{G}^{(k)}_{\mathrm{jump}} f$ is weakly continuous on compact sets $\{\rho \leq R\}$ for all $R \geq 0$ and all $f \in C_b(\mathcal{H}^+_{\mathrm{w}})$.

Moreover, one easily verifies that there exists a constant K (possibly depending on k) such that for all $x \in \mathcal{H}^+$ we have

$$\begin{aligned}\int_{\mathcal{H}^+\setminus\{0\}} \rho(y+x)\,\nu^{(k)}(x,\mathrm{d}y) &\leq \int_{\mathcal{H}^+\setminus\{0\}} (1+2\|x\|^2+2\|y\|^2)\,\nu^{(k)}(x,\mathrm{d}y)\\ &\leq K|\rho(x)|^2\,,\end{aligned} \tag{2.76}$$

similarly for the square-root

$$\begin{aligned}\int_{\mathcal{H}^+\setminus\{0\}} \sqrt{\rho(y+x)}\,\nu^{(k)}(x,\mathrm{d}y) &\leq \int_{\mathcal{H}^+\setminus\{0\}} (1+\|x\|+\|y\|)\,\nu^{(k)}(x,\mathrm{d}y)\\ &\leq K\rho(x)\,,\end{aligned} \tag{2.77}$$

and lastly

$$\int_{\mathcal{H}^+\setminus\{0\}} \nu^{(k)}(x,\mathrm{d}y) \leq K\sqrt{\rho(x)}\,. \tag{2.78}$$

Next, observe that by Lemma 2.30 and the fact that $(0,B,0,\mu)$ is also an admissible parameter set, we have $e^{t\tilde{B}^{(k)}}\xi \in \mathcal{H}^+$ whenever $\xi \in \mathcal{H}^+$. Thus

$$\begin{aligned}&P_t^{(\det,k)}\rho(\xi+x) = 1+\|\mathbf{x}_t^{(\xi+x,k)}\|^2 = 1+\|e^{t\tilde{B}^{(k)}}\xi+\mathbf{x}_t^{(\xi,k)}\|^2\\ &= P_t^{(\det,k)}\rho(x)+2\langle e^{t\tilde{B}^{(k)}}\xi,\mathbf{x}_t^{(\xi,k)}\rangle+\|\,\mathrm{e}^{t\tilde{B}^{(k)}}\,\xi\|^2 \geq P_t^{(\det,k)}\rho(x)\end{aligned} \tag{2.79}$$

for all $x,\xi \in \mathcal{H}^+$. This, together with estimates similar to (2.66) yields (note that $\|\,\mathrm{e}^{t\tilde{B}^{(k)}}\,\xi\| \leq \|\,\mathrm{e}^{t\tilde{B}^{(1)}}\,\xi\|$, and recall ω from (2.63))

$$\begin{aligned}&\left|\frac{\sup_{t\geq 0}\mathrm{e}^{-\omega t}\,P_t^{(\det,k)}\rho(\xi+x)-\sup_{t\geq 0}\mathrm{e}^{-\omega t}\,P_t^{(\det,k)}\rho(x)}{\sup_{t\geq 0}\mathrm{e}^{-\omega t}\,P_t^{(\det,k)}\rho(x)}\right|\\ &= \frac{\sup_{t\geq 0}\mathrm{e}^{-\omega t}\,P_t^{(\det,k)}\rho(\xi+x)-\sup_{t\geq 0}\mathrm{e}^{-\omega t}\,P_t^{(\det,k)}\rho(x)}{\sup_{t\geq 0}\mathrm{e}^{-\omega t}\,P_t^{(\det,k)}\rho(x)}\\ &\leq \frac{\sup_{t\geq 0}\mathrm{e}^{-\omega t}\left|\|\,\mathrm{e}^{t\tilde{B}^{(k)}}\,\xi\|^2+2\|\,\mathrm{e}^{t\tilde{B}^{(k)}}\,\xi\|\|\mathbf{x}_t^{(x,k)}\|\right|}{1+\|x\|^2} \leq \frac{\|\xi\|^2+2\|\xi\|(\|x\|+\sqrt{M})}{1+\|x\|^2}\\ &\leq \frac{(M+2\|\xi\|^2)(1+\|x\|)}{1+\|x\|^2} \leq \frac{2M+4\|\xi\|^2}{1+\|x\|}, \quad \forall\, x,\xi \in \mathcal{H}^+.\end{aligned} \tag{2.80}$$

It follows that for all $x \in \mathcal{H}^+$ we have

$$\begin{aligned}&\int_{\mathcal{H}^+\setminus\{0\}} \left| \frac{\sup_{t\geq 0} \mathrm{e}^{-\omega t}\left(P_t^{(\det,k)}\rho\right)(\xi+x) - \sup_{t\geq 0} \mathrm{e}^{-\omega t}\left(P_t^{(\det,k)}\rho\right)(x)}{\sup_{t\geq 0} \mathrm{e}^{-\omega t}\left(P_t^{(\det,k)}\rho\right)(x)} \right| \nu^{(k)}(x,\mathrm{d}\xi)\\ &\leq \sup_{y\in\mathcal{H}^+} \left| \int_{\mathcal{H}^+\setminus\{0\}} \left(\tfrac{2M+4\|\xi\|^2}{1+\|y\|}\right) \nu^{(k)}(y,\mathrm{d}\xi) \right| =: \tilde{\omega}_k < \infty.\end{aligned} \tag{2.81}$$

This ensures that all conditions of [45, Proposition 3.3] are satisfied.
Step 2: Presenting the assertions of [45, Proposition 3.3]. As in the proof of [45, Proposition 3.3], we introduce the operator $\mathcal{G}_{\mathrm{jump}}^{(k,n)} \in \mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+))$ which satisfies

$$(\mathcal{G}_{\mathrm{jump}}^{(k,n)} f)(x) = \int_{\mathcal{H}^+\setminus\{0\}} (f(\xi+x)-f(x)) \frac{n}{\rho(\xi+x)\wedge n}\, \nu^{(k)}(x,\mathrm{d}\xi)$$

for all $x \in \mathcal{H}^+$, $f \in \mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$. Note that $\mathcal{D} \subseteq \mathrm{dom}(\mathcal{G}_{\mathrm{jump}}^{(k)})$ by Lemma 2.33. For future reference (see Proposition 2.40 below) we also introduce $\tilde{\rho}_k \colon \mathcal{H}^+ \to \mathbb{R}$, $\tilde{\rho}_k(x) = \sup_{t\geq 0} e^{-\omega t} P_t^{(\det,k)}\rho(x)$. It follows from [45, Remark 2.9] that $\tilde{\rho}_k$ is an admissible weight function and that $\|\cdot\|_\rho \leq \|\cdot\|_{\tilde{\rho}_k} \leq M\,\|\cdot\|_\rho$. Moreover, it follows from the proof of [45, Proposition 3.3] (with $A = \mathcal{G}_{\det}^{(k)}$ and $B_n = \mathcal{G}_{\mathrm{jump}}^{(k,n)}$) that $\mathcal{G}_{\det}^{(k)} + \mathcal{G}_{\mathrm{jump}}^{(k,n)}$ is the generator of a generalized Feller semigroup $(P_t^{(k,n)})_{t\geq 0}$ on $\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$ for all $n \in \mathbb{N}$, such that

a) $\|P_t^{(k,n)}\|_{\mathcal{L}(\mathcal{B}_{\tilde{\rho}_k}(\mathcal{H}_{\mathrm{w}}^+))} \leq e^{(\omega+\tilde{\omega}_k)t}$ for all $t \geq 0$, $n \in \mathbb{N}$,

b) $\|P_t^{(k,n)}\|_{\mathcal{L}(\mathcal{B}_{\rho}(\mathcal{H}_{\mathrm{w}}^+))} \leq M\,\mathrm{e}^{(\omega+\tilde{\omega}_k)t}$ for all $t \geq 0$, $n \in \mathbb{N}$,

c) $\lim_{n\to\infty} \|(\mathcal{G}_{\mathrm{jump}}^{(k,n)} - \mathcal{G}_{\mathrm{jump}}^{(k)})f\|_\rho = 0$ for all $f \in \mathcal{D}$.

It moreover follows from the proof of [45, Proposition 3.3] that there exists a generalized Feller semigroup $(P_t^{(k)})_{t\geq 0}$ on $\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$ with generator $\mathcal{G}^{(k)}$ satisfying

$$\lim_{n\to\infty} \sup_{s\in[0,t]} \|(P_s^{(k,n)} - P_s^{(k)})f\|_\rho = 0, \text{ for all } f \in \mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+),\ t \geq 0. \tag{2.82}$$

Step 3: Proof of i) and ii). Fix $f \in \mathcal{D}$. Let $u_{k,n}(t) = P_t^{(k,n)} f$, $t \geq 0$ and $n \in \mathbb{N}$, let $u_k(t) = P_t^{(k)} f$, $t \geq 0$, and let $v_k(t) = P_t^{(k)}(\mathcal{G}_{\det}^{(k)} + \mathcal{G}_{\mathrm{jump}}^{(k)})f$. Observe that $u'_{k,n}(t) = P_t^{(k,n)}(\mathcal{G}_{\det}^{(k)} + \mathcal{G}_{\mathrm{jump}}^{(k,n)})f$. By a), b), and (2.82) we have, for all $T \geq 0$, that

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \left(\|u_{k,n}(t) - u_k(t)\|_\rho + \|u'_{k,n}(t) - v_k(t)\|_\rho\right) = 0. \tag{2.83}$$

This implies that u_k is differentiable and $u_k'(t) = v_k(t)$, which implies that $f \in \operatorname{dom}(\mathcal{G}^{(k)})$ and $\mathcal{G}^{(k)} f = u_k'(0) = (\mathcal{G}^{(k)}_{\text{det}} + \mathcal{G}^{(k)}_{\text{jump}})f$.

Step 4: Proof of iii). In order to verify that $P_t^{(k)}\mathbf{1} = \mathbf{1}$ for all $t \geq 0$, observe that $\mathcal{G}^{(k,n)}_{\text{jump}}\mathbf{1} = \mathbf{0}$ (the constant zero function), whence $e^{t\mathcal{G}^{(k,n)}_{\text{jump}}}\mathbf{1} = \mathbf{1}$ for all $t \geq 0$ and the Trotter product formula (see, e.g., [55, Chapter III, Corollary 5.8]) implies that $P_t^{(k,n)}\mathbf{1} = \mathbf{1}$ for all $t \geq 0$. It then follows that $P_t^{(k)}\mathbf{1} = \mathbf{1}$ for all $t \geq 0$.

Step 5: Proof of iv). Recall the functions $R^{(k)}$ and $F^{(k)}$ from (2.13a) and (2.13b). Recall from Lemmas 2.32 and 2.33 that $\mathrm{e}^{-\langle \cdot, u\rangle} \in \mathcal{D} \subseteq \operatorname{dom}(\mathcal{G}^{(k)}_{\text{det}}) \cap \operatorname{dom}(\mathcal{G}^{(k)}_{\text{jump}})$ for all $u \in \mathcal{H}^+$, and that

$$\begin{aligned}
\mathcal{G}^{(k)}(\mathrm{e}^{-\langle \cdot, u\rangle})(x) &= (\mathcal{G}^{(k)}_{\text{det}} + \mathcal{G}^{(k)}_{\text{jump}})(\mathrm{e}^{-\langle \cdot, u\rangle})(x) \\
&= \Big(-\langle \tilde{b}^{(k)} + \tilde{B}^{(k)}(x), u\rangle + \int_{\mathcal{H}^+\setminus\{0\}} \big(\mathrm{e}^{-\langle \xi, u\rangle} - 1\big)\nu^{(k)}(x, \mathrm{d}\xi)\Big)\,\mathrm{e}^{-\langle x, u\rangle} \\
&= \big(-F^{(k)}(u) - \langle x, R^{(k)}(u)\rangle\big)\,\mathrm{e}^{-\langle x, u\rangle},
\end{aligned} \tag{2.84}$$

for all $u, x \in \mathcal{H}^+$. On the other hand, Proposition 2.14 implies that

$$\begin{aligned}
&\frac{\partial}{\partial t}\,\mathrm{e}^{-\phi^{(k)}(t,u) - \langle x, \psi^{(k)}(t,u)\rangle} \\
&= \Big(-F^{(k)}(\psi^{(k)}(t,u)) - \langle x, R^{(k)}(\psi^{(k)}(t,u))\rangle\Big)\,\mathrm{e}^{-\phi^{(k)}(t,u) - \langle x, \psi^{(k)}(t,u)\rangle},
\end{aligned}$$

for all $u, x \in \mathcal{H}^+$. Therefore for all $u \in \mathcal{H}^+$, we conclude that the mapping

$$[0,\infty) \ni t \mapsto \mathrm{e}^{-\phi^{(k)}(t,u) - \langle \cdot, \psi^{(k)}(t,u)\rangle} \in \mathcal{D} \subseteq \operatorname{dom}(\mathcal{G}^{(k)})$$

is a classical solution to the following abstract Cauchy problem:

$$\begin{cases} \dfrac{\partial v(t)}{\partial t} = \mathcal{G}^{(k)} v(t), \\ \quad v(0) = \mathrm{e}^{-\langle \cdot, u\rangle}. \end{cases}$$

By the uniqueness of the classical solution we conclude (2.75). □

From Proposition 2.34, which ensures the existence of the generalized Feller semigroup $(P_t^{(k)})_{t\geq 0}$ with $P_t^{(k)}\mathbf{1} = \mathbf{1}$, together with the version of Kolmogorov's extension Theorem 2.26, we conclude that there exists a generalized Feller process associated to $(P_t^{(k)})_{t\geq 0}$, denoted by $(X_t^{(k)})_{t\geq 0}$, such that $\mathbb{E}_x\left[f(X_t^{(k)})\right] = P_t^{(k)} f(x)$ for every $f \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$.

Item a) and equation (2.82) in the proof of Proposition 2.34 result in exponential bounds on $\|P_t^{(k)}\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+))}$ *that depend on* $k \in \mathbb{N}$. In order to proceed, we need to establish bounds that are *uniform* in k. We begin with a lemma that builds on top of the results in Proposition 2.20:

Lemma 2.35. *Let* (b, B, m, μ) *be an admissible parameter set conform Definition 2.3. Moreover for every* $k \in \mathbb{N}$*, let* $(\phi^{(k)}(\cdot, u), \psi^{(k)}(\cdot, u))$ *be the solution of* (2.14)*, the existence of which is established in Proposition 2.14, and the mappings* $\mathrm{d}_+\phi(\cdot, 0)$*,* $\mathrm{d}_+\psi(\cdot, 0)$*,* $\mathrm{d}_+^2\phi^{(k)}(\cdot, 0)$ *and* $\mathrm{d}_+^2\psi^{(k)}(\cdot, 0)$ *be as in Proposition 2.20 for the admissible parameter set* $(b, B, m^{(k)}, \mu^{(k)})$*. Moreover, let* $(X_t^{(k)})_{t\geq 0}$ *be the generalized Feller process associated to* $(P_t^{(k)})_{t\geq 0}$*. Then for every* $v, w \in \mathcal{H}$ *and* $t \geq 0$ *the following formulas hold true:*

$$\mathbb{E}_x\left[\langle X_t^{(k)}, v\rangle\right] = \mathrm{d}_+\phi(t,0)(v) + \langle x, \mathrm{d}_+\psi(t,0)(v)\rangle, \tag{2.85}$$

and

$$\begin{aligned}\mathbb{E}_x\left[\langle X_t^{(k)}, v\rangle\langle X_t^{(k)}, w\rangle\right] = &-\mathrm{d}_+^2\phi^{(k)}(t,0)(v,w) - \langle x, \mathrm{d}_+^2\psi^{(k)}(t,0)(v,w)\rangle\\ &+ \big(\mathrm{d}_+\phi(t,0)(v) + \langle x, \mathrm{d}_+\psi(t,0)(v)\rangle\big)\\ &\times \big(\mathrm{d}_+\phi(t,0)(w) + \langle x, \mathrm{d}_+\psi(t,0)(w)\rangle\big).\end{aligned} \tag{2.86}$$

Proof. Let $k \in \mathbb{N}$ arbitrary, but fixed. Recall from Remark 2.27 that for all $t \geq 0$:

$$\mathbb{E}_x\left[\|X_t^{(k)}\|^2\right] < \infty, \quad \forall x \in \mathcal{H}^+. \tag{2.87}$$

We first show that the formulas (2.85) and (2.86) holds for $v, w \in \mathcal{H}^+$ and subsequently extend these to $v, w \in \mathcal{H}$. Let $u \in \mathcal{H}^+$, $x \in \mathcal{H}^+$, $t \geq 0$ and set

$$\Phi^{(k)}(t,u,x) \coloneqq \mathrm{e}^{-\phi^{(k)}(t,u) - \langle x, \psi^{(k)}(t,u)\rangle},$$

and by the affine property of $(X_t^{(k)})_{t\geq 0}$ from equation (2.75) we have

$$\mathbb{E}_x\left[\mathrm{e}^{-\langle X_t^{(k)}, u\rangle}\right] = \Phi^{(k)}(t,u,x). \tag{2.88}$$

By Proposition 2.20 the right-hand side of equation (2.88) is one-sided differentiable in $u \in \mathcal{H}^+$ in the direction v for every $v \in \mathcal{H}^+$. In particular, by applying the chain-rule at $u = 0$ we have:

$$\begin{aligned}\mathrm{d}_+\Phi^{(k)}(t,0,x)(v) &= \big(-\mathrm{d}_+\phi^{(k)}(t,0)(v) - \langle x, \mathrm{d}_+\psi^{(k)}(t,0)(v)\rangle\big)\Phi^{(k)}(t,0,x)\\ &= -\mathrm{d}_+\phi^{(k)}(t,0)(v) - \langle x, \mathrm{d}_+\psi^{(k)}(t,0)(v)\rangle,\end{aligned} \tag{2.89}$$

where $\mathrm{d}_+\phi^{(k)}(t,0) = \mathrm{d}_+\phi(t,0)$ and $\mathrm{d}_+\psi^{(k)}(t,0) = \mathrm{d}_+\psi(t,0)$ for all $t \geq 0$ and $k \in \mathbb{N}$, see Lemma 2.18.

Moreover, note that for $\theta \in \mathbb{R}^+$ the random variable $\mathrm{e}^{-\langle X_t^{(k)}, \theta v\rangle}$ is integrable and for $\mathbb{P}_x$-almost all $\omega \in \Omega$ the mapping $\theta \mapsto \mathrm{e}^{-\langle X_t^{(k)}(\omega), \theta v\rangle}$ is differentiable. Due to equation (2.87) the term

$$\sup_{\theta \in [0,1]} \left| \frac{\mathrm{d}}{\mathrm{d}\theta} \mathrm{e}^{-\langle X_t^{(k)}, \theta v\rangle} \right| = \sup_{\theta \in [0,1]} \left| -\langle X_t^{(k)}, v\rangle \, \mathrm{e}^{-\langle X_t^{(k)}, \theta v\rangle} \right|$$

is integrable. Hence, all the requirements for switching the derivative with respect to θ and the expectation with respect to $\mathbb{P}_x$ are fulfilled, thus the left-hand side of equation (2.88) together with equation (2.89) yields:

$$\mathbb{E}_x \left[\langle X_t^{(k)}, v\rangle \right] = \mathrm{d}_+\phi(t,0)(v) + \langle x, \mathrm{d}_+\psi(t,0)(v)\rangle. \tag{2.90}$$

Again due to equation (2.87) we obtain by differentiating both sides of equation (2.88) at $u = 0$ twice in the direction v and w the formula in (2.86). Note that for every $v \in \mathcal{H}$ there exist $v^+, v^- \in \mathcal{H}^+$ such that $v = v^+ - v^-$, by linearity of the formula (2.85) in v, we have:

$$\begin{aligned}
\mathbb{E}_x \left[\langle X_t^{(k)}, v\rangle \right] &= \mathbb{E}_x \left[\langle X_t^{(k)}, v^+\rangle \right] - \mathbb{E}_x \left[\langle X_t^{(k)}, v^-\rangle \right] \\
&= \mathrm{d}_+\phi(t,0)(v^+) - \mathrm{d}_+\phi(t,0)(v^-) \\
&\quad + \langle x, \mathrm{d}_+\psi(t,0)(v^+) - \mathrm{d}_+\psi(t,0)(v^-)\rangle \\
&= \mathrm{d}_+\phi(t,0)(v) + \langle x, \mathrm{d}_+\psi(t,0)(v)\rangle.
\end{aligned}$$

By introducing the linear functional

$$\langle\langle \cdot, v \otimes w\rangle\rangle \colon \mathcal{H} \otimes \mathcal{H} \to \mathbb{R} \text{ defined by } \langle\langle x \otimes x, v \otimes w\rangle\rangle := \langle x, v\rangle \langle x, w\rangle, \tag{2.91}$$

we can write

$$\mathbb{E}_x \left[\langle X_t^{(k)}, v\rangle \langle X_t^{(k)}, w\rangle \right] = \mathbb{E}_x \left[\langle\langle X_t^{(k)} \otimes X_t^{(k)}, v \otimes w\rangle\rangle \right]$$

for every $v, w \in \mathcal{H}^+$ and we have

$$\begin{aligned}
\mathbb{E}_x \left[\langle\langle X_t^{(k)} \otimes X_t^{(k)}, v \otimes w\rangle\rangle \right] = &-\langle\langle \mathrm{d}_+^2\phi^{(k)}(t,0) + \mathrm{d}_+^2\psi^{(k)}(t,0)^*(x), v \otimes w\rangle\rangle \\
&+ \langle\langle \mathrm{d}_+\phi(t,0) \otimes \mathrm{d}_+\phi(t,0), v \otimes w\rangle\rangle \\
&+ \langle\langle \mathrm{d}_+\phi(t,0) \otimes \mathrm{d}_+\psi(t,0)^*(x), v \otimes w\rangle\rangle \\
&+ \langle\langle \mathrm{d}_+\psi(t,0)(x) \otimes \mathrm{d}_+\phi(t,0), v \otimes w\rangle\rangle \\
&+ \langle\langle \mathrm{d}_+\psi(t,0)^*(x) \otimes \mathrm{d}_+\psi(t,0)^*(x), v \otimes w\rangle\rangle,
\end{aligned} \tag{2.92}$$

where we conveniently identified functionals on $\mathcal{H}$ with elements of $\mathcal{H}$.

Written in this form the right-hand side in formula (2.86) reveals its linearity in $v \otimes w$ and for $v \otimes w \in \mathcal{L}_2(\mathcal{H})$, we have

$$v \otimes w = v^+ \otimes w^+ - v^+ \otimes w^- - v^- \otimes w^+ + v^- \otimes w^-$$

and thus expanding both sides by linearity in equation (2.86), shows the validity of the formula for all $v, w \in \mathcal{H}$. ☐

Note that by inserting the formulas from (2.49)–(2.50) and (2.29)–(2.32) into the corresponding terms in (2.85) and (2.86), the latter become explicit up to the parameters (b, B, m, μ). To save some space, we give those explicit formulas only for the limit case in Proposition 2.38 below.
Using the formulas from Lemma 2.35, we establish uniform growth bounds for the semigroups $(P_t^{(k)})_{t\geq 0}$ in the next proposition. Let us note here that in general we do not obtain an uniform growth bound $w \in \mathbb{R}^+$ with $M = 1$:

Proposition 2.36. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3 and for every $k \in \mathbb{N}$ let $(P_t^{(k)})_{t\geq 0}$ be the generalized Feller semigroup on $\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$ associated with $(b, B, m^{(k)}, \mu^{(k)})$, the existence of which is guaranteed by Proposition 2.34. Then there exists a constant $w \in \mathbb{R}^+$ and $M \geq 1$, both independent of $k \in \mathbb{N}$, such that*

$$\|P_t^{(k)}\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+))} \leq M\,\mathrm{e}^{wt} \quad \textit{for all } k \in \mathbb{N}, t \geq 0. \tag{2.93}$$

Proof. Recall from Remark 2.27, that in order to show the existence of a $M \geq 1$ and $w \in \mathbb{R}^+$ such that equation (2.93) holds, it suffices to show the existence of a $\epsilon > 0$ and $C \geq 0$, independent of $k \in \mathbb{N}$, such that

$$\mathbb{E}_x\left[\rho(X_t^{(k)})\right] \leq C\rho(x)\,, \quad \forall t \in [0, \epsilon] \text{ and } x \in \mathcal{H}^+. \tag{2.94}$$

Let $k \in \mathbb{N}$ be arbitrary, but fixed and denote by $(e_n)_{n\in\mathbb{N}}$ an ONB of $\mathcal{H}$, then by Parseval's identity and monotone convergence we have:

$$\mathbb{E}_x\left[\rho(X_t^{(k)})\right] = \mathbb{E}_x\left[1 + \|X_t^{(k)}\|^2\right] = 1 + \sum_{n=1}^{\infty} \mathbb{E}_x\left[\langle X_t^{(k)}, e_n\rangle^2\right],$$

for every $t \geq 0$ and $x \in \mathcal{H}^+$. By equation (2.86), in particular using the notation in equation (2.92), we have for all $n \in \mathbb{N}$:

$$\begin{aligned}\mathbb{E}_x\left[\langle X_t^{(k)}, e_n\rangle^2\right] &= \langle\langle -\,\mathrm{d}_+^2\phi^{(k)}(t,0) - \mathrm{d}_+^2\psi^{(k)}(t,0)^*(x), e_n \otimes e_n\rangle\rangle \\ &\quad + \langle\langle \big(\mathrm{d}_+\phi(t,0) + \mathrm{d}_+\psi(t,0)^*(x)\big)^{\otimes 2}, e_n \otimes e_n\rangle\rangle. \end{aligned} \tag{2.95}$$

We show separately for the first and second terms on the right-hand side of equation (2.95) that, when summing over all $n \in \mathbb{N}$, we find a $\epsilon > 0$ and $C \geq 0$ such that equation (2.94) holds. Since

$$\sum_{n=1}^{\infty} \langle \mathrm{d}_+\phi(t,0) + \mathrm{d}_+\psi(t,0)^*(x), e_n\rangle^2 = \|\mathrm{d}_+\phi(t,0) + \mathrm{d}_+\psi(t,0)^*(x)\|^2,$$

we deduce for the second term ion the right hand side of (2.95):

$$\sum_{n=1}^{\infty} \langle\langle \big(\mathrm{d}_+\phi(t,0) + \mathrm{d}_+\psi(t,0)^*(x)\big)^{\otimes 2}, e_n \otimes e_n\rangle\rangle \leq C(t)(1 + \|x\|^2),$$

for

$$C(t) = \big(\|\mathrm{d}_+\phi(t,0)\| + \|\mathrm{d}_+\psi(t,0)^*\|_{\mathcal{L}(\mathcal{H})}\big)^2.$$

The terms $\|\mathrm{d}_+\phi(t,0)\|$ and $\|\mathrm{d}_+\psi(t,0)^*\|_{\mathcal{L}(\mathcal{H})}$ are bounded for all $t \geq 0$. Therefore, we deduce the existence of $\epsilon > 0$ and $C \geq 0$, independent of $k \in \mathbb{N}$, such that

$$\sum_{n=1}^{\infty} \langle\langle \big(\mathrm{d}_+\phi(t,0) + \mathrm{d}_+\psi(t,0)^*(x)\big)^{\otimes 2}, e_n \otimes e_n\rangle\rangle \leq C(1 + \|x\|^2), \tag{2.96}$$

for all $t \in [0,\epsilon]$ and $x \in \mathcal{H}^+$. We continue with the first term on the right hand side of (2.95). Recall formulas (2.30), (2.32), (2.43a), (2.49) and (2.50), from which we obtain:

$$\begin{aligned}&\langle\langle \mathrm{d}_+^2\psi^{(k)}(t,0)^*(x), e_n \otimes e_n\rangle\rangle \\ &\quad = -\int_0^t \int_{\mathcal{H}^+\setminus\{0\}} \langle \mathrm{e}^{s\,\mathrm{d}R(0)^*}\xi, e_n\rangle^2 \big\langle x, \mathrm{e}^{(t-s)\,\mathrm{d}R(0)}\big\rangle \frac{\mu^{(k)}(\mathrm{d}\xi)}{\|\xi\|^2}\,\mathrm{d}s,\end{aligned} \tag{2.97}$$

and

$$\begin{aligned}\langle\langle \mathrm{d}_+^2\phi^{(k)}(t,0), e_n \otimes e_n\rangle\rangle = -\int_0^t \Big(&\int_{\mathcal{H}^+\setminus\{0\}} \langle \mathrm{e}^{s\,\mathrm{d}R(0)^*}\xi, e_n\rangle^2\, m^{(k)}(\mathrm{d}\xi) \\ &+ \langle\langle \mathrm{d}_+^2\psi^{(k)}(s,0)^*(b), e_n \otimes e_n\rangle\rangle \Big)\,\mathrm{d}s \\ &+ \int_0^t \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \langle\langle \mathrm{d}_+^2\psi^{(k)}(s,0)^*(\xi), e_n \otimes e_n\rangle\rangle\, m(\mathrm{d}\xi)\,\mathrm{d}s\,.\end{aligned} \tag{2.98}$$

Hence, the two terms on the right hand side of equation (2.95) can be estimated by

$$\sum_{n=1}^{\infty}\int_0^t \int_{\mathcal{H}^+\setminus\{0\}} \langle \mathrm{e}^{s\,\mathrm{d}R(0)^*}\xi, e_n\rangle^2 \langle x, \mathrm{e}^{(t-s)\,\mathrm{d}R(0)}\rangle \frac{\mu^{(k)}(\mathrm{d}\xi)}{\|\xi\|^2}\,\mathrm{d}s$$
$$\leq \Big(\int_0^t \|\mathrm{e}^{s\,\mathrm{d}R(0)^*}\|^2_{\mathcal{L}(\mathcal{H})}\|\mathrm{e}^{(t-s)\,\mathrm{d}R(0)}\|_{\mathcal{L}(\mathcal{H})}\|\mu(\mathcal{H}^+\setminus\{0\})\|\,\mathrm{d}s\Big)\|x\|,$$

and

$$\sum_{n=1}^{\infty}\langle\langle \mathrm{d}_+^2\phi^{(k)}(t,0), e_n\otimes e_n\rangle\rangle$$
$$\leq 2\Big(\|b\| + \|\mu(\mathcal{H}^+\setminus\{0\})\| + \int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2 + \|\xi-\chi(\xi)\|m(\mathrm{d}\xi)\Big)$$
$$\times \int_0^t\int_0^s \|\mathrm{e}^{\tau\,\mathrm{d}R(0)^*}\|^2_{\mathcal{L}(\mathcal{H})}\|\mathrm{e}^{(s-\tau)\,\mathrm{d}R(0)}\|_{\mathcal{L}(\mathcal{H})}\,\mathrm{d}\tau\,\mathrm{d}s\,,$$

where we used that for all $k\in\mathbb{N}$:

$$\|\mu^{(k)}(\mathcal{H}^+\setminus\{0\})\| \leq \|\mu(\mathcal{H}^+\setminus\{0\})\| < \infty$$

and

$$\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2 + \|\xi-\chi(\xi)\|m^{(k)}(\mathrm{d}\xi) \leq \int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2 + \|\xi-\chi(\xi)\|m(\mathrm{d}\xi) < \infty.$$

Therefore, there exist $\epsilon>0$ and $\tilde{C}\geq 0$ such that

$$\sum_{n=1}^{\infty}\langle\langle -\,\mathrm{d}_+^2\phi^{(k)}(t,0) - \mathrm{d}_+^2\psi^{(k)}(t,0)^*(x), e_n\otimes e_n\rangle\rangle \leq \tilde{C}(1+\|x\|^2),$$

for all $t\in[0,\epsilon]$ and $x\in\mathcal{H}^+$. Taking the sum of the latter constant $\tilde{C}$ and the constant C found in equation (2.96) yields (2.94). □

In the next step we show that the family $(P_t)_{t\geq 0}$, defined by

$$P_t := \lim_{k\to\infty} P_t^{(k)}, \quad t\geq 0,$$

gives rise to a generalized Feller semigroup and deduce the existence of a generalized Feller process $(X_t)_{t\geq 0}$ with generator $\mathcal{G}$ as in formula (2.11).

Proposition 2.37. *Let (b, B, m, μ) be an admissible parameter set conform Definition 2.3. Then there exists a generalized Feller semigroup $(P_t)_{t\geq 0}$ on $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ such that*

$$\left(P_t\, \mathrm{e}^{-\langle \cdot, u\rangle}\right)(x) = \mathrm{e}^{-\phi(t,u)-\langle x, \psi(t,u)\rangle}, \tag{2.99}$$

for all $t \geq 0$ and $x, u \in \mathcal{H}^+$, where $(\phi(\cdot, u), \psi(\cdot, u))$ is the unique solution to the generalized Riccati equation (2.8a)-(2.8b)*. The semigroup $(P_t)_{t\geq 0}$ gives rise to a generalized Feller process $(X_t)_{t\geq 0}$ in $(\mathcal{H}^+_{\mathrm{w}}, \|\cdot\|^2 + 1)$ such that*

$$\mathbb{E}_x\left[f(X_t)\right] = P_t f(x), \quad t \geq 0, \quad x \in \mathcal{H}^+\,,$$

and the generator $\mathcal{G}$ of $(P_t)_{t\geq 0}$ is of the form in equation (2.11) *on $\mathcal{D}$.*

Proof. Hereto we check that the conditions of Theorem 3.2 in [45] hold. From Proposition 2.36, we know that the sequence of semigroups $(P_t^{(k)})_{t\geq 0, k\in\mathbb{N}}$ with generators $(\mathcal{G}^{(k)})_{k\in\mathbb{N}}$ satisfy the following growth bound

$$\|P_t^{(k)}\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}}))} \leq M\, \mathrm{e}^{wt}, \qquad \forall\, n \in \mathbb{N} \quad \text{and } t \geq 0\,, \tag{2.100}$$

where $w \in \mathbb{R}$.
Recall the definition of $\mathcal{D}$ from equation (2.57) and recall from Lemma 2.28 that $\mathcal{D}$ is a dense subspace of $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$. Thus (i) in Theorem 3.2 in [45] is satisfied.
Note that the operator $\mathcal{G}^{(n)}$, $n \in \mathbb{N}$, applied to the function $\mathrm{e}^{-\phi^{(k)}(s,u)-\langle \cdot, \psi^{(k)}(s,u)\rangle}$, with $(\phi^{(k)}(\cdot, u), \psi^{(k)}(\cdot, u))$ being a solution to (2.14), gives (see also (2.84))

$$\begin{aligned}
&\left(\mathcal{G}^{(n)}\, \mathrm{e}^{-\phi^{(k)}(s,u)-\langle \cdot, \psi^{(k)}(s,u)\rangle}\right)(x) \\
&\quad = \mathrm{e}^{-\phi^{(k)}(s,u)} \mathcal{G}^{(n)}\, \mathrm{e}^{-\langle \cdot, \psi^{(k)}(s,u)\rangle}(x) \\
&\quad = \left(-F^{(n)}(\psi^{(k)}(s,u)) - \langle R^{(n)}(\psi^{(k)}(s,u)), x\rangle\right) \mathrm{e}^{-\phi^{(k)}(s,u)-\langle x, \psi^{(k)}(s,u)\rangle}\,,
\end{aligned}$$

for $x, u \in \mathcal{H}^+$, $s \geq 0$ From the latter and equation (2.75), we infer

$$\begin{aligned}
&\frac{1}{\|x\|^2+1}\left|\mathcal{G}^{(n)} P_s^{(k)}\, \mathrm{e}^{-\langle \cdot, u\rangle}(x) - \mathcal{G}^{(k)} P_s^{(k)}\, \mathrm{e}^{-\langle \cdot, u\rangle}(x)\right| \\
&\leq \frac{\mathrm{e}^{-\phi^{(k)}(s,u)-\langle x, \psi^{(k)}(s,u)\rangle}}{\|x\|^2+1}\left[b_{s,u}^{(n,k)} + \|x\| a_{s,u}^{(n,k)}\right]\,,
\end{aligned} \tag{2.101}$$

where

$$\begin{aligned}
a_{s,u}^{(n,k)} &:= \left\|R^{(n)}(\psi^{(k)}(s,u)) - R^{(k)}(\psi^{(k)}(s,u))\right\|, \\
b_{s,u}^{(n,k)} &:= \left|F^{(n)}(\psi^{(k)}(s,u)) - F^{(k)}(\psi^{(k)}(s,u))\right|.
\end{aligned}$$

From the equations (2.5) and (2.26) we have, for all $0 \leq s \leq T < \infty$:

$$\begin{aligned}
&\Big|\Big(\mathrm{e}^{-\langle \xi, \psi^{(k)}(s,u)\rangle} - 1 - \langle \chi(\xi), \psi^{(k)}(s,u)\rangle\Big)\big(\mathbf{1}_{\{\|\xi\|>1/n\}} - \mathbf{1}_{\{\|\xi\|>1/k\}}\big)\Big| \\
&\quad \leq \|\psi^{(k)}(s,u)\|^2\|\xi\|^2\mathbf{1}_{\{\|\xi\|\leq 1\}} \\
&\quad \leq \sup_{s\in[0,T]} \|\psi^{(1)}(s,u)\|^2\|\xi\|^2\mathbf{1}_{\{\|\xi\|\leq 1\}} =: g(\xi)\,.
\end{aligned}$$

Observe that for $h \in \mathcal{H}^+$, we have $\int_{\mathcal{H}^+\setminus\{0\}} g(\xi)\,\frac{\langle \mu(\mathrm{d}\xi), h\rangle}{\|\xi\|^2} < \infty$. Hence, it follows from Lemma A.6 that $g(\cdot)/\|\cdot\|^2 \in \mathcal{L}^1(\mathcal{H}^+,\mu)$ and from Theorem A.8, we deduce that $\sup_{s\in[0,T]} a_{s,u}^{(n,k)}$ converges to 0 as $n,k \to \infty$. By the admissibility condition Definition 2.3 i), we infer $\int_{\mathcal{H}^+\setminus\{0\}} g(\xi)\, m(\mathrm{d}\xi) < \infty$ and applying the dominated convergence theorem we also deduce that $\sup_{s\in[0,T]} b_{s,u}^{(n,k)}$ converge to 0 as $n,k \to \infty$. Observing that $\phi^{(k)}(s,u) \in \mathbb{R}^+$ and $\psi^{(k)}(s,u) \in \mathcal{H}^+$ for all $s \geq 0$, we can bound $\mathrm{e}^{-\phi^{(k)}(s,u)-\langle x, \psi^{(k)}(s,u)\rangle}$ by 1 for all $x \in \mathcal{H}^+$ and get from equation (2.101), that for all $s > 0$:

$$\begin{aligned}
&\Big\|\mathcal{G}^{(n)}P_s^{(k)}\,\mathrm{e}^{-\langle\cdot,u\rangle} - \mathcal{G}^{(k)}P_s^{(k)}\,\mathrm{e}^{-\langle\cdot,u\rangle}\Big\|_\rho \\
&\quad \leq \sup_{x\in\mathcal{H}^+} \frac{\|x\|+1}{\|x\|^2+1}\Big(a_{s,u}^{(n,k)} + b_{s,u}^{(n,k)}\Big) \\
&\quad \leq \left(\sup_{s\in[0,T]} a_{s,u}^{(n,k)} + \sup_{s\in[0,T]} b_{s,u}^{(n,k)}\right) C_u\|\mathrm{e}^{-\langle\cdot,u\rangle}\|_\infty\,,
\end{aligned}$$

where $C_u = \sup_{x\in\mathcal{H}^+}(\|x\|+1)/(\|x\|^2+1)$. Thus condition (ii) in Theorem 3.2 in [45] is satisfied with $\|\cdot\|_{\mathcal{D}} = \|\cdot\|_\infty$ and we deduce the existence of a generalized Feller semigroup $(P_t)_{t\geq 0}$ with the same growth bound as the semigroup $(P_t^{(k)})_{t\geq 0}$ and such that $P_t f = \lim_{k\to\infty} P_t^{(k)} f$, for all $f \in \mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$, uniformly on compacts in time. Since $P_t\mathbf{1} = \mathbf{1}$, for all $t \geq 0$, we deduce from Theorem 2.26 that there exists a generalized Feller process $(X_t)_{t\geq 0}$ such that $P_t f(x) = \mathbb{E}_x[f(X_t)]$ for all $t \geq 0$ and $x \in \mathcal{H}^+$. The exponential affine formula (2.99) follows from formula (2.75) and the fact that $\lim_{k\to\infty}\phi^{(k)}(t,u) = \phi(t,u)$ and $\lim_{k\to\infty}\psi^{(k)}(t,u) = \psi(t,u)$ for all $t \geq 0$ and $u \in \mathcal{H}^+$. From this we further derive the particular form of the generator $\mathcal{G}$ on the space $\mathcal{D}$ by noting that $t \mapsto P_t\,\mathrm{e}^{-\langle\cdot,u\rangle}(x)$ uniquely solves the abstract Cauchy problem associated to $(\mathcal{G}, \mathrm{dom}(\mathcal{G}))$ and hence by mimicking the proof of the approximation case in Proposition 2.34, we conclude formula (2.11). □

Analogous to the approximating processes $(X_t^{(k)})_{t\geq 0}$, for $k \in \mathbb{N}$ in Lemma 2.35, we now deduce explicit formulas for the expressions $\mathbb{E}_x[\langle X_t, v\rangle]$ as well as for $\mathbb{E}_x\big[\langle X_t, v\rangle^2\big]$, where $x \in \mathcal{H}^+$, $t \geq 0$ and $v \in \mathcal{H}^+$.

Proposition 2.38. *Let* (b, B, m, μ) *be an admissible parameter set conform Definition 2.3 and recall from* (2.29)–(2.32) *the definition of* $\mathrm{d}R(0), \mathrm{d}^2R(0)$, $\mathrm{d}F(0)$, *and* $\mathrm{d}^2F(0)$. *Then for all* $v, w \in \mathcal{H}^+$ *the following formulas hold true:*

$$\mathbb{E}_x\left[\langle X_t, v\rangle\right] = \int_0^t \langle b, \mathrm{e}^{s\,\mathrm{d}R(0)}v\rangle + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \langle \xi, \mathrm{e}^{s\,\mathrm{d}R(0)}v\rangle\, m(\mathrm{d}\xi)\,\mathrm{d}s + \langle x, \mathrm{e}^{t\,\mathrm{d}R(0)}v\rangle \tag{2.102}$$

and

$$\begin{aligned}
&\mathbb{E}_x\left[\langle X_t, v\rangle\langle X_t, w\rangle\right]\\
&\quad = -\int_0^t \mathrm{d}^2F(0)(\mathrm{e}^{s\,\mathrm{d}R(0)}v, \mathrm{e}^{s\,\mathrm{d}R(0)}w)\ \mathrm{d}s\\
&\qquad - \int_0^t\int_0^s \mathrm{d}F(0)\left(\mathrm{e}^{(s-u)\,\mathrm{d}R(0)}\,\mathrm{d}^2R(0)(\mathrm{e}^{u\,\mathrm{d}R(0)}v, \mathrm{e}^{u\,\mathrm{d}R(0)}w)\right)\ \mathrm{d}u\ \mathrm{d}s\\
&\qquad - \int_0^t \left\langle x, \mathrm{e}^{(t-s)\,\mathrm{d}R(0)}\,\mathrm{d}^2R(0)(\mathrm{e}^{s\,\mathrm{d}R(0)}v, \mathrm{e}^{s\,\mathrm{d}R(0)}w)\right\rangle\ \mathrm{d}s\\
&\qquad + \left(\int_0^t \mathrm{d}F(0)(\mathrm{e}^{s\,\mathrm{d}R(0)}v)\ \mathrm{d}s + \left\langle x, \mathrm{e}^{t\,\mathrm{d}R(0)}v\right\rangle\right)\\
&\qquad\quad \times\left(\int_0^t \mathrm{d}F(0)(\mathrm{e}^{s\,\mathrm{d}R(0)}w)\ \mathrm{d}s + \left\langle x, \mathrm{e}^{t\,\mathrm{d}R(0)}w\right\rangle\right).
\end{aligned} \tag{2.103}$$

Moreover, for $v \in \mathcal{H}^+$, $\langle\cdot, v\rangle \in \mathrm{dom}(\mathcal{G})$ *and*

$$\mathcal{G}\langle\cdot, v\rangle(x) = \langle b + B(x), v\rangle + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \langle \xi, v\rangle\,\nu(x, \mathrm{d}\xi)\,, \qquad x \in \mathcal{H}^+\,. \tag{2.104}$$

Proof. Formulas (2.102) and (2.103) can be obtained analogous to the computation of the formulas (2.85) and (2.86) derived for the approximating case, combined with the explicit formulas (2.49)–(2.50). As in the proof of Lemma 2.35 we use Proposition 2.20 and the finite second moments of the process $(X_t)_{t\geq 0}$ to interchange the operations of the expectation and the one-sided derivatives. To obtain more explicit formulas, we consider the analogous of the formulas (2.85) and (2.86) and recall that $\mathrm{d}_+\phi(t,0)(v)$, $\mathrm{d}_+^2\phi(t,0)(v,v)$ can be expressed in terms of $\mathrm{d}F(0)$, $\mathrm{d}^2F(0)$, $\mathrm{d}_+\psi(t,0)(v)$, and $\mathrm{d}_+^2\psi(t,0)(v,w)$, see (2.42a) and (2.43a). Then, we recall the expressions (2.49) and (2.50) for $\mathrm{d}_+\psi(t,0)(v)$, and $\mathrm{d}_+^2\psi(t,0)(v,w)$.

To prove (2.104), observe that using the analogue of (2.85), we get

$$\begin{aligned}
&\frac{1}{t}\Big|P_t\langle\cdot,v\rangle(x)-\langle x,v\rangle-\langle b+B(x),v\rangle-\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\langle\xi,v\rangle\,\nu(x,\mathrm{d}\xi)\Big|\\
&\quad\leq\frac{1}{t}\Big|\,\mathrm{d}_+\phi(t,0)(v)-\langle b,v\rangle-\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\langle\xi,v\rangle\,m(\mathrm{d}\xi)\Big|\\
&\qquad+\frac{1}{t}\|x\|\Big\|\,\mathrm{d}_+\psi(t,0)(v)-v-B^*(v)-\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\langle\xi,v\rangle\,\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big\|.
\end{aligned}$$

The latter together with formulas (2.42a) and (2.42b), yield

$$\begin{aligned}
&\lim_{t\to0^+}\sup_{x\in\mathcal{H}^+}\frac{\frac{1}{t}\Big|P_t\langle\cdot,v\rangle(x)-\langle x,v\rangle-\langle b+B(x),v\rangle-\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\langle\xi,v\rangle\,\nu(x,\mathrm{d}\xi)\Big|}{1+\|x\|^2}\\
&\quad\leq\Big|\,\mathrm{d}F(0)(v)-\langle b,v\rangle-\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\langle\xi,v\rangle\,m(\mathrm{d}\xi)\Big|\\
&\qquad+\Big\|\,\mathrm{d}R(0)(v)-B^*(v)-\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\langle\xi,v\rangle\,\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big\|
\end{aligned}$$

and recalling the formulas for $\mathrm{d}R(0)$ and $\mathrm{d}F(0)$ respectively in (2.29) and (2.30), we conclude that $\langle\cdot,v\rangle\in\mathrm{dom}(\mathcal{G})$, for $v\in\mathcal{H}^+$ and that (2.104) holds. □

Remark 2.39. As observed in Remark 2.4, the second moment conditions are a consequence of our generalized Feller approach and weight function $\rho=\|\cdot\|^2+1$. More specifically, the uniform bounds established in Proposition 2.36 rely on the existence of second moments as established in Lemma 2.35. A natural question to ask is whether one could perform the analysis with a different (weaker) weight function. However, in the proof of Lemma 2.32 we need that $\sqrt{\rho(x)}\geq c\|x\|$, $x\in\mathcal{H}^+$, for some constant $c\in(0,\infty)$.
Note also that the existence of a *first* moment of $\|\xi\|^{-2}\mu(\mathrm{d}\xi)$ is already used in Lemma 2.12 to ensure that the mappings $R^{(k)}$ are Lipschitz continuous. Interestingly, also the alternative approach to proving existence in Chapter 6 requires the second moment condition.

In general we do not obtain a version of the process X in Proposition 2.37 with càdlàg paths (but see Chapter 6 for a positive result). By Theorem 2.13 in [45] a càdlàg version exists when the associated semigroup $(P_t)_{t\geq0}$ is quasi-contractive on $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$, i.e., if one can take $M=1$ in Proposition 2.36. We do not know whether this holds in general. However, we can show that X admits a càdlàg version in the finite activity setting:

Proposition 2.40. *Assume the setting of Proposition 2.37 and assume moreover that $m(\mathcal{H}^+\setminus\{0\})<\infty$ and that $\mathcal{H}^+\setminus\{0\}\ni\xi\mapsto\|\xi\|^{-2}$ is μ-integrable. Then there exists a version of X with càdlàg paths.*

Proof. By [45, Theorem 2.13] it in fact suffices to prove that the generalized Feller semigroup $(P_t)_{t\geq 0}$ associated to X is quasi-contractive on $\mathcal{B}_{\tilde{\rho}}(\mathcal{H}^+_{\mathrm{w}})$, where $\tilde{\rho}\colon \mathcal{H}^+ \to [0,\infty)$ is an admissible weight function such that its associated norm $\|\cdot\|_{\tilde{\rho}}$ is equivalent to $\|\cdot\|_{\rho}$. Note that in the finite activity setting we can apply Proposition 2.34 with $k=\infty$ (with the understanding that $m^{(\infty)} := m$ and $\mu^{(\infty)} := \mu$) to directly obtain $(P_t)_{t\geq 0}$ (i.e., no approximation over k is necessary). In particular $\tilde{\omega}_\infty < \infty$, where $\tilde{\omega}_\infty$ is defined by taking $k=\infty$ in (2.81). It then follows from statement a) on page 44 that $(P_t)_{t\geq 0}$ is quasi-contractive on $\mathcal{B}_{\tilde{\rho}_\infty}(\mathcal{H}^+_{\mathrm{w}})$ where $\tilde{\rho}_\infty$ is an admissible weight function with associated norm equivalent to $\|\cdot\|_{\rho}$. $\square$

In the next section we give the proof of Theorem 2.8. This is based on collecting the results from this section and transferring from a generalized Feller setting to the classical setting (in particular transferring from the weak to the strong topology on $\mathcal{H}^+$) that we used for presenting the results in Section 2.2.

2.4.3 Proof of Theorem 2.8

Let (b,B,m,μ) be an admissible parameter set. Then by Proposition 2.37 there exists a generalized Feller semigroup $(P_t)_{t\geq 0}$ and the associated generalized Feller process $(X_t)_{t\geq 0}$ in $\mathcal{H}^+$ such that $\mathbb{E}_x\left[f(X_t)\right] = P_t f(x)$ for $t\geq 0$ and the Markov property (2.54) holds. The existence of constants $M,\omega\in[1,\infty)$ such that (2.9) is satisfied follows from Remark 2.27. The space $\mathcal{H}$ is a separable Hilbert space and hence the Borel-σ-algebras $\mathcal{B}(\mathcal{H}^+)$ and $\mathcal{B}(\mathcal{H}^+_{\mathrm{w}})$ coincide. This means that the transition kernels $(p_t(x,\cdot))_{t\geq 0}$ defining the semigroup $(P_t)_{t\geq 0}$ stay unaffected under the change of topology and hence the process $(X_t)_{t\geq 0}$ is also a Markov process in $\mathcal{H}^+$ with the strong topology. This proves the first part of the assertion and we continue with the affine transform formula.
Indeed, the asserted affine transform formula (2.2) is precisely formula (2.99) from Proposition 2.37. This and Proposition 2.16 implies for all $x\in\mathcal{H}^+$ that

$$\begin{aligned}\lim_{t\to 0+}\frac{P_t\,\mathrm{e}^{-\langle\cdot,u\rangle}(x)-\mathrm{e}^{-\langle\cdot,u\rangle}(x)}{t} &= \lim_{t\to 0+}\frac{\mathrm{e}^{-\phi(t,u)-\langle x,\psi(t,u)\rangle}-\mathrm{e}^{-\langle x,u\rangle}}{t}\\ &= \left(-F(u)-\langle x,R(u)\rangle\right)\mathrm{e}^{-\langle x,u\rangle}. \end{aligned} \tag{2.105}$$

In particular, we see that $\mathcal{G}(\mathcal{D})\subseteq C_b(\mathcal{H}^+)$ and since $(P_t)_{t\geq 0}$ is a strongly continuous semigroup on $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ we have

$$\left(P_t\,\mathrm{e}^{-\langle\cdot,u\rangle}\right)(x) = \mathrm{e}^{-\langle\cdot,x\rangle}(x) + \int_0^t \left(P_s\mathcal{G}\,\mathrm{e}^{-\langle\cdot,u\rangle}\right)(x)\,\mathrm{d}s.$$

Consequently, we have shown that $\mathcal{D}\subseteq\operatorname{dom}(\mathcal{G})$ and from formula (2.105) we see that formula (2.11) holds true on $\mathcal{D}$.

2.5 Concluding remarks

With Theorem 2.8 we have proven the existence of affine Markov processes on the cone of positive self-adjoint Hilbert-Schmidt operators by a novel approach inspired by [45]. In particular, our approach relies on the theory of generalized Feller processes, taking the weight function $\rho = \|\cdot\|^2 + 1$. This approach requires the existence of first and second moments of the jump measures m and μ. A beneficial by-product is that we obtain explicit formulas for the first and second moments of the affine Markov process, see Proposition 2.38. See Remark 2.39 for a discussion regarding the necessity of the second-moment condition. Below, we make some more remarks, including possible further directions of research.

- ***On relaxing the condition on existence of moments***

A possible direction of further research is to investigate whether one can adapt the proof in such a way to allow for the weight function $\rho = \|\cdot\| + 1$. In this case a first moment conditions on m and μ should suffice. On a more abstract level, the question arises whether it is possible to establish existence without any moment conditions, as can be done in the finite dimensional setting where the cone of interest does not have empty interior. Note, however, that for instantaneous covariance processes one usually assumes the existence of second moments, i.e. for applications in stochastic covariance modeling the second moment assumptions on the affine processes are not a real restriction, but often required anyways.

- ***On the inclusion of a diffusion part***

Another tantalizing question is to what degree an infinite-dimensional affine process on the cone of positive self-adjoint Hilbert-Schmidt operators allows for diffusion. It is clear from [23] that certain constructions are possible in infinite-dimensions. See also [99] for an process with affine diffusion part on the cone of positive continuous functions, which, however, has non-empty interior.

- ***On considering a different state space for the covariance process***

Note that in the stochastic covariance model (Y, X) in (1.6), we take X to be the *square root* of an affine process in order to obtain that Y is again affine, see Chapter 3. However, this means that the 'natural' state space for X is not the cone of positive self-adjoint Hilbert-Schmidt operators, but the cone of positive self-adjoint *trace class* operators. Unfortunately, this is no longer a cone in a Hilbert space. As self-duality of the cone was used at various instances in the proof of Theorem 2.8, it is not clear how much can be salvaged if we consider trace class operators. This would be a further interesting direction of research.

CHAPTER 3

An Infinite-Dimensional Affine Stochastic Covariance Model

Abstract of the chapter In this chapter we present a flexible and tractable affine stochastic covariance model in infinite-dimensions. More specifically, we consider a Hilbert space valued linear stochastic differential equation driven by an infinite-dimensional Brownian motion that is modulated by an affine pure-jump process with values in the cone of positive self-adjoint Hilbert-Schmidt operators. The tractability of our model lies in the fact that the two processes involved are jointly *affine*, i.e. we show that their characteristic function can be given quasi-explicitly in terms of the solutions to a set of generalized mild Riccati equations. The flexibility of the model lies in the fact that we allow multiple modeling options for the instantaneous covariance process, including processes with state-dependent jump intensity. We discuss applications of our model in the context of forward curve dynamics described in the HJMM modeling framework. In this setting we discuss various examples: An infinite-dimensional version of the Barndorff-Nielsen–Shephard stochastic volatility model, as well as models with instantaneous covariance processes admitting state-dependent jump-intensity, that have the potential to model *volatility clustering*.

This chapter is based on [39]:
Cox, S., Karbach, S., and Khedher, A.
An infinite-dimensional affine stochastic volatility model.
Mathematical Finance 32, 3 (2022), 878–906.

3.1 Introduction

In this chapter we propose a novel class of infinite-dimensional *affine* stochastic covariance models (Y, X), where $Y = (Y_t)_{t\geq 0}$ takes values in an infinite-dimensional separable Hilbert space $(H, \langle\cdot,\cdot\rangle_H)$ and $X = (X_t)_{t\geq 0}$ is a time-homogeneous *affine* Markov process with values in the cone of positive self-adjoint Hilbert-Schmidt operators on H. The process X is taken from the class of affine processes introduced in Chapter 2. The process $(Y_t)_{t\geq 0}$ is modeled by the following stochastic differential equation

$$\mathrm{d}Y_t = \mathcal{A}Y_t \,\mathrm{d}t + X_t^{1/2}\,\mathrm{d}W_t^Q, \quad t \geq 0, \quad Y_0 = y \in H\,, \tag{3.1}$$

where $\mathcal{A}\colon \mathrm{dom}(\mathcal{A}) \subseteq H \to H$ is a possibly unbounded operator with dense domain $\mathrm{dom}(\mathcal{A})$ and $(W_t^Q)_{t\geq 0}$ is a Q-Brownian motion independent of X, with Q a positive self-adjoint trace-class operator on H. Assuming that X is progressively measurable and using the moment bounds on X established in Chapter 2, the existence of a solution to (3.1) is straightforward (see Lemma 3.7 below).
In Section 3.2.1 we show that under the assumption that the Markov process $(X_t)_{t\geq 0}$ has càdlàg paths, it is a square-integrable semimartingale. This follows from the formulation of an associated martingale problem in terms of what we call *a weak* generator (see Definition 3.2) of the Markov process $(X_t)_{t\geq 0}$ and yields the explicit representation of $(X_t)_{t\geq 0}$ as

$$X_t = x + \int_0^t \Big(b + B(X_s) + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \xi\,\nu(X_s,\mathrm{d}\xi)\Big)\,\mathrm{d}s + \bar{J}_t, \quad t \geq 0, \tag{3.2}$$

where $x, b \in \mathcal{H}^+$, $B \in \mathcal{L}(\mathcal{H})$ is a bounded linear operator, given $x \in \mathcal{H}^+$ the measure $\nu(x,\cdot)\colon \mathcal{B}(\mathcal{H}^+ \setminus \{0\}) \to \mathbb{R}$ is such that $\nu^X(\mathrm{d}t,\mathrm{d}\xi) = \nu(X_t,\mathrm{d}\xi)\,\mathrm{d}t$ is the predictable compensator of the jump-measure of $(X_t)_{t\geq 0}$, and $(\bar{J}_t)_{t\geq 0}$ is a purely discontinuous $\mathcal{H}$-valued square-integrable martingale. Here and throughout this chapter the parameters (b, B, m, μ) are admissible as in Definition 2.3.
By exploiting the results in Chapter 2 and [117], we adapt the proof of [80, Theorem II.2.42] to our infinite-dimensional setting to obtain the characteristic triplet (see Definition 3.1) of $(X_t)_{t\geq 0}$ explicitly and show its affine form (see Proposition 3.5).
Our *main contribution* lies in showing that our stochastic covariance model (Y, X) has the affine property, that is, we prove for all $t \geq 0$ that the mixed Fourier-Laplace transform of (Y_t, X_t) is exponentially affine in the initial value $(y, x) \in H \times \mathcal{H}^+$ and has a quasi-explicit formula in terms of a solution to generalized (mild) Riccati equations that are written in terms of the parameters of the model, see Theorem 3.14 below.

The proof Theorem 3.14, i.e., of the affine property of our stochastic covariance model $(Y_t, X_t)_{t\geq 0}$, is given in Section 3.3. It involves considering an approximation $(Y_t^{(n)}, X_t)_{t\geq 0}$ of $(Y_t, X_t)_{t\geq 0}$ obtained by replacing $\mathcal{A}$ in (3.1) by its Yosida approximation. The use of the approximation allows us to exploit the semimartingale theory and standard techniques in order to show that the approximating process is affine. To show that the affine property holds for the limiting process $(Y_t, X_t)_{t\geq 0}$, we study the convergence of the generalized Riccati equations associated with $(Y_t^{(n)}, X_t)_{t\geq 0}$ to those associated with $(Y_t, X_t)_{t\geq 0}$. We prove the existence of a unique solution to these generalized Riccati equations by exploiting infinite dimensional ODE results and using the quasi-monotonicity argument to show that the solution stays in the cone $\mathcal{H}^+$, see [50] and [113]. In order for the approach described above to succeed, we impose a commutativity-type condition on the covariance operator of the Q-Wiener process $(W_t^Q)_{t\geq 0}$ and the stochastic covariance $(X_t^{1/2})_{t\geq 0}$ (see Assumption $\mathcal{B}$ below). This condition is also imposed in [21] and we find it rather limiting. However, we show that it can be avoided by considering a slightly different version of the stochastic covariance model, see Remark 3.10.

In Section 3.4 we consider a number of examples. For the process Y we assume the setting proposed in [60, 18], which can be used to model arbitrage-free forward prices at time $t \geq 0$ of a contract delivering an asset, e.g. a commodity, at time $t + x$. In this case the operator $\mathcal{A}$ in (3.1) is given by $\mathcal{A} = \partial/\partial x$ and the space H is given by some weighted Sobolev space. For the process $(X_t)_{t\geq 0}$, we construct several examples in which we specify the drift and jump parameters: We first show that the infinite-dimensional lift of the multivariate Barndorff-Nielsen–Shephard model introduced in [21] is a particular example of our model. The instantaneous covariance process $(X_t)_{t\geq 0}$ in this example is given as the solution to a SDE driven by a *Lévy subordinator* in the space of self-adjoint Hilbert-Schmidt operators, as we show in Section 3.4.1. As mentioned above, this example does not involve state-dependent jump intensities. However, Sections 3.4.2, 3.4.3, and 3.4.4 provide explicit parameter choices that *do* involve state-dependent jump intensities. In Section 3.4.2 we construct a covariance process which is essentially one-dimensional; evolving along a fixed vector $z \in \mathcal{H}^+$. In Section 3.4.3, we construct a truly infinite-dimensional covariance process X. In this example both X_t, $t \geq 0$, and Q share a fixed orthonormal basis of eigenvectors. This is imposed to ensure that the commutativity condition given by Assumption $\mathcal{B}$ is satisfied. In Section 3.4.4, we avoid this commutativity condition by considering an example involving the alternative model from Remark 3.10.

In Chapter 4 we derive quasi-explicit formulas for option-prices on commodity forwards based on the introduced affine model. In practice, these computations require the study of finite dimensional approximations of the instantaneous covariance process and the Riccati equations, which is being tackled in Chapter 6.

3.1.1 Layout of the chapter

In Section 3.2 we give an in-depth analysis of our stochastic covariance model and introduce sufficient parameter assumptions that ensure the well-posedness of the proposed model. Subsequently, in Section 3.3 we prove the affine-property of our joint model $(Y_t, X_t)_{t\geq 0}$. We split the proof into two parts, first in Section 3.3.1 we show the existence and uniqueness of solutions to the associated generalized Riccati equations under admissible parameter assumptions and thereafter, in Section 3.3.2, we prove the affine transform formula. In Section 3.4, we give several examples of stochastic covariance models included in our model class by specifying various instantaneous covariance processes $(X_t)_{t\geq 0}$.

3.1.2 Notation and Hilbert-valued semimartingales

Throughout this chapter we let $(H, \langle\cdot,\cdot\rangle_H)$ be an infinite-dimensional separable Hilbert space and let $\chi\colon \mathcal{H} \to \mathcal{H}$ be the truncation function given by $\chi(x) = x\mathbf{1}_{\{\|x\|\leq 1\}}$. By slight abuse of notation we shall write $\langle y, u\rangle_H = -\mathrm{i}\langle y, \tilde{u}\rangle_H$, in case that $y \in H$ and $u = \mathrm{i}\tilde{u} \in \mathrm{i}H$, see also the notation section in Chapter 4.
For this section let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space, write $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ and denote by $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ another separable Hilbert space. Let $M = (M_t)_{t\geq 0}$ be an $\mathcal{H}$-valued locally square-integrable martingale. Then we know from [117, Theorem 21.6 and Section 23.3] that there exists a unique (up to a $\mathbb{P}$-null set) càdlàg predictable process $\langle\langle M\rangle\rangle$ of finite variation taking values in the set of positive self-adjoint elements of $\mathcal{L}_1(\mathcal{H})$ such that $\langle\langle M\rangle\rangle_0 = 0$ and the process

$$M \otimes M - \langle\langle M\rangle\rangle$$

is an $\mathcal{L}_1(\mathcal{H})$-valued local martingale.
Following [117, Definition 23.7], an $\mathcal{H}$-valued process $X = (X_t)_{t\geq 0}$ is called *a semimartingale* if

$$X_t = X_0 + M_t + A_t, \qquad t \geq 0, \tag{3.3}$$

where X_0 is $\mathcal{H}$-valued and $\mathcal{F}_0$-measurable, M is a $\mathcal{H}$-valued locally square integrable martingale with càdlàg paths such that $M_0 = 0$ and A is an adapted $\mathcal{H}$-valued càdlàg process of finite variation with $A_0 = 0$. When the process A in (3.3) is predictable, then X is said to be *a special semimartingale*. The decomposition (3.3) in this case is unique (see [117, Theorem 23.6]) and is called *the canonical decomposition* of X. For a semimartingale X, we write $\Delta X_t = X_t - X_{t-}$, where $X_{t-} = \lim_{s\to t-} X_s$. Notice that when $\|\Delta X\|$ is bounded, then X is a special semimartingale (see [117, Chapter 4, Exercise 11]).
Two $\mathcal{H}$-valued locally square-integrable martingales M and N are called *orthogonal* if the real-valued process $(\langle M_t, N_t\rangle)_{t\geq 0}$ is a local martingale.

Further, we call M a *purely discontinuous* local martingale if it is orthogonal to all continuous local martingales. An $\mathcal{H}$-valued semimartingale can be written as

$$X_t = X_0 + X_t^c + M_t^d + A_t, \quad t \geq 0, \tag{3.4}$$

where X_0 is $\mathcal{F}_0$-measurable, X^c is a continuous local martingale with $X_0^c = 0$, M^d is a locally square integrable martingale orthogonal to X^c with $M_0^d = 0$, and A is a càdlàg process of finite variation with $A_0 = 0$, see [117, Theorem 20.2]. The process X^c is unique (up to a $\mathbb{P}$-null set), see [117, Chapter 4, Exercise 13]. We associate with the $\mathcal{H}$-valued semimartingale X, the integer-valued random measure $\mu^X : \mathcal{B}([0,\infty) \times \mathcal{H}) \to \mathbb{N}$ given by

$$\mu^X(\mathrm{d}t, \mathrm{d}\xi) = \sum_{s \geq 0} \mathbf{1}_{\{\Delta X_s \neq 0\}} \delta_{(s, \Delta X_s)}(\mathrm{d}t, \mathrm{d}\xi), \tag{3.5}$$

where δ_a denotes the Dirac measure at point a. Recall from [80, Theorem II.1.8], the existence and uniqueness (up to a $\mathbb{P}$-null set) of *the predictable compensator* ν^X of μ^X.
Given a semimartingale X we define the 'large jumps' process $\check{X}$ by

$$\check{X} := \sum_{s \leq \cdot} \Delta X_s \mathbf{1}_{\{\|\Delta X_s\| > 1\}},$$

and we define the 'small jumps' process

$$\hat{X} = X - \check{X}. \tag{3.6}$$

Since $\|\Delta \hat{X}\| \leq 1$, $\hat{X}$ is a special semimartingale and hence it admits the unique decomposition

$$\hat{X}_t = X_0 + M_t^{\hat{X}} + A_t^{\hat{X}}, \quad t \geq 0, \tag{3.7}$$

where X_0 is $\mathcal{F}_0$-measurable, $M^{\hat{X}}$ is a local martingale with $M_0^{\hat{X}} = 0$, and $A^{\hat{X}}$ is a predictable process of finite variation with $A_0^{\hat{X}} = 0$. We are ready to introduce *the characteristic triplet* of an $\mathcal{H}$-valued semimartingale X:

Definition 3.1. Let X be an $\mathcal{H}$-valued semimartingale, let $A^{\hat{X}}$ be the predictable process of finite variation from decomposition (3.7), let X^c be the continuous martingale part of X as provided by (3.4), and let ν^X be the predictable compensator of μ^X, where μ^X is defined by (3.5). Then we call the triplet $(A^{\hat{X}}, \langle\langle X^c \rangle\rangle, \nu^X)$ the *characteristic triplet* of X. Note that the characteristic triplet consists of a predictable càdlàg $\mathcal{H}$-valued process of finite variation, a predictable càdlàg $\mathcal{L}_1(\mathcal{H})$-valued process of finite variation, and a predictable random measure on $\mathcal{B}([0,\infty) \times \mathcal{H})$.

3.2 The stochastic covariance model

In this section we specify our stochastic covariance model. First, in Section 3.2.1, we introduce the exact specification for the instantaneous covariance process X, for which we choose an affine Markov process on the cone of positive self-adjoint Hilbert-Schmidt operators, the existence of which we proved in Chapter 2. We show that whenever the process X admits for a version with càdlàg paths, this version is actually a Markovian semimartingale with characteristic triplet of an affine form and the representation (3.2) holds true. In Section 3.2.2, we show that given such an instantaneous covariance process X there exists a mild solution Y to equation (3.1) with initial value $y \in H$, which enables us to introduce our joint stochastic covariance model $Z = (Y, X)$ (see Definition 3.8).

3.2.1 The affine covariance process

We propose to model the instantaneous covariance process $(X_t)_{t\geq 0}$ as an *affine* Markov process on the state space $\mathcal{H}^+$ in the sense of Chapter 2. More precisely, we assume that (b, B, m, μ) is an admissible parameter set as in Definition 2.3. Given (b, B, m, μ) it follows from Theorem 2.8 that there exists an associated square-integrable time-homogeneous $\mathcal{H}^+$-valued affine Markov process X. More specifically, Theorem 2.8 and Proposition 2.38 imply Theorem 3.3 below, a version of the existence results that we need in our derivations later and which we recall for the readers convenience. In order to state this result we adapt the notion of a *weak generator*[1] from Section 2.2 slightly and let the state-space of test-functions be the space of continuous functions on $\mathcal{H}^+$ with at most quadratic growth (in the tails), in particular we allow for unbounded functions in the domain.

Definition 3.2 (Weak generator on functions with at most quadratic-growth)**.** Let X be a square-integrable $\mathcal{H}^+$-valued time-homogeneous Markov process with transition semigroup $(P_t)_{t\geq 0}$ acting on the space

$$C_{\mathrm{w}}(\mathcal{H}^+,\mathbb{R}) := \Big\{ f \in C(\mathcal{H}^+,\mathbb{R})\colon \sup_{x\in\mathcal{H}^+} \frac{f(x)}{\|x\|^2+1} < \infty \Big\}.$$

$C_{\mathrm{w}}(\mathcal{H}^+,\mathbb{R})$ Then the *weak generator* $\mathcal{G}\colon \operatorname{dom}(\mathcal{G}) \subseteq C_{\mathrm{w}}(\mathcal{H}^+;\mathbb{R}) \to C_{\mathrm{w}}(\mathcal{H}^+;\mathbb{R})$ of $(P_t)_{t\geq 0}$ is defined as follows: $f \in \operatorname{dom}(\mathcal{G})$ if and only if there exists a $g \in C_{\mathrm{w}}(\mathcal{H}^+,\mathbb{R})$ such that $g(x) = \lim_{t\downarrow 0} t^{-1}(P_t f(x) - f(x))$ and for all $x \in \mathcal{H}^+$

$$P_t f(x) = f(x) + \int_0^t P_s g(x)\,\mathrm{d}s, \quad \text{and in this case set } \mathcal{G}f := g.$$

[1] Alternatively, we could work in the framework of generalized Feller semigroups and their generators again, as we did in the latter part of Chapter 2.

We then recall the following result(s) from Chapter 2:

Theorem 3.3. *Let (b, B, m, μ) be an admissible parameter set according to Definition 2.3. Then there exist constants $M, \omega \in [1, \infty)$ and a square-integrable $\mathcal{H}^+$-valued time-homogeneous Markov process $(X_t)_{t\geq 0}$ with transition semigroup $(P_t)_{t\geq 0}$, acting on functions $f \in C_{\mathrm{w}}(\mathcal{H}^+, \mathbb{R})$, and weak generator $(\mathcal{G}, \mathrm{dom}(\mathcal{G}))$ such that the following holds:*

i) $\mathbb{E}[\|X_t\|^2 | X_0 = x] \leq M e^{\omega t}(\|x\|^2 + 1)$ *for all* $t \geq 0$,

ii) $\mathrm{lin}\left(\left\{ \mathrm{e}^{-\langle \cdot, u \rangle} : u \in \mathcal{H}^+ \right\} \cup \{ \langle \cdot, u \rangle : u \in \mathcal{H}^+ \} \right) \subseteq \mathrm{dom}(\mathcal{G})$, *and*

iii) for every $f \in \mathrm{lin} \left\{ \mathrm{e}^{-\langle \cdot, u \rangle} : u \in \mathcal{H}^+ \right\} \cup \{ \langle \cdot, u \rangle : u \in \mathcal{H}^+ \}$ *we have:*

$$\mathcal{G}f(x) = \langle b + B(x), f'(x) \rangle + \int\limits_{\mathcal{H}^+ \setminus \{0\}} \left(f(x + \xi) - f(x) - \langle \chi(\xi), f'(x) \rangle \right) \nu(x, \mathrm{d}\xi), \tag{3.8}$$

where $\nu(x, \mathrm{d}\xi) = m(\mathrm{d}\xi) + \frac{\langle \mu(\mathrm{d}\xi), x \rangle}{\|\xi\|^2}$.

An additional assumption we want to impose on the affine instantaneous covariance process X under consideration is the requirement, that X must admit for a version with càdlàg paths.

Assumption $\mathcal{A}$. The time-homogeneous Markov process X associated with the parameter set (b, B, m, μ) satisfying the conditions in Definition 2.3 has càdlàg paths.

Due to the lack of local compactness of the underlying state space, standard Feller theory were not employed to establish Theorem 3.3. We overcame this problem by using *generalized Feller semigroups*, but unfortunately the Markov processes associated to a generalized Feller semigroup need not have càdlàg paths (but see [45, Theorem 2.13] for a positive result). Some (rather limiting) conditions that ensure that Assumption $\mathcal{A}$ is satisfied are provided in the lemma below. In Chapter 6 we show that Assumption $\mathcal{A}$ is satisfied in general.

Lemma 3.4. *Assume that (b, B, m, μ) is an admissible parameter set that fulfill either one of the following two cases:*

i) (the Lévy-driven case) $\mu(\mathrm{d}\xi) = 0$,

ii) (finite activity jumps) $m(\mathcal{H}^+ \setminus \{0\}) < \infty$ *and* $\int_{\mathcal{H}^+ \setminus \{0\}} \langle x, \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \rangle < \infty$ *for all* $x \in \mathcal{H}^+$.

Then the affine Markov process $(X_t)_{t\geq 0}$ associated with (b, B, m, μ) admits for a version with càdlàg paths.

Proof. To prove i), observe that the weak generator (3.8) associated to the admissible parameters $(b, B, m, 0)$ is a weak generator of a Lévy driven SDE as described for example in [123, equation 9.37]) and hence the assertion follows from [123, Theorem 4.3]. In case of ii), the assertion follows from Proposition 2.40. □

We show in the next proposition that the version of X with càdlàg paths is in fact a Markovian semimartingale:

Proposition 3.5. *Suppose that (b, B, m, μ) is an admissible parameter set according to Definition 2.3 and such that the associated affine Markov process X satisfies Assumption $\mathcal{A}$. Then there exists a version of $(X_t)_{t\geq 0}$ which is an $\mathcal{H}^+$-valued semimartingale with characteristics triplet (A, C, ν^X) of the form:*

$$A_t = \int_0^t \big(b + B(X_s)\big)\,\mathrm{d}s\,,\ t \geq 0, \tag{3.9}$$

$$C_t = 0\,,\ t \geq 0, \tag{3.10}$$

$$\nu^X(\mathrm{d}t, \mathrm{d}\xi) = \nu(X_t, \mathrm{d}\xi)\,\mathrm{d}t = \Big(m(\mathrm{d}\xi) + \langle X_t, \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\rangle\Big)\,\mathrm{d}t. \tag{3.11}$$

Moreover, the following representation holds

$$X_t = X_0 + \int_0^t \Big(b + B(X_s) + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \xi\,\nu(X_s, \mathrm{d}\xi)\Big)\,\mathrm{d}s + \bar{J}_t, \quad t \geq 0, \tag{3.12}$$

where $\bar{J}$ is a purely discontinuous square-integrable martingale.

In order to prove Proposition 3.5, we need the following result, which can be obtained by mimicking the proof of [123, Proposition 9.38]:

Proposition 3.6. *Let X be a square-integrable time-homogeneous càdlàg Markov process on $\mathcal{H}^+$ with transition semigroup $(P_t)_{t\geq 0}$ acting on $C_{\mathrm{w}}(\mathcal{H}^+, \mathbb{R})$, let $\mathcal{G}$ be its weak generator and let $f \in \mathrm{dom}(\mathcal{G})$. Define $M_t = f(X_t) - f(X_0) - \int_0^t (\mathcal{G}f)(X_s)\,\mathrm{d}s$. Then $(M_t)_{t\geq 0}$ is a real-valued martingale.*

Proof of Proposition 3.5. Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$, then for every $n \in \mathbb{N}$, we have $e_n = e_n^+ - e_n^-$, for $e_n^+, e_n^- \in \mathcal{H}^+$. By Theorem 3.3 and Proposition 3.6 applied to $f = \langle \cdot, e_n\rangle$ there exists a square-integrable martingale $\bar{J}^{(n)}$ such that

$$\langle X_t, e_n\rangle = \langle X_0, e_n\rangle + \int_0^t \Big(\langle b + B(X_s), e_n\rangle + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \langle \xi, e_n\rangle \nu(X_s, \mathrm{d}\xi)\Big)\,\mathrm{d}s$$
$$+ \bar{J}_t^{(n)}, \qquad t \geq 0\,.$$

Noting that $X=\sum_{n=1}^{\infty}\langle X,e_n\rangle e_n$, we infer that X is an $\mathcal{H}^+$-valued semimartingale with the decomposition in (3.12), where $\bar{J}=\sum_{n=1}^{\infty}\bar{J}^{(n)}e_n$ is a square integrable $\mathcal{H}$-valued martingale.
We are left to show that $\bar{J}$ is purely discontinuous and to make the characteristic triplet of X explicit. These are known results in the finite-dimensional setting (see for instance [80, Theorem II.2.42]). Below, we adapt the proof of [80, Theorem II.2.42] to our setting. For that we decompose $X=A^{\hat{X}}+N^{\hat{X}}+\check{X}$ as in (3.6) and (3.7). Denote by $(A^{\hat{X}},C,\nu^X)$ the characteristic triplet of the semimartingale X. Let $u\in\mathcal{H}^+$ be arbitrary and consider the function $g_u=\mathrm{e}^{-\langle\cdot,u\rangle}$, $u\in\mathcal{H}^+$. On the one hand, applying the Itô formula to $g_u(X)$ (see for instance, [117, Theorem 27.2]), yields that $g_u(X)$ is a real-valued semimartingale and

$$
\begin{aligned}
&\mathrm{e}^{-\langle X_t,u\rangle}\\
&\quad=\mathrm{e}^{-\langle X_0,u\rangle}-\int_0^t\mathrm{e}^{-\langle X_{s-},u\rangle}\langle u,\mathrm{d}A_s^{\hat{X}}\rangle-\int_0^t\mathrm{e}^{-\langle X_{s-},u\rangle}\langle u,\mathrm{d}N_s^{\hat{X}}\rangle\\
&\qquad+\tfrac{1}{2}\int_0^t\mathrm{e}^{-\langle X_{s-},u\rangle}\langle u\otimes u,\mathrm{d}C_s\rangle_{\mathcal{L}_2(\mathcal{H})}+\int_0^t\int_{\mathcal{H}^+\setminus\{0\}}\mathrm{e}^{-\langle X_{s-},u\rangle}K(\xi,u)\nu^X(\mathrm{d}s,\mathrm{d}\xi)\\
&\qquad+\int_0^t\int_{\mathcal{H}^+\setminus\{0\}}\mathrm{e}^{-\langle X_{s-},u\rangle}K(\xi,u)\big(\mu^X(\mathrm{d}s,\mathrm{d}\xi)-\nu^X(\mathrm{d}s,\mathrm{d}\xi)\big)\,,
\end{aligned}
\tag{3.13}
$$

where $K(\xi,u)=\mathrm{e}^{-\langle\xi,u\rangle}-1+\langle\chi(\xi),u\rangle$. On the other hand, by Proposition 3.6 there exists a real-valued martingale I^u such that

$$
\begin{aligned}
\mathrm{e}^{-\langle X_t,u\rangle}&=\mathrm{e}^{-\langle X_0,u\rangle}+I_t^u-\int_0^t\mathrm{e}^{-\langle X_s,u\rangle}\big(\langle b+B(X_s),u\rangle\big)\,\mathrm{d}s\\
&\qquad+\int_0^t\int_{\mathcal{H}^+\setminus\{0\}}\mathrm{e}^{-\langle X_s,u\rangle}K(\xi,u)\nu(X_s,\mathrm{d}\xi)\,\mathrm{d}s\,,\quad t\geq 0\,.
\end{aligned}
\tag{3.14}
$$

Note that for every $t\geq 0$, the integrals with respect to $\mathrm{d}s$ on the right-hand side of (3.14) remain unchanged if we take the left-limits X_{s-} instead of X_s, as the number of jumps on $[0,t]$ is at most countable. Moreover, as X takes values in $\mathcal{H}^+$, we have that $g_u(X)$ is bounded and hence it is a special semimartingale and its canonical decomposition is unique. Therefore the finite variation part in formulas (3.13) and (3.14) must coincide, i.e. the following must hold for all $t\geq 0$ almost surely:

$$
\begin{aligned}
&-\int_0^t\mathrm{e}^{-\langle X_{s-},u\rangle}\Big(\langle u,\mathrm{d}A_s^{\hat{X}}\rangle+\tfrac{1}{2}\langle u\otimes u,\mathrm{d}C_s\rangle_{\mathcal{L}^2(\mathcal{H})}+\int_{\mathcal{H}^+\setminus\{0\}}K(\xi,u)\nu^X(\mathrm{d}s,\mathrm{d}\xi)\Big)\\
&\quad=-\int_0^t\mathrm{e}^{-\langle X_s,u\rangle}\Big(\langle b+B(X_s),u\rangle+\int_{\mathcal{H}^+\setminus\{0\}}K(\xi,u)\nu(X_s,\mathrm{d}\xi)\Big)\,\mathrm{d}s.
\end{aligned}
\tag{3.15}
$$

Now, by integrating $\mathrm{e}^{\langle X_{s-},u\rangle}$ with respect to both sides of (3.15) over $[0,t]$, we obtain

$$-\langle u, A_t^{\hat{X}}\rangle + \tfrac{1}{2}\langle u\otimes u, C_t\rangle_{\mathcal{L}^2(\mathcal{H})} + \int_{\mathcal{H}^+\setminus\{0\}} K(\xi,u)\nu^X([0,t],\mathrm{d}\xi)$$
$$= -\langle u, \int_0^t b + B(X_s)\,\mathrm{d}s\rangle + \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} K(\xi,u)\nu(X_s,\mathrm{d}\xi)\,\mathrm{d}s, \quad \forall t\geq 0 \text{ a.s.}$$

Now, following similar steps as in the proof of [80, Theorem II.2.42] we conclude that $C_t = 0$, $\nu^X([0,t],\mathrm{d}\xi) = \int_0^t \nu(X_s,\mathrm{d}\xi)\,\mathrm{d}s$ and $A_t^{\hat{X}} = \int_0^t \big(b + B(X_s)\big)\,\mathrm{d}s$, $t\geq 0$, and the statements of the proposition follow. □

3.2.2 The joint process $Z = (Y, X)$

In this section we finally present our stochastic covariance model, see Definition 3.8 below, which involves taking the square root $X^{1/2}$ of the process X from Theorem 3.3 as instantaneous covariance for the H-valued process Y given by equation (3.16) below.
Throughout this section we consider the following setting: Let (b, B, m, μ) be a parameter set according to Definition 2.3, let $x \in \mathcal{H}^+$ and $y \in H$, and let $Q \in \mathcal{L}_1(H)$ be self-adjoint and positive. Next, let X be the square-integrable time-homogeneous Markov process associated with the parameter set (b, B, m, μ) the existence of which is guaranteed by Theorem 3.3; we denote the filtered probability space on which X is defined by $(\Omega^1, \mathcal{F}^1, (\mathcal{F}^1_t)_{t\geq 0}, \mathbb{P}^1)$ and assume $\mathbb{P}^1(X_0 = x) = 1$. In addition, we let $(\Omega^2, \mathcal{F}^2, (\mathcal{F}^2_t)_{t\geq 0}, \mathbb{P}^2)$ be another filtered probability space, which satisfies the usual conditions and allows for a Q-Wiener process $W^Q\colon [0,\infty)\times\Omega \to H$. Now set

$$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}) \coloneqq \big(\Omega^1\times\Omega^2, (\mathcal{F}^1\otimes\mathcal{F}^2), (\mathcal{F}^1_t\otimes\mathcal{F}^2_t)_{t\geq 0}, \mathbb{P}^1\otimes\mathbb{P}^2\big),$$

and denote the expectation with respect to $\mathbb{P}$ by $\mathbb{E}$. With slight abuse of notation we consider X and W^Q to be processes on $(\Omega, \mathcal{F}, \mathbb{F})$ (note that they are independent).
In addition, we assume $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ to be the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H.
Now, consider the following SDE, for which Lemma 3.7 below establishes the existence of a mild solution:

$$\begin{cases} \mathrm{d}Y_t = \mathcal{A}Y_t\,\mathrm{d}t + X_t^{1/2}\,\mathrm{d}W_t^Q, & t > 0,\\ Y_0 = y. \end{cases} \tag{3.16}$$

Lemma 3.7. *Assume the setting described above, in particular, let* (b, B, m, μ) *be as in Definition 2.3 and let* X *be the associated affine process. Moreover, let Assumption* $\mathcal{A}$ *hold. Then* X *is progressive,*

$$\mathbb{E}\left[\int_0^t \|X_s^{1/2} Q^{1/2}\|^2 \,\mathrm{d}s\right] < \infty\,, \tag{3.17}$$

and moreover

$$Y_t = S(t)y + \int_0^t S(t-s) X_s^{1/2} \,\mathrm{d}W_s^Q\,, \quad t \geq 0\,, \tag{3.18}$$

is the unique mild solution to (3.16)*.*

Proof. The fact that X is progressive follows from the $\mathbb{F}$-adaptedness of X and Assumption $\mathcal{A}$. Moreover, it follows from Theorem 3.3 i) and Hölder's inequality that

$$\begin{aligned}\mathbb{E}\|X_t^{1/2} Q^{1/2}\|^2 &\leq \|Q\|_{\mathcal{L}_1(H)} \mathbb{E}\|X_t^{1/2}\|^2_{\mathcal{L}(H)} \leq \|Q\|_{\mathcal{L}_1(H)} \mathbb{E}\|X_t\| \\ &\leq \sqrt{M} \|Q\|_{\mathcal{L}_1(H)}\, \mathrm{e}^{\omega t/2} \sqrt{\mathbb{E}\|X_0\|^2 + 1},\end{aligned}$$

which implies (3.17). Standard theory on infinite dimensional SDEs (see for instance [68, Section 3]) now yields the existence of a unique mild solution to (3.16) given by (3.18). □

Definition 3.8. Assume the setting described above and let (b, B, m, μ) be as in Definition 2.3 and let X be the associated affine process. Moreover, let Assumption $\mathcal{A}$ hold and let Y be given by (3.18). Then we refer to the $H \times \mathcal{H}^+$-valued process $Z = (Y, X)$ as the *joint stochastic covariance model with affine pure-jump covariance* (and with parameters $(b, B, m, \mu, Q, \mathcal{A})$ and initial value (x, y)). Note that the process $(Z, (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}))$ is a (stochastically) weak solution to the following SDE in $H \times \mathcal{H}$:

$$\begin{cases} \mathrm{d}Z_t &= (\mathbf{b} + \mathbf{A} Z_t)\,\mathrm{d}t + \mathbf{\Sigma}(Z_t)\,\mathrm{d}\mathbf{W}_t + \mathrm{d}\mathbf{J}_t\,, \quad t \geq 0\,, \\ Z_0 &= (y, x) \in H \times \mathcal{H}^+\,, \end{cases} \tag{3.19}$$

where $\mathbf{b}, \mathbf{A}, \mathbf{\Sigma}, \mathbf{B}$, and $\mathbf{J}$ are as follows

$$\mathbf{b} := \begin{bmatrix} 0 \\ b + \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \xi\, m(\mathrm{d}\xi) \end{bmatrix}, \quad \mathbf{A} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} := \begin{bmatrix} \mathcal{A} z_1 \\ B(z_2) + \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \xi\, \frac{\langle z_2, \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2} \end{bmatrix},$$

$$\mathbf{\Sigma}(z) := \begin{bmatrix} (z_2)^{1/2} & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathrm{d}\mathbf{W} := \begin{bmatrix} \mathrm{d}W^Q \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathrm{d}\mathbf{J} := \begin{bmatrix} 0 \\ \mathrm{d}\bar{J} \end{bmatrix},$$

where $\bar{J}$ is the purely discontinuous square-integrable martingale obtained from Proposition 3.5.

Remark 3.9. The assumption that W^Q is a Q-Wiener process can be weakened whilst maintaining all results presented in this chapter. Indeed, as X itself is already $\mathcal{H}^+$ valued, it suffices to assume that $Q \in \mathcal{L}_2(H)$ (instead of Q being trace-class), see also the proof of Lemma 3.7.

In order to show that our joint model is affine (see Theorem 3.14 below), we need one further assumption. This assumption is also imposed in [21], see Proposition 3.2 of that article.

Assumption $\mathcal{B}$. There exists a positive and self-adjoint operator $D \in \mathcal{L}(H)$ such that

$$X_t^{1/2} Q X_t^{1/2} = D^{1/2} X_t D^{1/2}\,, \quad \text{for all } t \geq 0.$$

To the best of our knowledge, all examples for which Assumption $\mathcal{B}$ holds are such that Q and X_t commute for all $t \geq 0$. In fact, as commuting self-adjoint and compact operators are jointly diagonalizable, this is difficult to ensure without assuming there exists a fixed orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of H that forms the eigenvectors of Q and of X_t, $t \geq 0$. Note that this essentially reduces the state space of X to the cone of positive, square integrable sequences ℓ_2^+, i.e., we only model the eigenvalues of X, as the eigenvectors are fixed, see also Section 3.4.3. In conclusion, Assumption $\mathcal{B}$ is rather limiting. However, it can be circumvented if one considers a slightly different model, see Remarks 3.10 and 3.11 below.

Remark 3.10. Assumption $\mathcal{B}$ can be omitted if, instead of equation (3.16), one assumes that the process Y in the joint model satisfies the following stochastic differential equation:

$$\begin{cases} \mathrm{d}Y_t = \mathcal{A} Y_t \,\mathrm{d}t + D^{1/2} X_t^{1/2} \,\mathrm{d}W_t\,, \quad t \geq 0, \\ Y_0 = y, \end{cases} \tag{3.20}$$

where W is an H-cylindrical Brownian motion (i.e., $\mathrm{d}W_t$ is white noise) and $D \in \mathcal{L}_1(H)$ is positive and self-adjoint (in fact, $D \in \mathcal{H}^+$ suffices, see Remark 3.9). In this case, provided that (b, B, m, μ) is as in Definition 2.3 and $\mathcal{A}$ hold, we have

$$\mathbb{E}\left[\int_0^t \|D^{1/2} X_s^{1/2}\|^2 \,\mathrm{d}s\right] < \infty\,, \tag{3.21}$$

and

$$Y_t = S(t)y + \int_0^t S(t-s) D^{1/2} X_s^{1/2} \,\mathrm{d}W_s\,, \quad t \geq 0\,, \tag{3.22}$$

is the unique mild solution to (3.20), see also [48, Chapter 4, Section 3].

Moreover, Theorem 3.14 below remains valid: if Y is given by (3.20), (b, B, m, μ) is as in Definition 2.3 and Assumption $\mathcal{A}$ hold, we obtain *exactly* the same expression for $\mathbb{E}\left[\mathrm{e}^{\langle Y_t,u_1\rangle_H-\langle X_t,u_2\rangle}\right]$. In particular the joint model involving (3.20) under Assumption $\mathcal{A}$ coincides with the joint model involving (3.16) under Assumptions $\mathcal{A}$, and $\mathcal{B}$, in the sense that for every fixed time $t \geq 0$ the distribution of (Y_t, X_t) is the same. See Section 3.4.4 for an example of a model involving (3.20).

Remark 3.11. If $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ is the generator of an analytic semigroup and moreover $\mathcal{A}^{-\alpha} \in \mathcal{L}_4(H)$ (equivalently, $\mathcal{A}^{-2\alpha} \in \mathcal{H}$) for some $\alpha \in [0, \frac{1}{2})$, then a mild solution to (3.16) exists even if W^Q is an H-cylindrical Brownian motion. These conditions are satisfied, e.g., when $\mathcal{A}$ is the Laplacian on $\mathbb{R}^d$ for $d \in \{1, 2, 3\}$. We refer to [48] for details.
Although this provides another way to circumvent Assumption $\mathcal{B}$ (as Q is the identity in this case), we will not investigate this setting any further: For the applications we have in mind $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ fails to be the generator of an analytic semigroup. Note that to obtain the assertions of Theorem 3.14 in this setting, one would have to adapt its proof: one would not only have to approximate the operator $\mathcal{A}$ but also the noise.

3.3 The joint stochastic covariance model is affine

In this section we present our main result, namely that the stochastic volatility model $Z = (Y, X)$ in Definition 3.8 has the *affine property*, see Theorem 3.14. In particular, this means that we can express the mixed Fourier-Laplace transform $\mathbb{E}[\mathrm{e}^{i\langle Y_t,u\rangle_H-\langle X_t,v\rangle}]$ $(u \in H, v \in \mathcal{H}^+)$ in terms of the solution to *generalized mild Riccati equations* associated to the model parameters (b, B, m, μ), $\mathcal{A}$ and Q (respectively D). In the upcoming subsection we discuss the well-posedness of these generalized Riccati equations. Our main result, Theorem 3.14, is contained and proven in Section 3.3.2.

3.3.1 Analysis of the mild generalized Riccati equations

Let us fix an admissible parameter set (b, B, m, μ) according to Definition 2.3 and a positive self-adjoint operator $D \in \mathcal{L}(H)$. Recall $F\colon \mathcal{H}^+ \to \mathbb{R}$ and $R\colon \mathcal{H}^+ \to \mathcal{H}$ from (2.6) and define, by slight abuse of notation, the function $R : \mathrm{i}H \times \mathcal{H}^+ \to \mathcal{H}$ as

$$R(h, u) = R(u) - \tfrac{1}{2} D^{1/2}h \otimes D^{1/2}h \tag{3.23}$$

Let $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ be the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ and let $(\mathcal{A}^*, \mathrm{dom}(\mathcal{A}^*))$ be its adjoint. It is well known that $(\mathcal{A}^*, \mathrm{dom}(\mathcal{A}^*))$ generates the strongly continuous semigroup $(S^*(t))_{t\geq 0}$ on H, see for instance [72, Theorem 4.3].

Let $T \geq 0$, $u_1 \in \mathrm{i}H$ and $u_2 \in \mathcal{H}^+$. We consider the following system of differential equations on $[0,T]$, called the *generalized mild Riccati equations*

$$\begin{cases} \dfrac{\partial \Phi}{\partial t}(t,u) = F(\psi_2(t,u)), & \Phi(0,u) = 0, \quad (3.24a) \\ \psi_1(t,u) = u_1 - \mathrm{i}\mathcal{A}^* \left(\mathrm{i}\displaystyle\int_0^t \psi_1(s,u)\,\mathrm{d}s\right), & \psi_1(0,u) = u_1, \quad (3.24b) \\ \dfrac{\partial \psi_2}{\partial t}(t,u) = R(\psi_1(t,u),\psi_2(t,u)), & \psi_2(0,u) = u_2. \quad (3.24c) \end{cases}$$

Definition 3.12. Let $u = (u_1,u_2) \in \mathrm{i}H \times \mathcal{H}^+$. We say that $(\Phi(\cdot,u),\Psi(\cdot,u)) := (\Phi(\cdot,u),(\psi_1(\cdot,u),\psi_2(\cdot,u)))\colon [0,T] \to \mathbb{R}\times\mathrm{i}H\times\mathcal{H}$ is a *mild solution* of (3.24) whenever $\Phi(\cdot,u) \in C^1([0,T];\mathbb{R}^+)$, $\psi_1(\cdot,u) \in C([0,T];\mathrm{i}H)$, $\psi_2(\cdot,u) \in C^1([0,T];\mathcal{H}^+)$ and the map $(\Phi(\cdot,u),\Psi(\cdot,u))$ satisfies the equations (3.24)-(3.24c).

In the following proposition we show for every $u = (u_1,u_2) \in \mathrm{i}H \times \mathcal{H}^+$ the existence of a unique mild solution $(\Phi(\cdot,u),\Psi(\cdot,u))$ of (3.24).

Proposition 3.13. *Let (b,B,m,μ) be an admissible parameter set according to Definition 2.3, let $(\mathcal{A},\mathrm{dom}(\mathcal{A}))$ be the generator of a strongly continuous semigroup, and let $D \in \mathcal{L}(H)$ be positive and self-adjoint. Then for every $u \in \mathrm{i}H\times\mathcal{H}^+$ and $T \geq 0$ there exists a unique mild solution $(\Phi(\cdot,u),\Psi(\cdot,u))$ of (3.24) on the interval $[0,T]$.*

Proof. We set for $k \in \mathbb{N}$,

$$m^{(k)}(\mathrm{d}\xi) = \mathbf{1}_{\{\|\xi\|>1/k\}} m(\mathrm{d}\xi) \quad \text{and} \quad \mu^{(k)}(\mathrm{d}\xi) = \mathbf{1}_{\{\|\xi\|>1/k\}} \mu(\mathrm{d}\xi)\,.$$

Then for each $k \in \mathbb{N}$ let $F^{(k)}\colon \mathcal{H}^+ \to \mathbb{R}$ be as in (2.13a) and define $R^{(k)}\colon \mathrm{i}H \times \mathcal{H}^+ \to \mathcal{H}$ as

$$R^{(k)}(h,u) := \tilde{R}^{(k)}(u) - \tfrac{1}{2} D^{1/2}h \otimes D^{1/2}h\,, \tag{3.25}$$

where $\tilde{R}^{(k)}(u) = B^*(u) - \int_{\mathcal{H}^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi,u\rangle} - 1 + \langle \chi(\xi),u\rangle\right) \frac{\mu^{(k)}(\mathrm{d}\xi)}{\|\xi\|^2}$, $u \in \mathcal{H}^+$ is as in (2.13b). Consider on the interval $[0,T]$ the equations

$$\begin{cases} \dfrac{\partial \Phi^{(k)}}{\partial t}(t,u) = F^{(k)}(\psi_2^{(k)}(t,u)), & \Phi^{(k)}(0,u) = 0, \quad (3.26a) \\ \psi_1(t,u) = u_1 - \mathrm{i}\mathcal{A}^* \left(\mathrm{i}\displaystyle\int_0^t \psi_1(s,u)\,\mathrm{d}s\right), & \psi_1(0,u) = u_1, \quad (3.26b) \\ \dfrac{\partial \psi_2^{(k)}}{\partial t}(t,u) = R^{(k)}(\psi_1(t,u),\psi_2^{(k)}(t,u)), & \psi_2^{(k)}(0,u) = u_2\,. \quad (3.26c) \end{cases}$$

Standard semigroup theory (see, e.g., [55, Chapter II, Lemma 1.3]) ensures that the unique mild solution to (3.26b) is given by

$$\psi_1(t,(u_1,u_2)) = -\,\mathrm{i}S^*(t)(\mathrm{i}u_1)\,, \qquad t\in[0,T]$$

and $\psi_1(\cdot,u)\in C([0,T];\mathrm{i}H)$. Plugging $\psi_1(t,u)$ into (3.26c), yields

$$\frac{\partial\psi_2^{(k)}}{\partial t}(t,u) = \tilde{R}^{(k)}(\psi_2^{(k)}(t,u)) + \tfrac{1}{2}D^{1/2}S^*(t)(\mathrm{i}u_1)\otimes D^{1/2}S^*(t)(\mathrm{i}u_1)\,.$$

For $k\in\mathbb{N}$, $u_1\in\mathrm{i}H$, $t\in[0,T]$, define $\mathcal{R}_{u_1}^{(k)}(t,\cdot)\colon\mathcal{H}^+\to\mathcal{H}$, by

$$\mathcal{R}_{u_1}^{(k)}(t,h) = \tilde{R}^{(k)}(h) + \tfrac{1}{2}D^{1/2}S^*(t)(\mathrm{i}u_1)\otimes D^{1/2}S^*(t)(\mathrm{i}u_1).$$

By Lemma 2.12 the function $\tilde{R}^{(k)}$ is Lipschitz continuous on $\mathcal{H}^+$ and since the term $\frac{1}{2}D^{1/2}S^*(t)(\mathrm{i}u_1)\otimes D^{1/2}S^*(t)(\mathrm{i}u_1)$ does not depend on h, we conclude that for every $t\in[0,T]$ and $u_1\in\mathrm{i}H$ the function $\mathcal{R}_{u_1}^{(k)}(t,\cdot)$ is Lipschitz continuous on $\mathcal{H}^+$ as well, with the same Lipschitz constant as $\tilde{R}^{(k)}$. By Lemma 2.11, for every $k\in\mathbb{N}$ the function $\tilde{R}^{(k)}$ is quasi-monotone with respect to $\mathcal{H}^+$ (recall the notion of quasi-monotonicity from Definition 2.10, and see [50, Lemma 4.1 and Example 4.1] for relevant equivalent definitions). From this we conclude that $\mathcal{R}_{u_1}^{(k)}(t,\cdot)$ is also quasi-monotone for every $t\in[0,T]$ and $u_1\in\mathrm{i}H$. Moreover, the growth condition

$$\|\mathcal{R}_{u_1}^{(k)}(t,u_2)\| \le \big(\|B\|_{\mathcal{L}(\mathcal{H})}+2k\|\mu(\mathcal{H}^+\setminus\{0\})\|\big)\|u_2\| + \tfrac{1}{2}M^2\,\mathrm{e}^{2wt}\|D^{1/2}\|^2_{\mathcal{L}(H)}\|u_1\|_H^2,$$

for every $t\in[0,T]$, $u_1\in\mathrm{i}H$ holds, where the constants $M\ge 1$ and $w\in\mathbb{R}$ are such that $\|S^*(t)\|_{\mathcal{L}(H)}\le M\,\mathrm{e}^{wt}$, for all $t\ge 0$ which exist for every strongly continuous semigroup, see [55, Chapter I, Proposition 5.5]. Thus the conditions of [113, Chapter 6, Theorem 3.1 and Proposition 3.2] are satisfied and we conclude from this the existence of a unique solution $\psi_2^{(k)}(\cdot,u)$ on $[0,T]$ to the equation

$$\frac{\partial\psi_2^{(k)}}{\partial t}(t,(u_1,u_2)) = \mathcal{R}_{u_1}^{(k)}(t,\psi_2^{(k)}(t,u)),$$

such that $\psi_2^{(k)}(0,(u_1,u_2)) = u_2$, hence $\psi_2^{(k)}(\cdot,u)$ is the unique solution to equation (3.26c). By setting $\Phi^{(k)}(t,u) = \int_0^t F^{(k)}(\psi_2^{(k)}(s,u))\,\mathrm{d}s$ and the continuity of $F^{(k)}$ it follows that $(\Phi^{(k)}(\cdot,u),\psi_1(\cdot,u),\psi_2^{(k)}(\cdot,u))$ is the unique mild solution to equations (3.26a)-(3.26c) on $[0,T]$.

Now, let $\mathcal{R}_{u_1}\colon[0,T]\times\mathcal{H}^+\to\mathcal{H}$ be defined as the $\mathcal{R}_{u_1}^{(k)}$ above, only with $\tilde{R}^{(k)}$ replaced by $\tilde{R}$. By a similar reasoning as above and by Lemma 2.11 and Remark 2.13, we conclude that $\mathcal{R}_{u_1}(t,\cdot)$ is locally Lipschitz continuous on $\mathcal{H}^+$ and quasi-monotone with respect to $\mathcal{H}^+$ for every $t\in[0,T]$ and $u_1\in\mathrm{i}H$.

Thus, by [113, Chapter 6, Theorem 3.1] for every $t_0 \leq T$ and $u_2 \in \mathcal{H}^+$, there exists a $t_0 < t_{\max} \leq T$ and a mapping $\psi_{2,t_0}(\cdot,u)\colon [t_0,t_{\max}) \to \mathcal{H}^+$ such that

$$\frac{\partial \psi_{2,t_0}}{\partial t}(t,(u_1,u_2)) = \mathcal{R}_{u_1}\big(t,\psi_{2,t_0}(t,(u_1,u_2))\big), \quad \text{for } t \in [t_0,t_{\max}),$$

and $\psi_{2,t_0}(t_0,(u_1,u_2)) = u_2$. The function $\mathcal{R}_{u_1}$ maps bounded sets of $[0,\infty)\times\mathcal{H}^+$ into bounded sets of $\mathcal{H}$, thus by [113, Chapter 6, Proposition 1.1] it suffices to show that $t \mapsto \psi_2(t,u) := \psi_{2,0}(t,u)$ is bounded throughout its lifetime, to conclude that $t_{\max} = T$. By arguing as in the proof of Proposition 2.16 we see that for every $t \geq 0$ and $(u_1,u_2) \in \mathrm{i}H \times \mathcal{H}^+$ the sequence $(\psi_2^{(k)}(t,u))_{k\in\mathbb{N}}$ is a non-increasing sequence in $\mathcal{H}^+$ converging to $\psi_2(t,u) \geq 0$ for $t \in [0,t_{\max})$, hence

$$\|\psi_2(t,u)\| \leq \|\psi_2^{(k)}(t,u)\| \leq \|\psi_2^{(1)}(t,u)\|,$$

where the right-hand side is bounded on the whole $[0,T]$. Thus we conclude that $t_{\max} = T$ and $\psi_2(\cdot,u)$ is the unique solution to (3.24c). Then again by inserting $\psi_2(\cdot,u)$ into (3.24a) and the continuity of F, we conclude the existence of a unique solution $\Phi(\cdot,u)$ of (3.24a) on $[0,T]$, and thus also of $(\Phi(\cdot,u),\Psi(\cdot,u))$, the unique mild solution of (3.24) on the interval $[0,T]$. □

3.3.2 The affine property of our joint stochastic covariance model

Given the existence of a solution of the generalized mild Riccati equations (3.24), we show in the following theorem that our joint stochastic covariance model $Z = (X,Y)$ in Definition 3.8 has indeed the affine property.

Theorem 3.14. *Let $Z = (Y,X)$ be the stochastic covariance model in Definition 3.8 and let Assumption $\mathcal{B}$ hold. Moreover, let $u = (u_1,u_2) \in \mathrm{i}H \times \mathcal{H}$ and denote by $(\Phi(\cdot,u),(\psi_1(\cdot,u),\psi_2(\cdot,u)))$ the mild solution of the generalized mild Riccati equations* (3.24)*, the existence of which is guaranteed by Proposition 3.13. Then for all $t \geq 0$, it holds that*

$$\mathbb{E}\left[\mathrm{e}^{\langle Y_t,u_1\rangle_H - \langle X_t,u_2\rangle}\right] = \mathrm{e}^{-\Phi(t,u)+\langle y,\psi_1(t,u)\rangle_H - \langle x,\psi_2(t,u)\rangle}. \tag{3.27}$$

In applications, such as option-pricing, we are usually interested in distributional properties of the process $(Y_t)_{t\geq 0}$. Setting $u_2 = 0$ in equation (3.27) we obtain a quasi-explicit formula for the characteristic function of Y_t for $t \geq 0$. Due to its importance and easier reference we state it as a (trivial) corollary of Proposition 3.14.

Corollary 3.15. *Let the assumption of Theorem 3.14 hold. Then the characteristic function of the process Y is exponential-affine in its initial value $y \in H$ and the initial value $x \in \mathcal{H}^+$ of the instantaneous covariance process X, more specifically, for all $t \geq 0$ and $u_1 \in \mathrm{i}H$ we have:*

$$\mathbb{E}\left[\mathrm{e}^{\langle Y_t, u_1\rangle_H}\right] = \mathrm{e}^{-\Phi(t,(u_1,0)) + \langle y, \psi_1(t,(u_1,0))\rangle_H - \langle x, \psi_2(t,(u_1,0))\rangle}. \tag{3.28}$$

In order to prove Theorem 3.14, we first consider the joint process $(Y^{(n)}, X)$ obtained by replacing $\mathcal{A}$ in (3.16) by $\mathcal{A}^{(n)} := n\mathcal{A}(nI - \mathcal{A})^{-1}$, i.e. its *Yosida approximation.* The use of the approximation will allow us to exploit the semimartingale theory and to apply the Itô formula and standard techniques in order to show that the approximating process $(Y^{(n)}, X)$ is affine. Then, we study the affine property for the limiting process (see (3.30) below), when n goes to infinity. Given the assumptions of Lemma 3.7, we know that inequality (3.17) holds. Therefore from standard theory on infinite dimensional SDEs ([48, Proposition 6.4]) we know there exists a continuous adapted process $Y^{(n)} \colon [0,\infty) \times \Omega \to H$ such that

$$Y_t^{(n)} = y + \int_0^t \mathcal{A}^{(n)} Y_s^{(n)} \,\mathrm{d}s + \int_0^t X_s^{1/2} \,\mathrm{d}W_s^Q\,, \quad t \geq 0. \tag{3.29}$$

Moreover, [48, Proposition 7.5] ensures that

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{0\leq t\leq T} \|Y_t^{(n)} - Y_t\|_H^2\right] = 0\,. \tag{3.30}$$

See also [40, Theorem 5.1, Definition 2.6] where convergence rates are obtained for Yosida approximations of SPDEs in the case that the linear part of the drift is the generator of an analytic semigroup, e.g., a Laplacian.
Regarding the corresponding Riccati equations, we have the following result:

Proposition 3.16. *Let (b, B, m, μ) be as in Definition 2.3, let $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ be the generator of a strongly continuous semigroup, let $D \in \mathcal{L}(H)$ be a positive self-adjoint operator, and let $u \in iH \times \mathcal{H}^+$. Moreover, let $(\Phi(\cdot,u), (\psi_1(\cdot,u), \psi_2(\cdot,u)))$ be the mild solution of the generalized mild Riccati equation (3.24), and for $n \in \mathbb{N}$, let $(\Phi^{(n)}(\cdot,u), (\psi_1^{(n)}(\cdot,u), \psi_2^{(n)}(\cdot,u)))$ be the solution to (3.24) with $\mathcal{A} = \mathcal{A}^{(n)}$. Then*

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |\Phi^{(n)}(t,u) - \Phi(t,u)| = 0$$

and

$$\lim_{n\to\infty} \Big(\sup_{t\in[0,T]} \|\psi_1^{(n)}(t,u) - \psi_1(t,u)\|_H + \sup_{t\in[0,T]} \|\psi_2^{(n)}(t,u) - \psi_2(t,u)\| \Big) = 0\,.$$

Proof. The uniform convergence of $\psi_1^{(n)}(\cdot, u)$ to $\psi_1(\cdot, u)$ on $[0, T]$ is a well-known property of the Yosida approximation, see, e.g. [121, Proof of Theorem I.3.1]. Once this is established, the uniform convergence of $\psi_2^{(n)}(\cdot, u)$ to $\psi_2(\cdot, u)$ follows from [113, Chapter 6, Theorem 3.4]. The uniform convergence of $\Phi^{(n)}(\cdot, u)$ to $\Phi(\cdot, u)$ follows from the uniform convergence of $\psi_i^{(n)}(\cdot, u)$ to $\psi_i(\cdot, u)$, $i \in \{1, 2\}$. Hence the statement of the proposition is proved. □

With Proposition 3.16 and classical tools from stochastic calculus we can now prove Theorem 3.14:

Proof of Theorem 3.14. Let $T \geq 0$ and $u = (u_1, u_2) \in \mathrm{i}H \times \mathcal{H}^+$ be arbitrary. Moreover, let $(\Phi^{(n)}(\cdot, u), \Psi^{(n)}(\cdot, u))$, $n \in \mathbb{N}$, be the solution to (3.24a)-(3.24c) with $\mathcal{A} = \mathcal{A}^{(n)}$ (the n^{th} Yosida approximation). Note that as $\mathcal{A}^{(n)}$ is bounded,

$$t \mapsto \Psi^{(n)}(t, u) = (\psi_1(t, u), \psi_2(t, u))$$

is differentiable. We then define the function $f_u^{(n)}(t, y, x) \colon [0, T] \times H \times \mathcal{H}^+ \to \mathbb{C}$ as follows

$$f_u^{(n)}(t, y, x) := \mathrm{e}^{-\Phi^{(n)}(T-t,u)+\langle y,\psi_1^{(n)}(T-t,u)\rangle_H-\langle x,\psi_2^{(n)}(T-t,u)\rangle} .$$

Observe that $f_u^{(n)} \in C_b^{1,2,1}([0, T] \times H \times \mathcal{H}^+)$ and it holds

$$\begin{aligned}
&\frac{\partial}{\partial t} f_u^{(n)}(t, y, x) \\
&= \Big(\frac{\partial \Phi^{(n)}}{\partial t}(T-t, u) - \langle y, \frac{\partial \psi_1^{(n)}}{\partial t}(T-t, u)\rangle_H + \langle x, \frac{\partial \psi_2^{(n)}}{\partial t}(T-t, u)\rangle\Big) f_u^{(n)}(t, y, x) \\
&= \Big(F(\psi_2^{(n)}(T-t, u)) - \langle y, (\mathcal{A}^{(n)})^* \psi_1^{(n)}(T-t, u)\rangle_H \\
&\qquad + \langle x, R(\psi_1^{(n)}(T-t, u), \psi_2^{(n)}(T-t, u))\rangle\Big) f_u^{(n)}(t, y, x) . \qquad (3.31)
\end{aligned}$$

As before, we define $K \colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ as

$$K(u, v) = \mathrm{e}^{-\langle u,v\rangle} - 1 + \langle \chi(u), v\rangle$$

and $\tilde{K} \colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ for $\tilde{K}(u, v) = \mathrm{e}^{-\langle u,v\rangle} - 1 + \langle u, v\rangle$. Moreover, we set

$$\bar{\mu}^X(\mathrm{d}s, \mathrm{d}\xi) := \mu^X(\mathrm{d}s, \mathrm{d}\xi) - \nu(X_s, \mathrm{d}\xi)\, \mathrm{d}s .$$

Note that the joint process $(Y_t^{(n)}, X_t)_{t\geq 0}$ is a semimartingale on $H \times \mathcal{H}^+$ and $f_u^{(n)} \in C_b^{1,2,1}([0, T] \times H \times \mathcal{H}^+)$ for all $n \in \mathbb{N}$.

Thus, by an application of Itô's formula to $(f_u^{(n)}(t, Y_t^{(n)}, X_t))_{0\leq t\leq T}$ we obtain:

$$\begin{aligned}
f_u^{(n)}&(t, Y_t^{(n)}, X_t) \\
&= f_u^{(n)}(0, Y_0, X_0) + \int_0^t \frac{\partial}{\partial t} f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\,\mathrm{d}s \\
&\quad - \int_0^t f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\langle b + B(X_{s-}), \psi_2^{(n)}(T-s,u)\rangle\,\mathrm{d}s \\
&\quad + \int_0^t f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\langle \mathcal{A}^{(n)} Y_s^{(n)}, \psi_1^{(n)}(T-s,u)\rangle_H\,\mathrm{d}s \\
&\quad + \tfrac{1}{2}\int_0^t f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\langle X_{s-}^{1/2} Q X_{s-}^{1/2}, \psi_1^{(n)}(T-s,u)\otimes\psi_1^{(n)}(T-s,u)\rangle\,\mathrm{d}s \\
&\quad + \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} f_u^{(n)}(s, Y_s^{(n)}, X_{s-}) K(\xi, \psi_2^{(n)}(T-s,u))\nu(X_s, \mathrm{d}\xi)\,\mathrm{d}s \\
&\quad + \int_0^t f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\langle \psi_1^{(n)}(T-s,u), X_{s-}^{1/2}\,\mathrm{d}W_s^Q\rangle_H \\
&\quad + \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} f_u^{(n)}(s, Y_s^{(n)}, X_{s-}) \tilde{K}(\xi, \psi_2^{(n)}(T-s,u))\,\bar{\mu}^X(\mathrm{d}s, \mathrm{d}\xi) \\
&\quad - \int_0^t f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\langle \psi_2^{(n)}(T-s,u), \mathrm{d}\bar{J}_s\rangle\,. \qquad (3.32)
\end{aligned}$$

From (3.31), we infer

$$\begin{aligned}
f_u^{(n)}(t, Y_t^{(n)}, X_t) &= \int_0^t f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\langle \psi_1^{(n)}(T-s,u), X_{s-}^{1/2}\,\mathrm{d}W_s^Q\rangle_H \\
&\quad + \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} f_u^{(n)}(s, Y_s^{(n)}, X_{s-}) \tilde{K}(\xi, \psi_2^{(n)}(T-s,u))\,\bar{\mu}^X(\mathrm{d}s, \mathrm{d}\xi) \\
&\quad - \int_0^t f_u^{(n)}(s, Y_s^{(n)}, X_{s-})\langle \psi_2^{(n)}(T-s,u), \mathrm{d}\bar{J}_s\rangle\,. \qquad (3.33)
\end{aligned}$$

We hence conclude that the process $f_u^{(n)}(t, Y_t^{(n)}, X_t)$, $t\in[0,T]$ is a local martingale. Furthermore, since it is bounded on $[0,T]$, it is a martingale and it holds

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{\langle Y_T^{(n)}, u_1\rangle_H - \langle X_T, u_2\rangle}\right] &= \mathbb{E}\left[\mathrm{e}^{-\Phi^{(n)}(T,u) + \langle Y_0^{(n)}, \psi_1^{(n)}(T,u)\rangle_H - \langle X_0, \psi_2^{(n)}(T,u)\rangle}\right] \\
&= \mathrm{e}^{-\Phi^{(n)}(T,u) + \langle y, \psi_1^{(n)}(T,u)\rangle_H - \langle x, \psi_2^{(n)}(T,u)\rangle}.
\end{aligned}$$

Now, taking limits for $n\to\infty$, evoking (3.30) and Proposition 3.16 and since $T\geq 0$ was arbitrary, we conclude the proof. □

3.4 Examples of affine stochastic covariance models in infinite-dimensions

In this section we discuss several examples that are included in our class of joint stochastic covariance models with affine pure-jump instantaneous covariance processes. In all the examples we assume that the first component Y is modeled in the abstract setting of Definition 3.8, that means we do not specify Q or $\mathcal{A}$ any further, however we stress here that the HJMM modeling framework as described in Chapter 1 is our main example. For a more detailed introduction to stochastic covariance models in the HJMM framework see Section 4.3 below. In the current section our focus is on the correct specifications of the parameter set (b, B, m, μ) and initial value $X_0 = x \in \mathcal{H}^+$ such that the conditions in Definition 2.3 hold and the associated process $(X_t)_{t\geq 0}$ satisfies Assumption $\mathcal{A}$ as well as the joint process (Y, X) satisfies Assumption $\mathcal{B}$.
In Section 3.4.1 we show that Ornstein-Uhlenbeck processes driven by Lévy subordinators on $\mathcal{H}^+$ are included in our model class for the instantaneous covariance process X, which is implied by the parameter choice $\mu = 0$. Consequently, in Section 3.4.1 we conclude that our class of stochastic covariance models extends the infinite-dimensional lift of the BNS stochastic volatility model introduced in the work [21].
In the subsequent examples we focus on instantaneous covariance processes admitting for state-dependent jump intensities. In contrast to the Lévy driven case, these examples have the advantage to model *the volatility clustering* phenomenon. This was, for example, illustrated in [104] in a finite-dimensional setting where the state space is the cone of symmetric positive semi-definite matrices. In this latter paper, it was shown in a numerical example that for this type of models, the volatilities and jump intensities are time-varying leading to a clustering of jump events in phases of high jump intensities.
In Section 3.4.2 we construct a variance process X which takes values in $\{x + \lambda z \colon \lambda \geq 0\}$ for some fixed $z \in \mathcal{H}^+$. This is somewhat of a toy model: although the instantaneous covariance process is infinite-dimensional, its randomness is one-dimensional. In Section 3.4.3 we consider *a truly* infinite-dimensional stochastic variance process X. However, to ensure that Assumption $\mathcal{B}$ is satisfied, we assume that both Q and X_t, $t \geq 0$, are diagonalizable with respect to the same fixed ONB. We close this section with Section 3.4.4 in which we show the benefits of the model discussed in Remark 3.10, which does not require Assumption $\mathcal{B}$ and thus allows for a more general instantaneous covariance processes.
Note that the examples of affine stochastic covariance models presented in this section are either Lévy driven or have jumps of finite-activity. Indeed, the examples are such that Lemma 3.4 guarantees Assumption $\mathcal{A}$. A more general example, going beyond these restrictions, is given in Section 6.4 below.

3.4.1 An operator-valued BNS stochastic covariance model

In [21] the authors introduced an operator-valued BNS stochastic covariance model, that is an extension of the finite-dimensional model introduced in [10] (which explains the nomenclature *operator-valued BNS SV model*). In their model, it is assumed that the instantaneous covariance process X is driven by a Lévy process $(L_t)_{t\geq 0}$. In order to ensure that X is positive, it is assumed that $t \mapsto L_t$ is almost surely increasing with respect to $\mathcal{H}^+$, i.e. that L is an $\mathcal{H}^+$-*subordinator*. This holds if and only if for any fixed $t \geq 0$, we have $\mathbb{P}(L_t \in \mathcal{H}^+) = 1$, (see also [122, Proposition 9]). Roughly speaking, the model considered in [21] amounts to taking $\mu \equiv 0$ in our setting (i.e, to considering a stochastic covariance model $Z = (Y, X)$ in Definition 3.8 with parameters $(b, B, m, 0, Q, A)$). Indeed, in Subsection 3.4.1 below we demonstrate that the model introduced in [21] is fully contained in our setting.
First, however, we show in the following proposition that for this stochastic covariance model the characteristic function of Y_t, $t \in [0, T]$, can be made explicit up to the Laplace exponent of the driving Lévy subordinator.

Proposition 3.17. *Let $(b, B, m, 0)$ be as in Definition 2.3 and let X be the associated affine process with $X_0 = x \in \mathcal{H}^+$. Moreover, let $Q \in \mathcal{L}_1(H)$ be positive and self-adjoint such that Assumption $\mathcal{B}$ holds and $\mathcal{A}\colon \operatorname{dom}(\mathcal{A}) \subseteq H \to H$ be the generator of the strongly continuous semigroup $(S(t))_{t\geq 0}$. Then for every $y \in H$, the mild solution Y of* (3.16) *exists and for all $v_1 \in H$ and $t \geq 0$ it holds that*

$$\begin{aligned}\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle Y_t, v_1\rangle_H}\right] &= \exp\left(\mathrm{i}\langle y, S^*(t)v_1\rangle_H\right)\\ &\quad\times \exp\left(-\int_0^t \varphi_L\left(\tfrac{1}{2}\int_0^s \mathrm{e}^{(s-\tau)B^*}(D^{1/2}S^*(\tau)v_1)^{\otimes 2}\,\mathrm{d}\tau\right)\mathrm{d}s\right)\\ &\quad\times \exp\left(-\tfrac{1}{2}\langle x, \int_0^t \mathrm{e}^{\tau B^*}(D^{1/2}S^*(t-\tau)v_1)^{\otimes 2}\,\mathrm{d}\tau\rangle\right), \end{aligned} \tag{3.34}$$

where $\varphi_L\colon \mathcal{H} \to \mathbb{C}$ denotes the Laplace exponent of the Lévy process L with characteristics $(b, 0, m)$ and is given by

$$\varphi_L(u) = \langle b, u\rangle - \int_{\mathcal{H}^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, u\rangle} - 1 + \langle \chi(\xi), u\rangle\right) m(\mathrm{d}\xi)\,, \quad u \in \mathcal{H}^+. \tag{3.35}$$

Proof. The admissible parameter set $(b, B, m, 0)$ corresponds to the solution X of a linear stochastic differential equation driven by a Lévy process $(L_t)_{t\geq 0}$ with characteristics $(b, 0, m)$. It is easy to see that X has càdlàg paths and hence Assumption $\mathcal{A}$ is satisfied. Thus we are in the situation of Corollary 3.15 and conclude that the affine transform formula (3.28) holds for the mild solution $(\Phi(\cdot, v), (\psi_1(\cdot, v), \psi_2(\cdot, v)))$ of the generalized Riccati equations associated with $(b, B, m, 0)$ and initial value $v = (v_1, 0)$ for $v_1 \in H$.

Hence, it is left to show that the solutions have the explicit form as indicated by formula (3.34). Indeed, observe that the unique mild solution to equation (3.24b) is given by $\psi_1(t,(v_1,0)) = \mathrm{i}S^*(t)v_1$. Then inserting $\psi_1(\cdot,(v_1,0))$ into (3.24c) and recalling that $\mu = 0$ yields

$$\frac{\partial \psi_2}{\partial s}(s,(v_1,0)) = B^*(\psi_2(s,(v_1,0))) + \tfrac{1}{2}D^{1/2}S^*(t)v_1 \otimes D^{1/2}S^*(t)v_1\,.$$

By the variation of constant formula and recalling that $\psi_2(0,(v_1,0)) = 0$, we conclude that the unique solution $\psi_2(\cdot,(v_1,0))$ is given by

$$\begin{aligned}\psi_2(t,(v_1,0)) &= \tfrac{1}{2}\int_0^t \mathrm{e}^{(t-s)B^*}\left(D^{1/2}S^*(s)v_1 \otimes D^{1/2}S^*(s)v_1\right)\mathrm{d}s\\ &= \tfrac{1}{2}\int_0^t \mathrm{e}^{\tau B^*}\left(D^{1/2}S^*(t-\tau)v_1 \otimes D^{1/2}S^*(t-\tau)v_1\right)\mathrm{d}\tau.\end{aligned}$$

Lastly, by inserting $\psi_2(\cdot,(v_1,0))$ into (3.24a) and since F is a continuous function, integrating (3.24a) with respect to t gives

$$\begin{aligned}\Phi(t,(v_1,0)) = &\int_0^t \Big(\langle b, \psi_2(s,(v_1,0))\rangle\\ &\qquad - \int_{\mathcal{H}^+\setminus\{0\}}\left(\mathrm{e}^{-\langle \xi, \psi_2(s,(v_1,0))\rangle} - 1 + \langle \chi(\xi), \psi_2(s,(v_1,0))\rangle\right) m(\mathrm{d}\xi)\Big)\,\mathrm{d}s\\ = &\int_0^t \varphi_L(\psi_2(s,(v_1,0))\,\mathrm{d}s.\end{aligned}$$

Now, by inserting those formulas of $\Phi(t,(v_1,0))$, $\psi_1(t,(v_1,0))$ and $\psi_2(t,(v_1,0))$ into (3.28) we obtain the desired formula. □

Comparison with the operator-valued BNS model

In [21] the following infinite dimensional volatility model is considered for $t \geq 0$:

$$\begin{cases}\mathrm{d}Y_t &= \mathcal{A}Y_t\,\mathrm{d}t + X_t^{1/2}\,\mathrm{d}W_t^Q,\\ \mathrm{d}X_t &= B(X_t)\,\mathrm{d}t + \mathrm{d}L_t,\end{cases} \tag{3.36}$$

where $(L_t)_{t\geq 0}$ is an $\mathcal{L}_2(H)$-valued Lévy process satisfying $\mathbb{P}(L_t \in \mathcal{H}^+) = 1$ for every $t \geq 0$. Moreover, it is assumed that $B\colon \mathcal{L}_2(H) \to \mathcal{L}_2(H)$ is of the form $B(v) = cvc^*$ or $B(v) = cv + vc^*$ for some $c \in \mathcal{L}(H)$. Finally, $\mathcal{A}\colon \operatorname{dom}(\mathcal{A}) \subseteq H \to H$ is assumed to be an unbounded operator generating a strongly continuous semigroup and $(W_t)_{t\geq 0}$ is assumed to be an H-valued Brownian motion which (at least, in the part of [21] involving the affine property of (Y,X)) is assumed to be independent of $(L_t)_{t\geq 0}$ and with a covariance operator Q that satisfies Assumption $\mathcal{B}$.

In this section we show that the joint volatility model (3.36) is a special case of our model in the case that $\mu \equiv 0$, more specifically, that [21, Proposition 3.2] is a special case of Proposition 3.17 above. To this end, we first remark that if $\gamma \in \mathcal{L}_2(H)$, $C \in \mathcal{L}_1(\mathcal{L}_2(H))$, and $\eta\colon \mathcal{B}(\mathcal{L}_2(H)) \to [0,\infty]$ are the characteristics of L, then $C|_{\mathcal{H}} \equiv 0$ thanks to [21, Proposition 2.10]. Moreover, in view of Lemma 3.20, we have that $\gamma \in \mathcal{H}$, $C = 0$, and $\operatorname{supp}(\eta) \subseteq \mathcal{H}$ (this answers an open question in [21]: see the discussion prior to Proposition 2.11 in that article). Finally, it is easily verified that $B(\mathcal{H}) \subset \mathcal{H}$ in both cases described above, so although the 'ambient' space for X is $\mathcal{L}_2(H)$ in [21], one can, without loss of generality, take $\mathcal{H}$ as ambient space for X.

Next, note that the process X in (3.36) has càdlàg paths by construction (see also Lemma 3.4), so Assumption $\mathcal{A}$ is satisfied. It remains to verify that the conditions in Definition 2.3 are met. Note that Definition 2.3 iii) is immediately satisfied as $\mu \equiv 0$. To verify that the two choices for B described above satisfy Definition 2.3 iv), we recall from [21, Lemma 2.2] that in these cases one has $\mathrm{e}^{tB}(\mathcal{H}^+) \subseteq \mathcal{H}^+$ for all $t \geq 0$, which, by [106, Theorem 1], implies that B is quasi-monotone. Finally, Definition 2.3 ii) and Definition 2.3 i) hold due to the following result from [122]:

Theorem 3.18. *Let $(L_t)_{t\geq 0}$ be an $\mathcal{H}$-valued Lévy process with characteristic triplet (γ, C, η). Then the following two statements are equivalent:*

i) for all $t \geq 0$ we have $\mathbb{P}(L_t \in \mathcal{H}^+) = 1$;

ii) $C = 0$, $\operatorname{supp}(\eta) \subseteq \mathcal{H}^+$ and there exists an $I_\eta \in \mathcal{H}$ such that $\xi \mapsto |\langle \chi(\xi), h\rangle|$ is η-integrable and $\int_{\mathcal{H}^+\setminus\{0\}} \langle \chi(\xi), h\rangle\, \eta(\mathrm{d}\xi) = \langle I_\eta, h\rangle$ for all $h \in \mathcal{H}$, and such that $\gamma - I_\eta \in \mathcal{H}^+$.

Proof. First, note that $\mathcal{H}^+$ is *regular* (see, e.g., [92, Theorem 1]), i.e., any sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ satisfying $A_1 \leq_{\mathcal{H}^+} A_2 \leq_{\mathcal{H}^+} \ldots \leq_{\mathcal{H}^+} A$ for some $A \in \mathcal{H}$ is convergent in $\mathcal{H}$. The cone is also *normal*: its dual $\mathcal{H}^+$ is generating for $\mathcal{H}$. Thus $\mathcal{H}^+$ is a regular normal proper cone in the terminology of [122]. Now, note that the implication "i)⇒ii)" follows from [122, Theorem 18], and reverse implication follows from [122, Theorem 10]. □

3.4.2 An essentially one-dimensional variance process

In this section we present a simple example of a pure-jump affine process $(X_t)_{t\geq 0}$ on $\mathcal{H}^+$ with state-dependent jump intensity. Starting from its initial value $X_0 = x \in \mathcal{H}^+$ this process moves along a single vector $z \in \mathcal{H}^+ \setminus \{0\}$ and is thus essentially one-dimensional. For this case we specify an admissible parameter set (b, B, m, μ) such that the associated affine process X has càdlàg paths and is driven by a pure-jump process $(\bar{J}_t)_{t\geq 0}$ with jumps of size $\xi \in (0,\infty)$ in the single direction $z \in \mathcal{H}^+$ with $\|z\| = 1$ and such that the jump-intensity depends on the current state of the process X. For the sake of simplicity, we let the constant parameters b and m be zero. Moreover, we shall fix the dependency structure by means of a fixed vector $g \in \mathcal{H}^+ \setminus \{0\}$. We then take a measure $\eta\colon \mathcal{B}((0,\infty)) \to [0,\infty)$ such that $\int_0^\infty \lambda^{-2}\eta(\mathrm{d}\lambda) < \infty$ and define the vector valued measure $\mu\colon \mathcal{B}(\mathcal{H}^+ \setminus \{0\}) \to \mathcal{H}^+$ by

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

From the assumption that $\int_0^\infty \lambda^{-2}\eta(\mathrm{d}\lambda) < \infty$ it follows that for every $x \in \mathcal{H}^+$ the measure $M(x,\mathrm{d}\xi)$ on $\mathcal{B}(\mathcal{H}^+ \setminus \{0\})$ defined by

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

is finite and thus also

$$\int_{\mathcal{H}^+\setminus\{0\}} \langle \chi(\xi), u\rangle \frac{\langle x, \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2} = \int_0^1 \lambda^{-1}\eta(\mathrm{d}\lambda)\langle z, u\rangle\langle g, x\rangle < \infty, \quad \forall u, x \in \mathcal{H}^+.$$

We now must find a linear operator $B\colon \mathcal{H} \to \mathcal{H}$ such that

$$\langle B^*(u), x\rangle - \int_{\mathcal{H}^+\setminus\{0\}} \langle \chi(\xi), u\rangle \frac{\langle x, \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2} \geq 0, \tag{3.37}$$

whenever $\langle x, u\rangle = 0$ for $x, u \in \mathcal{H}^+$. The simplest example is obtained by taking

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

From this it can be seen that B and μ indeed satisfy condition (3.37) and conclude that the parameter set $(0, B, 0, \mu)$ is an admissible parameter set according to Definition 2.3. Thus the existence of an associated affine process X on $\mathcal{H}^+$ is guaranteed by Theorem 3.3. Since $\int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^{-2}\langle x, \mu(\mathrm{d}\xi)\rangle < \infty$ for all $x \in \mathcal{H}^+$, it follows from Lemma 3.4 that Assumption $\mathcal{A}$ is satisfied as well. It remains to ensure that Assumption $\mathcal{B}$ is satisfied. For this purpose it suffices to assume that x and z commute with Q. Indeed, note that for $u \in \{x + \lambda z \colon \lambda \in [0,\infty)\}$ we have $B(u) \in \{\lambda z \colon \lambda \in [0,\infty)\}$. Thus from the semimartingale representation (3.12), we see that

$$X_t \in \{x + \lambda z \colon \lambda \in [0,\infty)\}, \quad \forall t \geq 0,$$

which implies that X_t commutes with Q for all $t \geq 0$ and therefore Assumption $\mathcal{B}$ is satisfied.

3.4.3 Stochastic covariance models on fixed orthonormal bases with state-dependent jumps

In this example we specify an admissible parameter set (b, B, m, μ) giving more general affine dynamics of the associated variance process X on $\mathcal{H}^+$. In the previous Section 3.4.2 we imposed additional commutativity assumptions on the initial value $X_0 = x \in \mathcal{H}^+$, the jump direction z and the covariance operator Q. In this example we allow for a more general jump behavior, while maintaining Assumption $\mathcal{B}$. To do so, we pick up the discussion preceding Remark 3.10 and note here that Assumption $\mathcal{B}$ is satisfied, whenever Q and X_t commute for all $t \geq 0$. Recall that Q and $(X_t)_{t\geq 0}$ commute if and only if they are jointly diagonalizable. This motivates the consideration of a variance process X that is diagonalizable with respect to a fixed ONB.
More concretely, let $(e_n)_{n\in\mathbb{N}}$ be an ONB of eigenvectors of the operator Q. We model X such that X_t $(t \geq 0)$ is diagonalizable with respect to the ONB $(e_n)_{n\in\mathbb{N}}$, i.e.

$$X_t = \sum_{i\in\mathbb{N}} \lambda_i(t) e_n \otimes e_n, \quad t \geq 0,$$

for the sequence of eigenvalues $(\lambda_i(t))_{i\in\mathbb{N}}$ of X_t in ℓ_2^+. Concerning the modeling of the dynamics of $(X_t)_{t\geq 0}$, this essentially means that we model the dynamics of the sequence of eigenvalues $(\lambda_i(t))_{i\in\mathbb{N}}$ in ℓ_2^+ only.
We now come to a specification of the parameters (b, B, m, μ) such that the conditions in Definition 2.3 are satisfied and moreover such that X_t is indeed diagonalizable with respect to $(e_n)_{n\in\mathbb{N}}$ for all $t \geq 0$. Let $m\colon \mathcal{B}(\mathcal{H}^+ \setminus \{0\}) \to [0, \infty)$ be such that for $A \in \mathcal{B}(\mathcal{H}^+ \setminus \{0\})$ we have

$$m(A) \coloneqq \sum_{n\in\mathbb{N}} m_n(\{\lambda \in (0,\infty) \colon \lambda(e_n \otimes e_n) \in A\}), \tag{3.38}$$

for a sequence $(m_n)_{n\in\mathbb{N}}$ of finite measures on $\mathcal{B}((0,\infty))$ such that

$$\sum_{n\in\mathbb{N}} m_n((0,\infty)) < \infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \int_1^\infty \lambda^2 m_n(\mathrm{d}\lambda) < \infty. \tag{3.39}$$

Then let $\tilde{b} \in \mathcal{H}^+$ be diagonalizable with respect to $(e_n)_{n\in\mathbb{N}}$ and set

$$b \coloneqq \tilde{b} + \int_{\mathcal{H}^+\setminus\{0\}} \chi(\xi) m(\mathrm{d}\xi) = \tilde{b} + \sum_{n\in\mathbb{N}} \int_0^1 \lambda\, m_n(\mathrm{d}\lambda) e_n \otimes e_n.$$

We see that b and m satisfy their respective conditions in Definition 2.3.

Now, let $(g_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}^+$ and let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of finite measures on $\mathcal{B}((0,\infty))$ such that

$$\sum_{n\in\mathbb{N}} g_n\mu_n((0,\infty)) \in \mathcal{H}^+ \quad \text{and} \quad \sum_{n\in\mathbb{N}} \int_0^1 \lambda^{-2}\mu_n(\mathrm{d}\lambda)\langle g_n, x\rangle < \infty, \quad \forall x \in \mathcal{H}^+, \tag{3.40}$$

and define $\mu(\mathrm{d}\xi)\colon \mathcal{B}(\mathcal{H}^+ \setminus \{0\}) \to \mathcal{H}^+$ by

$$\mu(A) \coloneqq \sum_{n\in\mathbb{N}} g_n\mu_n(\{\lambda \in (0,\infty)\colon \lambda(e_n \otimes e_n) \in A\}). \tag{3.41}$$

Moreover, let $G \in \mathcal{H}$ be diagonalizable with respect to $(e_n)_{n\in\mathbb{N}}$, note that this implies that for any $x \in \mathcal{H}^+$ that is diagonalizable with respect to $(e_n)_{n\in\mathbb{N}}$, we have that $Gx + xG^*$ is diagonalizable with respect to $(e_n)_{n\in\mathbb{N}}$ as well. We thus define the linear operator $B\colon \mathcal{H} \to \mathcal{H}$ by

$$B(u) = Gu + uG^* + \int_{\mathcal{H}^+\setminus\{0\}} \chi(\xi)\frac{\langle \mu(\mathrm{d}\xi), u\rangle}{\|\xi\|^2}, \quad u \in \mathcal{H}. \tag{3.42}$$

Now, one can check that B and μ indeed satisfy their respective conditions in Definition 2.3. Due to the first condition on m in (3.39) and the second on μ in (3.40), it follows from Lemma 3.4 that Assumption $\mathcal{A}$ is satisfied.
Again from the semimartingale representation (3.12) we conclude that for all $t \geq 0$ the operator X_t is diagonalizable with respect to $(e_n)_{n\in\mathbb{N}}$ and thus Assumption $\mathcal{B}$ is satisfied as well.

3.4.4 A general infinite-dimensional affine stochastic covariance model with state-dependent jumps

In this example we show that modeling under the alternative formulation of the model (Y, X) provided by Remark 3.10 gives considerably more freedom in the model parameter specification. We write $\hat{b} = b + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \xi\, m(\mathrm{d}\xi)$ and for every $u \in \mathcal{H}$ we set $\hat{B}(u) = B(u) + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \xi\, \frac{\langle u,\mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2}$. We then see that for the stochastic volatility model (Y, X) given by the SDE

$$\mathrm{d}(Y_t, X_t) = \begin{bmatrix} 0 \\ \hat{b} \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \mathcal{A}Y_t \\ \hat{B}(X_t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} D^{1/2}X_t^{1/2} & 0 \\ 0 & 0 \end{bmatrix} \mathrm{d}\begin{bmatrix} W_t \\ 0 \end{bmatrix} + \mathrm{d}\begin{bmatrix} 0 \\ \bar{J}_t \end{bmatrix}, \quad t \geq 0,$$

with $(Y_0, X_0) = (y, x) \in H \times \mathcal{H}^+$ and $W = (W_t)_{t\geq 0}$ a cylindrical Brownian motion, the Assumption $\mathcal{B}$ can be dropped. Therefore, every admissible parameter set (b, B, m, μ), such that the associated affine process X satisfies Assumption $\mathcal{A}$ is a valid parameter choice.

To emphasize the gained flexibility, we compare it with the example in Section 3.4.3. For simplicity, we let $(e_n)_{n\in\mathbb{N}}$ be some ONB of H and specify m and μ as in (3.38) and (3.41), respectively, with respect to this ONB. This means that the noise in the instantaneous covariance process X again occurs on the diagonal only. However, Q need not be diagonalizable with respect to $(e_n)_{\in\mathbb{N}}$ and instead of taking b to be diagonalizable with respect to the ONB $(e_n)_{\in\mathbb{N}}$ and B of the particular form (3.42), we allow for a general drift $\hat{b}\in\mathcal{H}$ such that $\hat{b}-\int_{\mathcal{H}^+\setminus\{0\}}\chi(\xi)m(\mathrm{d}\xi)\geq 0$. Moreover, let C be a bounded linear operator on H and define $B\in\mathcal{L}(\mathcal{H})$ by

$$B(u)=Cu+uC^*+\Gamma(u),$$

for some $\Gamma\in\mathcal{L}(\mathcal{H})$ with $\Gamma(\mathcal{H}^+)\subseteq\mathcal{H}^+$ and such that

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

We can again check, that (b,B,m,μ) satisfies the conditions of Definition 2.3 and the associated affine process X Assumption $\mathcal{A}$. Then, according to (3.12) the instantaneous covariance process X has the representation

$$\begin{aligned}X_t=X_0+\int_0^t\Big(b+CX_s+X_sC^*+\Gamma(X_s)+\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\xi\nu(X_s,\mathrm{d}\xi)\Big)\,\mathrm{d}s\\+\bar{J}_t,\qquad t\geq 0\,,\end{aligned}\tag{3.43}$$

where $\nu(x,\mathrm{d}\xi)$ is as in (3.11) and $\bar{J}$ is a purely discontinuous square integrable martingale with a compensator ν^X. The dynamics (3.43) has a similar structure as the affine dynamics of covariance processes in finite dimensions presented in [42, equation 1.2] in the pure-jump case. Indeed, both models have an affine drift and are driven by a pure-jump process whose compensator is an affine function of X. As mentioned above, this model will also demonstrate clustering behavior.

3.5 Auxiliary results

Lemma 3.19. *Let $(\mathcal{H},\|\cdot\|,\langle\cdot,\cdot\rangle)$ be a separable real Hilbert space, let $K\subseteq\mathcal{H}$ be a cone such that $\mathcal{H}=K-K$ and let $\mu_1,\mu_2\colon\mathcal{B}(\mathcal{H})\to\mathbb{R}$ be measures such that* $\operatorname{supp}(\mu_1),\operatorname{supp}(\mu_2)\subseteq K$ *and*

$$\int_{\mathcal{H}}\mathrm{e}^{-\langle x,y\rangle}\,\mu_1(\mathrm{d}x)=\int_{\mathcal{H}}\mathrm{e}^{-\langle x,y\rangle}\,\mu_2(\mathrm{d}x),$$

for all $y\in K$. Then $\mu_1=\mu_2$.

Proof. We assume that $\mathcal{H}$ is infinite-dimensional, the proof for finite dimensional $\mathcal{H}$ is analogous. For $n \in \mathbb{N}$ let $\mathbb{R}_n^+ := [0,\infty)^n \subseteq \mathbb{R}^n$. We proof the assertion in two steps:
Step 1: reduction to the cone $\mathbb{R}_n^+$, $n \in \mathbb{N}$. Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{H}$. For all $n \in \mathbb{N}$ let $e_n^+, e_n^- \in K$ be such that $e_n = e_n^+ - e_n^-$ and define

$$D_n := \left\{ \cap_{i=1}^n \{x \in \mathcal{H} \colon \langle x, e_i^+ \rangle \in B_i^+, \langle x, e_i^- \rangle \in B_i^- \} \colon B_i^+, B_i^- \in \mathcal{B}(\mathbb{R}) \forall\, 1 \le i \le n \right\}.$$

Note that $\mathcal{B}(\mathcal{H}) = \sigma(\cup_{n\in\mathbb{N}} D_n)$, so by Dynkin's lemma it suffices to prove that $\mu_1|_{D_n} = \mu_2|_{D_n}$ for all $n \in \mathbb{N}$. For $n \in 2\mathbb{N}$ let $\mu_{1,n}, \mu_{2,n} \colon \mathcal{B}(\mathbb{R}^{2n}) \to \mathbb{R}$ be the measures defined by

$$\mu_{i,n}(B_1^+ \times B_1^- \times \ldots \times B_n^+ \times B_n^-) := \mu_i(\cap_{i=1}^n \{x \in \mathcal{H} \colon \langle e_i^+, x\rangle \in B_i^+, \langle e_i^-, x \rangle \in B_i^- \}),$$

$i \in \{1,2\}$, $B_1^+, B_1^-, \ldots, B_n^+, B_n^- \in \mathcal{B}(\mathbb{R})$. Note that $\operatorname{supp}(\mu_{i,n}) \subseteq \mathbb{R}_{2n}^+$ and $\int_{\mathbb{R}^{2n}} \mathrm{e}^{-\langle x,y\rangle}\, \mathrm{d}\mu_{1,n}(x) = \int_{\mathbb{R}^{2n}} \mathrm{e}^{-\langle x,y\rangle}\, \mathrm{d}\mu_{2,n}(x)$ for all $y \in \mathbb{R}_{2n}^+$. Thus it suffices to prove the lemma for the case that $\mathcal{H} = \mathbb{R}^n$ and $K = [0,\infty)^n$, for all $n \in \mathbb{N}$.
Step 2: the case $\mathcal{H} = \mathbb{R}^n$ and $K = \mathbb{R}_n^+$, $n \in \mathbb{N}$. Fix $n \in \mathbb{N}$ and consider the set

$$\mathcal{A} := \operatorname{lin}\{x \mapsto \mathrm{e}^{-\langle x,y\rangle} \colon y \in \mathbb{R}_n^+\} \subseteq C(\mathbb{R}).$$

By the Stone-Weierstrass theorem the set $\{f|_{[-R,R]^n} \colon f \in \mathcal{A}\}$ is dense in the space $C([-R,R]^n)$ for all $R > 0$. Thus we can mimic the proof of e.g. [79, Theorem E.1.14] to obtain that $\mu_1 = \mu_2$. □

Lemma 3.20. *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space, let $U \subseteq H$ be a closed linear subspace and let $P_U : H \to U$ be the orthogonal projection of H onto U. Moreover, let $(L_t)_{t\ge 0}$ be a H-valued Lévy process satisfying $\mathbb{P}(L_1 \in U) = 1$ and let $\gamma \in H$, $C \in \mathcal{L}_1(H)$ and $\eta \colon \mathcal{B}(H \setminus \{0\}) \to [0,\infty]$ be its characteristics. In addition, let $\gamma_s \in U$, $C_s \in \mathcal{L}_1(U)$ and $\eta_s \colon \mathcal{B}(U \setminus \{0\}) \to [0,\infty]$ be the characteristics of L when interpreted as a $(U, \langle\cdot,\cdot\rangle)$-valued process. Then $\gamma = \gamma_s$, $C = C_s P_U$, and $\eta(A) = \eta_s(A \cap U)$ for all $A \in \mathcal{B}(H \setminus \{0\})$. In particular, $Ch = 0$ whenever $h \in U^\perp$ and $\operatorname{supp}(\eta) \subseteq U$.*

Proof. Define $\tilde\eta \colon \mathcal{B}(H \setminus \{0\}) \to [0,\infty]$ by $\tilde\eta(A) = \eta_s(A \cap U)$, $A \in \mathcal{B}(H \setminus \{0\})$. Then for all $h \in H$ and $t \ge 0$ we have, using that $L_t \in U$ a.s.:

$$\begin{aligned}
&\mathbb{E}(\mathrm{e}^{i\langle L_t, h\rangle}) = \mathbb{E}(\mathrm{e}^{i\langle L_t, P_U h\rangle}) \\
&= \exp\left(t\left(i\langle \gamma_s, P_U h\rangle - \langle C_s P_U h, P_U h\rangle\right)\right) \\
&\quad \times \exp\Big(\Big(t \int_{U\setminus\{0\}} \left(\mathrm{e}^{i\langle \xi, P_U h\rangle} - 1 + i\langle \xi, P_U h\rangle 1_{\{\|\xi\|<1\}}\right) \eta_s(\mathrm{d}\xi)\Big)\Big) \\
&= \exp\Big(t\big(i\langle\gamma_s, h\rangle - \langle C_s P_U h, h\rangle + \int_{H\setminus\{0\}} \left(\mathrm{e}^{i\langle\xi,h\rangle} - 1 + i\langle \xi, h\rangle 1_{\{\|\xi\|<1\}}\right)\tilde\eta(\mathrm{d}\xi)\big)\Big).
\end{aligned}$$

The result now follows from the uniqueness of the characteristic triplet. □

3.6 Concluding remarks

In this chapter we introduced an infinite-dimensional affine stochastic covariance model, where the instantaneous covariance process X is modeled by an affine pure-jump process with possibly state-dependent jump intensities; its existence has been established in the previous Chapter 2 under the admissibility conditions Definition 2.3 posed on the parameters (b, B, m, μ). The proposed stochastic covariance model extends the operator-valued BNS model that was introduced in [21], where X is driven by a suitably chosen Lévy process (see Section 3.4.1). Moreover, in Section 3.4, we provided several other concrete examples of affine stochastic covariance models on Hilbert spaces.

- ***On the càdlàg paths assumption***

In the derivation of the affine transform formula, we make use of Hilbert valued semimartingale calculus, for this reason we assumed that X has càdlàg paths (see Assumption $\mathcal{A}$) and proved the existence of càdlàg paths under rather restrictive conditions (see Lemma 3.4). We actually relax the conditions in Chapter 6, where we consider finite-dimensional approximations of the instantaneous covariance processes in the Skorohod space.

- ***On the 'commutativity'-type assumption***

After we introduced our joint model, we proved that it is affine (see Theorem 3.14). To this end, we need an additional 'commutativity'-type assumption, see Assumption $\mathcal{B}$. This assumption is avoided by considering a slightly different model, see Remark 3.10 and Section 3.4.4. However, we either need Assumption $\mathcal{B}$ or the additional operator $D^{1/2}$ modulating the cylindrical Brownian noise. Another way to avoid Assumption $\mathcal{B}$ would be to construct an instantaneous covariance process X that takes values in the space of self-adjoint *trace class* operators. Indeed, in this case we can assume that the process W^Q driving Y is a cylindrical Brownian motion (i.e., Q is the identity). However, as already discussed in Section 2.5, taking the trace class operators as a state space is not trivial as this is a non-reflexive Banach space. This direction of research can be pursued in a forthcoming work.

CHAPTER 4

Option-Pricing in Infinite-Dimensional Affine Stochastic Covariance Models

Abstract of the chapter In this chapter we are concerned with the pricing of options in infinite-dimensional affine stochastic covariance models. More specifically, we consider a *geometric affine stochastic covariance model* for forward curve dynamics in, e.g. commodity markets, formulated in the HJMM framework and we derive quasi-explicit formulas for *plain vanilla call options* written on forward contracts. In this model the logartihmic forward curve dynamics are given by an affine stochastic covariance model as introduced in Chapter 3. Due to the affine structure of this model, it is tempting to use Fourier techniques for deriving option price formulas. Indeed, this approach is inspired by finite-dimensional affine models and allows us to derive formulas for option prices in terms of the solutions of associated generalized Riccati equations. Using these Fourier techniques require the computation of real- and complex exponential moments and, moreover, complex extensions of the affine transform formula. Extending the affine transform formula in this infinite-dimensional setting is the main objective of the first part of this chapter. In the second part we provide the quasi-explicit option-pricing formulas and discuss examples.

This chapter is part of the working paper:
Cox, S., He, J., Karbach, S., and Khedher, A.:
Option-Pricing in Infinite-Dimensional Affine Stochastic Volatility Models.

4.1 Introduction

The affine class is widely used in finance due to its tractability entailed by the affine-transform formula. In particular, the virtues of affine stochastic covariance models are the quasi-explicit form of their Fourier-Laplace transform, which makes option-pricing by means of Fourier-inversion methods effective [33, 53, 52]. In this chapter we show that this paradigm continues to hold in infinite dimensions. Indeed, we use standard Fourier techniques to derive quasi-explicit formulas for *plain vanilla put-* and *call options* written on commodity forwards modeled in a *geometric affine stochastic covariance model* in the HJMM framework. The geometric forward curve model (see Section 4.3 for the precise definition) is based on an affine stochastic covariance model (Y, X) given by Definition 3.8.

For the Fourier-method approach to work, it is necessary to ensure the existence of real- and complex exponential moments of (Y, X), and moreover to extend the affine-transform formula (3.27), respectively (3.28), to real and complex inputs, i.e. extensions such that input variables $u_1 \in H$ and $u_1 \in H \oplus_{\mathbb{R}} \mathrm{i}H$ are permitted. This is our main concern in Section 4.2. Once we established the extended versions of the affine transform formula (Proposition 4.2 and 4.5), we present formulas for European call- and put-options on forwards in terms of the solutions to some associated extended generalized Riccati equations (see Proposition 4.7).

4.1.1 Notation and preliminaries

Throughout this chapter we use the same notations as before and fix a separable, infinite-dimensional and real Hilbert space $(H, \langle\cdot,\cdot\rangle_H)$. Moreover, we denote the complexification of a Hilbert space $(V, \langle\cdot,\cdot\rangle_V)$ by $V^{\mathbb{C}}$, i.e.

$$V^{\mathbb{C}} := V \oplus_{\mathbb{R}} \mathrm{i}V = \{x + \mathrm{i}y \colon x, y \in V\}$$

$V^{\mathbb{C}} := V \oplus_{\mathbb{R}} \mathrm{i}V$ is the complex Hilbert space equipped with the (complex-valued) inner-product by $\langle x+\mathrm{i}y, u+\mathrm{i}v\rangle_{V^{\mathbb{C}}} := \langle x, u\rangle_V + \langle y, v\rangle_V + \mathrm{i}\langle y, u\rangle_V - \mathrm{i}\langle x, v\rangle_V$, for all $x, y, u, v \in V$, i.e. here we slightly change the convention for the inner-product compared to Chapter 3. Recall that for $z = x + \mathrm{i}y \in V^{\mathbb{C}}$, we say that $x = \Re(z)$ and $y = \mathcal{I}(z)$ are the real and imaginary parts of z. Moreover, $\bar{z} = \Re(z) - \mathrm{i}\,\mathcal{I}(z)$ denotes the complex conjugate of z. For the space of bounded linear operators from H to H, as usual denoted by $\mathcal{L}(H)$, we let $\mathcal{L}_1(H) \subseteq \mathcal{L}(H)$ denote subspace of *trace class operators*. Recall that $\mathcal{L}_1(H)$ is a Banach space with the norm

$$\|A\|_1 = \sum_{n=1}^{\infty} \langle (A^*A)^{1/2} e_n, e_n\rangle_H,$$

where $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis for H and where the norm is independent of the choice of $(e_n)_{n\in\mathbb{N}}$.

4.1.2 Setting

In this chapter we are working in the setting of Chapter 3. In particular, we let (b, B, m, μ) be an admissible parameter set as in Definition 2.3 and denote by $X = (X_t)_{t\geq 0}$ the associated affine process on positive self-adjoint Hilbert-Schmidt operators, the existence of which is guaranteed by Theorem 2.8. Recall from Proposition 3.5 that whenever $(X_t)_{t\geq 0}$ admits a version with càdlàg paths, this version is a square-integrable semimartingale that admits representation (4.1b) below. Throughout this chapter we assume that $(X_t)_{t\geq 0}$ has càdlàg paths.
In Section 4.2 we study the existence of real- and complex-exponential moments for affine stochastic covariance models (Y, X) given by(4.1) below. In particular, we do this under the following finite-variation assumption on the constant and linear jump measures:

Assumption $\mathcal{C}$. The jump measures m and μ satisfy

$$\int_{\mathcal{H}^+\cap\{\|\xi\|\leq 1\}} \|\xi\|\, m(\mathrm{d}\xi) < \infty\,, \quad \int_{\mathcal{H}^+\cap\{\|\xi\|\leq 1\}} \|\xi\|^{-1}\langle x, \mu(\mathrm{d}\xi)\rangle < \infty\,, \quad \forall\, x \in \mathcal{H}^+.$$

Moreover, we assume that $\mu(\mathrm{d}\xi)$ takes values in $\mathcal{L}_1(H) \cap \mathcal{H}^+$.

In case that $\mu = 0$, i.e. the Lévy case we can drop Assumption $\mathcal{C}$ for the constant jump measure m. The general case, i.e. without assuming Assumption $\mathcal{C}$, is left-open for the moment and will be treated in a future work.

Now, let us fix $y, \Upsilon \in H$, $x \in \mathcal{H}^+$, and let $D \in \mathcal{L}_2(H)$ be self-adjoint and positive. As before, we assume that $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H. Following Definition 3.8 and Remark 3.10, we consider the process $(Y, X, \Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ given by the (stochastically) weak solution to the following stochastic differential equation with parameters $(b, B, m, \nu, D, \Upsilon, \mathcal{A})$ and initial value $(Y_0, X_0) = (y, x)$:

$$\begin{cases} \mathrm{d}Y_t = \Big(\mathcal{A}Y_t + D^{1/2}X_t D^{1/2}\Upsilon\Big)\,\mathrm{d}t + D^{1/2}X_t^{1/2}\,\mathrm{d}W_t\,, \qquad t > 0\,, & \text{(4.1a)} \\ \mathrm{d}X_t = \Big(b + B(X_t) + \displaystyle\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \xi\,\nu(X_t, \mathrm{d}\xi)\Big)\,\mathrm{d}t + \mathrm{d}\bar{J}_t, \qquad t > 0\,, & \text{(4.1b)} \end{cases}$$

where $\bar{J} = (\bar{J}_t)_{t\geq 0}$ is a purely discontinuous square-integrable martingale, W is a cylindrical Brownian motion, independent of $\bar{J}$, and $\nu(x, \mathrm{d}\xi) = m(\mathrm{d}\xi) + \|\xi\|^{-2}\langle x, \mu(\mathrm{d}\xi)\rangle$. The existence of a (stochastically) weak solution of the stochastic differential equation (4.1) on $H \times \mathcal{H}^+$ follows by analogous arguments as in Proposition 3.6 and Lemma 3.7. Notice that in this chapter, in view of the application in Section 4.3, we added the linear part $G(X_t) = D^{1/2}X_t D^{1/2}\Upsilon$ in the drift of Y, which is a slight extension of the model specification in Definition 3.8. However, adding the linear drift-term G does not change well-posedness or the affine-property, as the proofs of Lemma 3.7 and Theorem 3.14 show.

4.2 Exponential moments of affine stochastic covariance models

Our main objective in this chapter is the pricing of options written on forwards modeled with log-dynamics of type (4.1). For this purpose we use Fourier transform techniques that allow us to write option prices in terms of solutions to associated generalized Riccati equations. However, this approach requires the computation of real- and complex exponential moments of the model (Y, X). This is the purpose of the present section, where we are concerned with real exponential moments in Section 4.2.1 and subsequently in Section 4.2.2 with the case of complex exponential moments.

4.2.1 Real exponential moments

For $x \in \mathcal{H}^+$, an admissible parameter set (b, B, m, μ) satisfying Assumption $\mathcal{C}$, and kernel $\nu(x, \mathrm{d}\xi)$ we define the set

$$\mathcal{U} \coloneqq \left\{ u \in \mathcal{H} \colon \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \mathrm{e}^{-\langle \xi, u\rangle}\, \nu(x, \mathrm{d}\xi) < \infty, \quad \forall x \in \mathcal{H}^+ \right\}.$$

We note that $\mathcal{U}$ is a convex subset in $\mathcal{H}$ and since $\mathcal{H}^+ \subseteq \mathcal{U}$, it is non-empty. Moreover, we assume that $\nu(x, \mathrm{d}\xi)$ is such that $\mathcal{U}$ is an open subset of $\mathcal{H}$. We then consider $F \colon \mathcal{U} \to \mathbb{R}$ and $R_\Upsilon \colon H \times \mathcal{U} \to \mathcal{H}$, the extensions of F and R from (2.6a) and (3.23) on $\mathcal{U}$ (with additional drift term $G^*(h) = -D^{1/2}h \otimes D^{1/2}\Upsilon$ added to function R), defined for $(h, u) \in H \times \mathcal{U}$ by

$$F(u) \coloneqq \langle \tilde{b}, u\rangle - \int_{\mathcal{H}^+\setminus\{0\}} \big(\mathrm{e}^{-\langle \xi, u\rangle} - 1\big) m(\mathrm{d}\xi), \tag{4.2}$$

$$\begin{aligned} R_\Upsilon(h, u) \coloneqq {}& \tilde{B}^*(u) - \tfrac{1}{2}(D^{1/2}h)^{\otimes 2} - \int_{\mathcal{H}^+\setminus\{0\}} \big(\mathrm{e}^{-\langle \xi, u\rangle} - 1\big) \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \\ & - D^{1/2}h \otimes D^{1/2}\Upsilon, \end{aligned} \tag{4.3}$$

where we set $\tilde{b} \coloneqq b - \int_{\mathcal{H}^+\setminus\{0\}} \chi(\xi)\, m(\mathrm{d}\xi)$ and for all $v \in \mathcal{H}^+$

$$\tilde{B}(v) \coloneqq B(v) - \int_{\mathcal{H}^+\setminus\{0\}} \chi(\xi) \frac{\langle v, \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2}.$$

Note that $\tilde{b}$ and $\tilde{B}$ are well defined due to the finite-variation Assumption $\mathcal{C}$.

For $T \geq 0$ and $u = (u_1, u_2) \in H \times \mathcal{U}$ we consider the following *extended generalized (mild) Riccati equations*:

$$\begin{cases} \dfrac{\partial P(t,u)}{\partial t} = F(q_2(t,u)), & 0 < t \leq T, \quad P(0,u) = 0, \\ q_1(t,u) = u_1 + \mathcal{A}^* \left(\displaystyle\int_0^t q_1(s,u)\,\mathrm{d}s \right), & 0 < t \leq T, \quad q_1(0,u) = u_1, \\ \dfrac{\partial q_2(t,u)}{\partial t} = R_\Upsilon(q_1(t,u), q_2(t,u)), & 0 < t \leq T, \quad q_2(0,u) = u_2. \end{cases} \begin{matrix} (4.4a) \\ (4.4b) \\ (4.4c) \end{matrix}$$

For $u = (u_1, u_2) \in H \times \mathcal{U}$ and $T \geq 0$, we say that $(P(\cdot,u), q_1(\cdot,u), q_2(\cdot,u))$, a mapping from $[0,T]$ to $\mathbb{R} \times H \times \mathcal{H}$, is a mild solution to (4.4) whenever $P(\cdot,u) \in C^1([0,T],\mathbb{R})$, $q_1(\cdot,u) \in C([0,T],H)$ and $q_2(\cdot,u) \in C^1([0,T],\mathcal{U})$ satisfy the equations (4.4a)-(4.4c).
In the following proposition we show the existence of a unique solution of the extended Riccati equations (4.4) up to its maximal lifetime.

Proposition 4.1. *Let (b, B, m, μ) be an admissible parameter set such that Assumption $\mathcal{C}$ is satisfied. Let $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ be the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H, let $D \in \mathcal{L}(H)$ be a positive self-adjoint operator, and let $\Upsilon \in H$. Then for every $u = (u_1, u_2) \in H \times \mathcal{U}$ there exist a real number $T_{q_2} \geq 0$ and a unique solution $(P(\cdot,u), q_1(\cdot,u), q_2(\cdot,u))$ of (4.4) on the interval $[0, T_{q_2})$.*

Proof. Standard semigroup theory ensures that for any $u_1 \in H$ and $T > 0$ the unique mild solution of (4.4b) is given by $q_1(t,u) = S^*(t)u_1$ for $t \in [0,T]$. We define $\mathcal{R}_\Upsilon^{u_1}(t,\cdot)\colon \mathcal{U} \to \mathcal{H}$, by

$$\mathcal{R}_\Upsilon^{u_1}(t,u) := B^*(u) - \tfrac{1}{2}\big(D^{1/2}S^*(t)u_1\big)^{\otimes 2} - D^{1/2}S^*(t)u_1 \otimes D^{1/2}\Upsilon \\ - \int_{\mathcal{H}^+\setminus\{0\}} \big(\mathrm{e}^{-\langle \xi, u\rangle} - 1\big)\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}.$$

By plugging $q_1(t,u)$ into (4.4c) we thus obtain the equation

$$\begin{cases} \frac{\partial q_2}{\partial t}(t,u) = \mathcal{R}_\Upsilon^{u_1}(t, q_2(t,u)), \\ q_2(0,u) = u_2\,. \end{cases} \tag{4.5}$$

Observe that the function $\mathcal{R}_\Upsilon^{u_1}(t,\cdot)$ is locally Lipschitz continuous on $\mathcal{U}$ for every $t \geq 0$ and $u_1 \in H$. Since $\mathcal{U}$ is assumed to be open, it follows from standard ODE results (see, e.g. [113, Chapter 6, Proposition 1.2]) that for every $u_2 \in \mathcal{U}$ there exist a $T_{q_2} > 0$ and a unique solution $q_2(\cdot,u)$ of (4.5) on $[0, T_{q_2})$ with

$$T_{q_2} = \liminf_{n\to\infty} \{t \geq 0\colon \|q_2(t,u)\| \geq n \text{ or } q_2(t,u) \in \partial\mathcal{U}\}\,.$$

Finally, by inserting $q_2(\cdot,u)$ into (4.4a) and observing that F is continuous on $\mathcal{U}$, the statement follows. ☐

In order to compute exponential moments for the process (Y, X), we first consider the joint process $(Y^{(n)}, X)$ obtained by replacing the unbounded operator $\mathcal{A}$ in (4.1a) by its Yosida approximation $\mathcal{A}^{(n)} := n\mathcal{A}(nI - \mathcal{A})^{-1}$, as we did in Section 3.3.2. We therefore consider the process $Y^{(n)} \colon [0, \infty) \times \Omega \to H$ given by the solution to (see [48, Proposition 6.4] for the existence of the solution to $Y^{(n)}$)

$$Y_t^{(n)} = y + \int_0^t (\mathcal{A}^{(n)} Y_s^{(n)} + D^{1/2} X_s D^{1/2} \Upsilon) \, \mathrm{d}s + \int_0^t D^{1/2} X_s^{1/2} \, \mathrm{d}W_s \,, \quad t \geq 0. \tag{4.6}$$

Using this approximation of Y allows us to exploit the semimartingale theory. In particular, we can apply Itô's formula and standard techniques in order to compute exponential moments for $(Y^{(n)}, X)$. We suppose that the results are maintained when taking the limit n to infinity, since by [48, Proposition 7.5] it holds that

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{0\leq t\leq T} \|Y_t^{(n)} - Y_t\|_H^2\right] = 0\,. \tag{4.7}$$

Indeed, we have the following:

Proposition 4.2. *Let (b, B, m, μ) be an admissible parameter set satisfying Assumption $\mathcal{C}$, let (Y, X) be the stochastic covariance model in (4.1), let $u = (u_1, u_2) \in H \times \mathcal{U}$, and let $(P(\cdot, u), q_1(\cdot), u), q_2(\cdot, u))$ be the mild solution to the Riccati equations (4.4) up to time T, the existence of which is guaranteed by Proposition 4.1. Then we have*

$$\mathbb{E}\left[\mathrm{e}^{\langle Y_T, u_1\rangle_H - \langle X_T, u_2\rangle}\right] < \infty. \tag{4.8}$$

Moreover, for every $t \leq T$, it holds

$$\mathbb{E}\left[\mathrm{e}^{\langle Y_T, u_1\rangle_H - \langle X_T, u_2\rangle} \mid \mathcal{F}_t\right] = \mathrm{e}^{-P(T-t,u) + \langle Y_t, q_1(T-t,u)\rangle_H - \langle X_t, q_2(T-t,u)\rangle}. \tag{4.9}$$

Proof. Let $(P^{(n)}(\cdot, u), q_1^{(n)}(\cdot, u), q_2^{(n)}(\cdot, u))$ be the solution of (4.4) with $\mathcal{A} = \mathcal{A}^{(n)}$ (the n-th Yosida approximation). We define the function $g_u^{(n)}(t, y, x) \colon [0, T] \times H \times \mathcal{H}^+ \to \mathbb{R}$ as follows

$$g_u^{(n)}(t, y, x) := \exp\left(-P^{(n)}(T - t, u) + \langle y, q_1^{(n)}(T - t, u)\rangle_H - \langle x, q_2^{(n)}(T - t, u)\rangle\right) .$$

Similarly to the proof of Theorem 3.14, we can apply Itô's formula to the process $g_u^{(n)}(t, Y_t^{(n)}, X_t)$ to deduce that it is in fact a local martingale.

Then, by observing that $g_u^{(n)}(t, Y_t^{(n)}, X_t)$ is strictly positive for all $t \in [0,T]$, we infer it must be a supermartingale and we thus conclude that

$$\begin{aligned}\mathbb{E}\left[\mathrm{e}^{\langle Y_T^{(n)}, u\rangle_H - \langle X_T, u\rangle}\right] &= \mathbb{E}\left[g_u^{(n)}(T, Y_T^{(n)}, X_T)\right] \\ &\leq g_u^{(n)}(0, y, x) < \infty\,, \quad (y,x) \in H \times \mathcal{H}\,.\end{aligned}$$

From this it follows that $(g_u^{(n)}(t, Y_t^{(n)}, X_t))_{t\in[0,T]}$ is actually a proper martingale and hence for every $t \in [0,T]$ we see that

$$\mathbb{E}\left[\mathrm{e}^{\langle Y_T^{(n)}, u_1\rangle_H - \langle X_T, u_2\rangle} \mid \mathcal{F}_t\right] = \mathrm{e}^{-P^{(n)}(T-t,u) + \langle Y_t^{(n)},\, q_1^{(n)}(T-t,u)\rangle_H - \langle X_t,\, q_2^{(n)}(T-t,u)\rangle}.$$

Following similar arguments as in the proof of Proposition 3.16, we deduce that

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |P^{(n)}(t,u) - P(t,u)| = 0$$

and

$$\lim_{n\to\infty}\left(\sup_{t\in[0,T]} \|q_1^{(n)}(t,u) - q_1(t,u)\|_H + \sup_{t\in[0,T]} \|q_2^{(n)}(t,u) - q_2(t,u)\|\right) = 0\,.$$

Hence taking limits for $n \to \infty$ and invoking (4.7) yields the statement. □

4.2.2 Complex exponential moments

In this section, and as a next step, we extend the exponential moments for affine stochastic covariance models from the real set $H \times \mathcal{U}$ to a set of complex vectors. More precisely, let $\mathcal{H}^{\mathbb{C}}$ denote the complexification of $\mathcal{H}$ and define the following *complex strip* in $\mathcal{H}^{\mathbb{C}}$:

$$S(\mathcal{U}) \coloneqq \left\{u \in \mathcal{H}^{\mathbb{C}} \colon\ \Re(u) \in \mathcal{U}\right\}.$$

Note that for every $u \in S(\mathcal{U})$ and $x \in \mathcal{H}^+$ we have

$$\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} |\,\mathrm{e}^{-\langle \xi, u\rangle_{\mathcal{H}^{\mathbb{C}}}}|\nu(x,\mathrm{d}\xi) = \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \mathrm{e}^{-\langle \xi, \Re(u)\rangle}\nu(x,\mathrm{d}\xi) < \infty\,.$$

Moreover, it follows from [132, Theorem 25.17] that the functions $u \mapsto F(u)$ and $u \mapsto R_\Upsilon(h,u)$ ($h \in H$) in (4.2) and (4.3) can be analytically extended to the complex strip $S(\mathcal{U}) \subseteq \mathcal{H}^{\mathbb{C}}$. We therefore consider F and R_Υ throughout the remainder of this section as functions from $S(\mathcal{U})$ to $\mathbb{C}$, respectively, from $H^{\mathbb{C}} \times S(\mathcal{U})$ to $\mathcal{H}^{\mathbb{C}}$.

Next, we consider the following *complex generalized (mild) Riccati equations*:

$$\begin{cases} \dfrac{\partial \phi(t,u)}{\partial t} = F(\psi_2(t,u)), & 0 < t \leq T, \quad \phi(0,u) = 0, \quad (4.10a) \\ \psi_1(t,u) = u_1 + \mathcal{A}^* \displaystyle\int_0^t \psi_1(s,u)\,\mathrm{d}s, & 0 < t \leq T, \quad \psi_1(0,u) = u_1\,, \quad (4.10b) \\ \dfrac{\partial \psi_2(t,u)}{\partial t} = R_\Upsilon(\psi_1(t,u), \psi_2(t,u)), & 0 < t \leq T, \quad \psi_2(0,u) = 0. \quad (4.10c) \end{cases}$$

Let $T \geq 0$ and $u = (u_1, 0)$ for some $u_1 \in H^{\mathbb{C}}$. Analogously to the extended generalized Riccati equations (4.4), we say that the map $(\phi(\cdot,u), \psi_1(\cdot,u), \psi_2(\cdot,u))$ is a solution to (4.10) if $\phi(\cdot,u) \in C^1([0,T],\mathbb{C}), \psi_1(\cdot,u) \in C([0,T],H^{\mathbb{C}})$ and $\psi_2(\cdot,u) \in C^1([0,T],S(\mathcal{U}))$ satisfy equations (4.10a)-(4.10c).
Before proving the existence and uniqueness of a solution to (4.10), we give the following lemma that follows along the lines of [95, Lemma 5.12] under Assumption $\mathcal{C}$.

Lemma 4.3. *Let* R_Υ *be as in* (4.3) *and let Assumption* $\mathcal{C}$ *be satisfied. Then there exists a locally Lipschitz continuous function* g *on* $\mathcal{H}$ *such that for all* $h \in H^{\mathbb{C}}$ *and* $u \in S(\mathcal{U})$ *we have*

$$\Re(\langle \overline{u}, R_\Upsilon(h,u)\rangle) \leq g(\Re(u))(1 + \|u\|^2)(1 + \|h\|^2), \tag{4.11}$$

Proof. We first split $\Re(\langle \overline{u}, R_\Upsilon(u,h)\rangle)$ into three parts as follows:

$$\Re(\langle \overline{u}, R_\Upsilon(h,u)\rangle) = I_1 + I_2 + I_3\,, \tag{4.12}$$

where

$$\begin{aligned} I_1 &= \Re(\langle \overline{u}, \tilde{B}^*(u)\rangle) + \Re(\langle \overline{u}, D^{1/2}h \otimes D^{1/2}h\rangle) + \Re(\langle \overline{u}, D^{1/2}h \otimes D^{1/2}\Upsilon\rangle)\,, \\ I_2 &= -\Re\Big(\int_{\mathcal{H}^+ \cap \{0 < \|\xi\| \leq 1\}} \big(\mathrm{e}^{-\langle u,\xi\rangle_{\mathcal{H}^{\mathbb{C}}}} - 1\big) \langle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}, \overline{u}\rangle \Big)\,, \\ I_3 &= -\Re\Big(\int_{\mathcal{H}^+ \cap \{\|\xi\| > 1\}} \big(\mathrm{e}^{-\langle u,\xi\rangle_{\mathcal{H}^{\mathbb{C}}}} - 1\big) \langle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}, \overline{u}\rangle \Big). \end{aligned}$$

For the first term on the right-hand side of (4.12), it holds

$$\begin{aligned} I_1 &\leq \|\tilde{B}^*\|_{\mathcal{L}(\mathcal{H})}\|u\|^2 + \|D\|_{\mathcal{L}(H)}\|h\|^2\|u\| + \|D\|_{\mathcal{L}(H)}\|h\|\|\Upsilon\|\|u\| \\ &\leq g_1(1 + \|u\|^2)(1 + \|h\|^2), \end{aligned}$$

where $g_1 \coloneqq \|\tilde{B}^*\|_{\mathcal{L}(\mathcal{H})} + \|D\|_{\mathcal{L}(H)} + \|D\|_{\mathcal{L}(H)}\|\Upsilon\|$. For the second and third term, i.e. I_2 and I_3, respectively. We can argue along the lines of [95, Lemma 5.12]. □

In the following proposition we assert the existence of a unique solution of equations (4.10) on the interval $[0,T]$. Here again we follow the lines of [95] and show that the arguments can be extended to our infinite-dimensional setting.

Proposition 4.4. *Let (b, B, m, μ) be an admissible parameter set satisfying Assumption $\mathcal{C}$, let $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ be the generator of a strongly continuous semigroup, let $D \in \mathcal{L}(H)$ be a positive self-adjoint operator, and let $\Upsilon \in H$. Assume that the extended generalized Riccati equations* (4.4) *have a unique solution on* $[0,T]$, *for $(u_1, u_2) \in H \times \mathcal{U}$. Then for every $u = (u_1, 0)$, $u_1 \in H^{\mathbb{C}}$, there exist a unique solution $(\phi(\cdot, u), \psi_1(\cdot, u), \psi_2(\cdot, u))$ to* (4.10) *on the interval* $[0,T]$.

Proof. Let $u = (u_1, u_2) \in H^{\mathbb{C}} \times S(\mathcal{U})$ and $T \geq 0$. By standard semigroup theory, we see that the unique mild solution to (4.10c) on $[0,T]$ is given by $\psi_1(t,u) = S^*(t)u_1 \in H^{\mathbb{C}}$. For every $t \in [0,T]$ we consider the extension of the map $\mathcal{R}^{u_1}_{\Upsilon}$, from the proof of Proposition 4.1, to a function $\mathcal{R}^{u_1}_{\Upsilon}(t, \cdot)\colon S(\mathcal{U}) \to \mathcal{H}^{\mathbb{C}}$. Then by plugging $\psi_1(t,u)$ into (4.4c), we obtain the equation

$$\begin{cases} \frac{\partial \psi_2(t,u)}{\partial t} &= \mathcal{R}^{u_1}_{\Upsilon}(t, \psi_2(t,u))\,, \\ \psi_2(0,u) &= 0. \end{cases} \tag{4.13}$$

The map $\mathcal{R}^{u_1}_{\Upsilon}(t,\cdot)$ is locally Lipschitz continuous on $S(\mathcal{U})$, which again by standard arguments implies the existence of a unique solution $\psi_2(\cdot, u)$ on some interval $[0, T_{\psi_2})$ with values in $S(\mathcal{O}) \subseteq \mathcal{H}^{\mathbb{C}}$. We want to show that the lifetime T_{ψ_2} of $\psi_2(\cdot, u_1, 0)$ is always greater or equal than the lifetime T of $q_2(\cdot, \Re(u_1), 0)$. First observe that $-\big(D^{1/2}S^*(t)\Im(u)\big)^{\otimes 2} \leq_{\mathcal{H}^+} 0$ and

$$\int_{\mathcal{H}^+\setminus\{0\}} \Re(e^{-\langle \xi, u\rangle_{\mathcal{H}^{\mathbb{C}}}} - 1)\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \leq_{\mathcal{H}^+} \int_{\mathcal{H}^+\setminus\{0\}} (e^{-\langle \xi, \Re(u)\rangle} - 1)\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}.$$

This together with the fact that $q_2(0, \Re(u_1), 0) = \Re(\psi_2(0, u_1, 0)) = 0$ imply for all $t < T \wedge T_{\psi_2}$

$$\begin{aligned} &\frac{\partial q_2(t, \Re(u_1), 0)}{\partial t} - \mathcal{R}^{\Re(u_1)}_{\Upsilon}\big(t, q_2(t, \Re(u_1), 0)\big) \\ &\quad = \frac{\partial \Re(\psi_2(t, u_1, 0))}{\partial t} - \Re\big(\mathcal{R}^{u_1}_{\Upsilon}\big(t, \psi_2(t, u_1, 0)\big)\big) \\ &\quad \leq_{\mathcal{H}^+} \frac{\partial \Re(\psi_2(t, u_1, 0))}{\partial t} - \mathcal{R}^{\Re(u_1)}_{\Upsilon}\big(t, \Re(\psi_2(t, u_1, 0))\big)\,. \end{aligned} \tag{4.14}$$

Note that $\mathcal{R}^{u}_{\Upsilon}(t,\cdot)$ is quasi-monotone with respect to $\mathcal{H}^+$, for $u \in H$ (see [144, Section 6]) for the notion of quasi-monotonicity). Then by Volkmann's comparison result [144, Satz 2] and (4.14), we have

$$q_2(t, \Re(u_1), 0) \leq_{\mathcal{H}^+} \Re(\psi_2(t, u_1, 0)), \quad \text{for all } t < T \wedge T_{\psi_2}. \tag{4.15}$$

We now prove that this implies that $T \leq T_{\psi_2}$. First, note that whenever $u \in \mathcal{U}$ and $v \in \mathcal{H}$ are such that $u \leq_{\mathcal{H}^+} v$, then $v \in \mathcal{U}$, hence $\Re(\psi_2(t, u_1, 0))$ does not approach the boundary of $\mathcal{U}$ before $q_2(t, u_1, 0)$ does. Next we show that also the function $t \mapsto \|\Re(\psi_2(t, u))\|$ does not explode before $t \mapsto \|q_2(t, u)\|$.
Indeed, it follows along the lines of [95, Lemma 5.12] together with (4.15) that

$$\begin{aligned}
&\|\Re(\mathcal{R}_{\Upsilon}^{u_1}(\psi_2(t,u)))\| \\
&\quad \leq \|\tilde{B}\|_{\mathcal{L}(\mathcal{H})}\|\Re(\psi_2(t,u))\| + \frac{1}{2}M^2\,\mathrm{e}^{2\omega t}\|D\|_{\mathcal{L}(\mathcal{H})}\|u_1\| + M\,\mathrm{e}^{\omega t}\|D\|_{\mathcal{L}(\mathcal{H})}\|u_1\|\|\Upsilon\| \\
&\qquad + \left\| \int_{\mathcal{H}^+\cap\{0<\|\xi\|\leq 1\}} \|\xi\|^{-1}\mu(\mathrm{d}\xi) \right\| \mathrm{e}^{\|q_2(t,\Re(u))\|}\|\Re(\psi_2(t,u))\| \\
&\qquad + \|\mu(\mathcal{H}^+\cap\{\|\xi\|>1\})\| + \left\| \int_{\mathcal{H}^+\cap\{\|\xi\|>1\})} \mathrm{e}^{-\langle \xi, q_2(t,\Re(u))\rangle}\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \right\|,
\end{aligned} \tag{4.16}$$

where the constants $M \geq 1$ and $\omega \in \mathbb{R}$ are such that $\|S^*(t)\|_{\mathcal{L}(H)} \leq M\,\mathrm{e}^{\omega t}$. Hence by an application of Gronwall's inequality we conclude that $\|\Re(\psi_2(t,u))\|$ is bounded as long as $\|q_2(t,u)\|$ is bounded. Now, by Lemma 4.3, we have

$$\Re(\langle \overline{u}, R_{\Upsilon}(h,u)\rangle) \leq g(\Re(u))(1+\|u\|^2)(1+\|h\|^2),$$

where g is a locally Lipschitz continuous function on $\mathcal{H}$ and thus

$$\begin{aligned}
\frac{\partial}{\partial t}\|\psi_2(t,u)\|^2 &= 2\Re\big(\langle \overline{\psi_2(t,u)}, R_{\Upsilon}(\psi_1(t,u),\psi_2(t,u))\rangle\big) \\
&\leq g(\Re(\psi_2(t,u)))(1+\|\psi_2(t,u)\|^2)(1+\|\psi_1(t,u)\|^2)\,.
\end{aligned}$$

Again by an application of Gronwall's inequality and since we already proved that $\Re(\psi_2(\cdot, u_1, 0))$ does not explode before $q_2(\cdot, \Re(u_1), 0)$ we conclude that $T_{\psi_2} \geq T$, which proves the assertion. □

From Proposition 4.4 and following similar derivations as in [95, Theorem 2.26 and Section 5.3], we deduce the following result on the existence of complex moments and the legitimacy of the affine transform formula (4.9) on complex vectors $u_1 \in H^{\mathbb{C}}$.

Proposition 4.5. *Let (b, B, m, μ) be an admissible parameter set satisfying Assumption $\mathcal{C}$, let (Y, X) be the stochastic covariance model in* (4.1)*, and assume that the equations* (4.4) *have a unique solution on $[0, T]$ for $u = (u_1, u_2) \in H \times \mathcal{U}$. Then for every $t \in [0, T]$, and $u_1 \in H^{\mathbb{C}}$, we have $\mathbb{E}\left[|\,\mathrm{e}^{\langle u_1, Y_t\rangle_{H^{\mathbb{C}}}}|\right] < \infty$, and*

$$\mathbb{E}\left[\mathrm{e}^{\langle u_1, Y_T\rangle_{H^{\mathbb{C}}}} \mid \mathcal{F}_t\right] = \mathrm{e}^{-\phi(T-t,u_1,0)+\langle \psi_1(T-t,u_1,0), Y_t\rangle_{H^{\mathbb{C}}} - \langle \psi_2(T-t,u_1,0), X_t\rangle_{\mathcal{H}^{\mathbb{C}}}},$$

where $(\phi(\cdot, u_1, 0), \psi_1(\cdot, u_1, 0), \psi_2(\cdot, u_1, 0))$ is the unique solution of (4.10) *on the interval $[0, T]$, the existence of which is guaranteed by Proposition 4.4.*

4.3 Pricing options written on forwards

In the following we recall some preliminaries on commodity forward curve modeling in the Heath-Jarrow-Morton-Musiela framework as described in [18]. Moreover, we connect our stochastic covariance model (Y, X) to this setting and study the pricing of options written on forwards, the logarithmic dynamics of which are modeled by (Y, X) on a suitable Hilbert space H.
Suppose $w\colon \mathbb{R}^+ \to [1, \infty)$ is a non-decreasing measurable function such that $w(0) = 1$ and $\int_0^\infty w^{-1}(x)\, \mathrm{d}x < \infty$. Let H_w be the space of absolutely continuous functions $f\colon \mathbb{R}^+ \to \mathbb{R}$ such that

$$\|f\|_w^2 \coloneqq f(0)^2 + \int_0^\infty w(x)|f'(x)|^2\, \mathrm{d}x < \infty\,,$$

where f' denotes the weak derivative of f. The space H_w is a separable Hilbert space when endowed with the inner product

$$\langle f, g\rangle_w \coloneqq f(0)g(0) + \int_0^\infty w(x)f'(x)g'(x)\, \mathrm{d}x.$$

The space H_w is proposed in [60] as a class of Hilbert spaces that match the economic reasoning about the forward rate curve in fixed income markets (see also [18], where H_w is considered as the state space of forward curves in commodity markets). An example for a weight function $w\colon \mathbb{R}^+ \to [1, \infty)$ satisfying the assumptions above, is given by $w(x) = \mathrm{e}^{\beta x}$, for $\beta > 0$, in which case we write $(H_\beta, \langle\cdot,\cdot\rangle_\beta) \coloneqq (H_{\mathrm{e}^\beta}, \langle\cdot,\cdot\rangle_{\mathrm{e}^\beta})$.
It is well known that the left-shift semigroup is a strongly continuous semigroup on H_w with infinitesimal generator $\mathcal{A}$ being the operator of differentiation (in space). Moreover, the point evaluation map $\delta_x(u) = u_x$ is a continuous linear functional on H_w, for all $x \geq 0$. It can be expressed as (see [60, Lemma 5.3.1])

$$\delta_x = \langle\cdot, u_x\rangle_w\,, \tag{4.17}$$

where for $x \in \mathbb{R}^+$, $u_x : \mathbb{R}^+ \to \mathbb{R}$, $y \mapsto 1 + \int_0^{x\wedge y} w^{-1}(z)\, \mathrm{d}z$.
Let $F(t, \hat{T})$ be the forward price at time $t > 0$ of a contract delivering an asset (commodity) at time $\hat{T} > t$. The forward curve f_t in the Musiela parametrization is given in terms of time-to-maturity, i.e. $f_t(x) \coloneqq F(t, t + x)$, for $x \geq 0$. Fix a weight function w and assume that the underlying process $(Y_t)_{t\geq 0}$ is modeled by

$$\mathrm{d}Y_t = (\mathcal{A}Y_t + C_t)\, \mathrm{d}t + D^{1/2}X_t^{1/2}\, \mathrm{d}W_t\,, \quad Y_0 = y \in H_w, \tag{4.18}$$

where $\mathcal{A} = \frac{\partial}{\partial x}$ denotes the derivative operator, C is an integrable adapted process taking values in H_w, $D \in \mathcal{L}_2(H_w)$ is self-adjoint and positive, W is a cylindrical Brownian motion, independent of the affine process X on $\mathcal{H}^+$ given by (4.1a).

Let $T>0$ be a fixed time horizon and $(Y,\Omega,(\mathcal{F}_t)_{0\leq t\leq T},\mathbb{F},\mathbb{P})$ be the solution to (4.18) where we assume $\mathbb{P}$ to be the *physical measure*. We model the *arbitrage-free* under the HJMM framework as

$$f_t(x) \coloneqq \exp\big(\delta_x(Y_t)\big) = \exp\big(\langle Y_t, u_x\rangle_w\big)\,,\quad t\geq 0,\, x\geq 0, \tag{4.19}$$

where u_x is as described in (4.17) and Y is as in (4.18). Assume ξ is an H_w-valued adapted process such that

$$\mathbb{E}\left[\exp\left(-\int_0^T \langle \xi_s, \mathrm{d}W_s\rangle_w - \frac{1}{2}\int_0^T \|\xi_s\|_w^2 \,\mathrm{d}s\right)\right] = 1\,, \tag{4.20}$$

where $\mathbb{E}$ is the expectation under $\mathbb{P}$. Then, by Girsanov's theorem with respect to cylindrical Brownian motion (see e.g., [47, Theorem 13]) we know that the process

$$\tilde{W}_t = W_t - \int_0^t \xi_s \,\mathrm{d}s$$

is a cylindrical Brownian motion under the measure $\mathbb{Q}$ given by

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(-\int_0^t \langle \xi_s, \mathrm{d}W_s\rangle_w - \frac{1}{2}\int_0^t \|\xi_s\|_w^2 \,\mathrm{d}s\right)\,,\qquad 0\leq t\leq T\,.$$

The dynamics of Y under the new measure $\mathbb{Q}$ are given by

$$\mathrm{d}Y_t = (\mathcal{A}Y_t + C_t - D^{1/2}X_t^{1/2}\xi_t)\,\mathrm{d}t + D^{1/2}X_t^{1/2}\,\mathrm{d}\tilde{W}_t\,,\quad 0\leq t\leq T\,, \tag{4.21}$$

and we denote the expectation with respect to $\mathbb{Q}$ by $\mathbb{E}_{\mathbb{Q}}$. In the following lemma we state a drift condition under which $t\mapsto F(t,\tau)$, $0\leq t\leq\tau\leq T$ is a $\mathbb{Q}$-martingale measure for $t\leq T$. The proof follows the same lines as in [22, Proposition 6.8 and Lemma 6.3].

Lemma 4.6. *Let ξ be an H_w-valued adapted process satisfying* (4.20) *and let $\mathcal{A}$, C, D, and Y be as described after equation* (4.18)*. Let S be the shift semigroup associated with $\mathcal{A}$. Moreover, assume*

$$C_t - D^{1/2}X_t^{1/2}\xi_t = -\frac{1}{2}D^{1/2}X_tD^{1/2}S^*(\tau-t)u_0\,,\qquad 0\leq t\leq\tau. \tag{4.22}$$

Then $t\mapsto F(t,\tau)$, $0\leq t\leq\tau\leq T$ is a local $\mathbb{Q}$-martingale.

From now on we assume that (4.22) holds. That is we consider dynamics under $\mathbb{Q}$ of type (4.1a) with

$$\Upsilon = \Upsilon^{\vartheta} = -\frac{1}{2}S^*(\vartheta)u_0,\qquad \vartheta\geq 0. \tag{4.23}$$

Notice that the process Y and the associated extended Riccati equations (4.10) in this case will depend on the time to delivery, denoted by ϑ and it follows as in [22, Lemma 6.3] that the function $\vartheta \mapsto -\frac{1}{2}D^{1/2}X_tD^{1/2}S^*(\vartheta)u_0(\vartheta)$ is contained in the space H_w. Sometimes, we shall indicate the dependence on ϑ by writing Y^ϑ and $(\phi^\vartheta, \psi_1^\vartheta, \psi_2^\vartheta)$.
In the sequel we consider options at time t written on forwards given by

$$F(T_0,T_1) = f_{T_0}(T_1 - T_0) = \exp\big(\delta_{T_1-T_0} Y_{T_0}^{T_1-T_0}\big), \qquad T_0 \le T_1 \le T. \tag{4.24}$$

Options on forwards occur for example in classical commodity markets like oil, metals or agriculture (see [15] for a good overview). In the following proposition we derive semi-explicit expressions for a *plain-vanilla* call option (assuming zero interest-rate) written on forwards in terms of the Fourier transform of the pay-off function and the solutions of the complex generalized Riccati equations (4.10).

Proposition 4.7. *Let $T > 0$ and $0 \le T_0 \le T_1 \le T$. Set $\vartheta = T_1 - T_0$ and assume that Y^ϑ is given by* (4.1a) *with $\Upsilon = -\frac{1}{2}S^*(\vartheta)u_0$. Moreover, let the forward price F be as in* (4.24)*, set $u = (\nu + \mathrm{i}\lambda)u_\vartheta$ with $\nu > 1, \lambda \in \mathbb{R}$ and for every $u \in H^{\mathbb{C}}$ let $(\phi^\vartheta(\cdot,u,0), \psi_1^\vartheta(\cdot,u,0), \psi_2^\vartheta(\cdot,u,0))$ be the unique solution of* (4.10) *on the interval $[0,T]$. Then the price of a call option at time $t \ge 0$ with exercise time T_0 and strike $K > 0$, written on the forward F is given by*

$$\begin{aligned}&\mathbb{E}_{\mathbb{Q}}[(F(T_0,T_1) - K)^+ \mid \mathcal{F}_t] \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} g(\lambda)\, \mathrm{e}^{-\phi^\vartheta(T_0-t,u_1,0) + \langle \psi_1^\vartheta(T_0-t,u_1,0), Y_t\rangle_{H_w^{\mathbb{C}}} - \langle \psi_2^\vartheta(T_0-t,u_1,0), X_t\rangle_{\mathcal{H}^{\mathbb{C}}}}\, \mathrm{d}\lambda\,,\end{aligned} \tag{4.25}$$

where the function $g\colon \mathbb{R} \to \mathbb{C}$ is given by

$$g(\lambda) := \frac{K^{-(\nu-1+\mathrm{i}\lambda)}}{(\nu + \mathrm{i}\lambda)(\nu - 1 + \mathrm{i}\lambda)}\,. \tag{4.26}$$

Proof. Standard Fourier techniques for option pricing (see e.g., [53] or [61, Theorem 10.6]), yield

$$\mathbb{E}_{\mathbb{Q}}[(F(T_0,T_1) - K)^+ | \mathcal{F}_t] = \frac{1}{2\pi}\int_{\mathbb{R}} g(\lambda)\mathbb{E}_{\mathbb{Q}}\big[\,\mathrm{e}^{\langle(\nu+\mathrm{i}\lambda)u_\vartheta, Y_{T_0}^\vartheta\rangle_{H_w^{\mathbb{C}}}}\, |\mathcal{F}_t\big]\ \mathrm{d}\lambda\,.$$

Then using Proposition 4.5 yields (4.25). □

Consider a put option with strike $K > 0$ and exercise time T_0, i.e., the payoff is given by $(K - F(T_0,T_1))^+$. Similarly to the case of a call option, we obtain the expression in the right hand side of (4.25) for the price of a put option with $u = (\nu + \mathrm{i}\lambda)u_\vartheta$, for $\nu < 0, \lambda \in \mathbb{R}$.

In the following two examples we can make the option-pricing formula in (4.25) more explicit, since for constant instantaneous covariance, but also for Lévy driven ones, the associated generalized Riccati equations are explicitly solvable.

Example 4.8 (Constant volatility)**.** In this example we derive the option-price formula in case of constant instantaneous covariance, i.e. the case that we want to disrupt with our stochastic instantaneous covariance processes. Assume that under the measure $\mathbb{Q}$, the process Y is modeled by

$$Y_t = (\mathcal{A}Y_t + D^{1/2}\sigma D^{1/2}S^*(\cdot)u_0)\,\mathrm{d}t + D^{1/2}\sigma^{1/2}\,\mathrm{d}\tilde{W}_t\,,$$
$$Y_0 = x \in H_w\,, \qquad t \geq 0\,, \tag{4.27}$$

where $\sigma \in \mathcal{L}_2(H_w)$ is self-adjoint and positive, $\mathcal{A}$ and D are as described after equation (4.18), S is the shift semigroup associated with $\mathcal{A}$, and $\tilde{W}$ is a cylindrical Brownian motion under $\mathbb{Q}$. The adjoint S^* of the left shift semigroup on H_w satisfies

$$S^*(t)h(x) = h(0) + h(0)\int_0^{x\wedge t} \frac{1}{w(s)}\,\mathrm{d}s + \int_t^{x\vee t} \frac{h'(s-t)w(s-t)}{w(s)}\,\mathrm{d}s\,.$$

Computing the characteristic function of Y and using similar derivations as in the proof of Proposition 4.7 yield the following result.

Lemma 4.9. *Let $T_0 \leq T_1$, set $\vartheta = T_1 - T_0$, assume that Y is as in* (4.27)*, let $F(T_0,T_1) = \exp\left(\delta_\vartheta Y_{T_0}^\vartheta\right)$, and let $u = (\nu + \mathrm{i}\lambda)u_\vartheta$, for $\nu > 1$, $\lambda \in \mathbb{R}$. It holds*

$$\mathbb{E}_{\mathbb{Q}}[(F(T_0,T_1) - K)^+ \mid \mathcal{F}_t] = \frac{1}{2\pi}\int_{\mathbb{R}} g(\lambda)\,\mathrm{e}^{-\phi^\vartheta(T_0-t,u,0) + \langle \psi^\vartheta(T_0-t,u,0), Y_t\rangle_{H_w^{\mathbb{C}}}}\,\mathrm{d}\lambda\,,$$

where g is as in (4.26) *and $(\phi^\vartheta(\cdot,u,0), \psi^\vartheta(\cdot,u,0))$ is the solution of*

$$\begin{cases} \frac{\partial\phi(t,u)}{\partial t} = -\frac{1}{2}\langle D^{1/2}\sigma D^{1/2}S^*(\vartheta), \psi^\vartheta(t,u)\rangle_{H_w^{\mathbb{C}}} \\ \qquad\qquad + \frac{1}{2}\langle D^{1/2}\sigma D^{1/2}\psi^\vartheta(t,u), \overline{\psi^\vartheta(t,u)}\rangle_{H_w^{\mathbb{C}}}\,, \quad \phi^v(0,u) = 0\,, \\ \frac{\partial\psi^\vartheta(t,u)}{\partial t} = \mathcal{A}^*\psi^\vartheta(t,u)\,, \qquad \psi^\vartheta(0,u) = u\,. \end{cases}$$

Example 4.10 (The operator-valued Barndorff–Nielsen, Shepard model BNS)**.** We assume that under the measure $\mathbb{Q}$, the process (Y,X) is modeled by

$$\mathrm{d}Y_t^\vartheta = \left(\mathcal{A}Y_t + D^{1/2}X_tD^{1/2}S^*(\vartheta)u_0\right)\mathrm{d}t + D^{1/2}X_t^{1/2}\,\mathrm{d}\tilde{W}_t\,, \qquad Y_0^\vartheta = y \in H\,,$$
$$\mathrm{d}X_t = (CY_t + Y_tC^*)\,\mathrm{d}t + \mathrm{d}L_t, \qquad X_0 = x \in \mathcal{H}^+\,.$$

Where we assume that $\mathcal{A}$, D and $\tilde{W}$ are as described after equation (4.18), $C \in L(H)$ and L is an $\mathcal{H}^+$-subordinator, i.e., $t \mapsto L_t$ is almost surely increasing with respect to the cone $\mathcal{H}^+$ and has the characteristic triplet $(b,0,m)$. From Proposition 4.7 we conclude the following lemma.

Lemma 4.11. *Let* $\vartheta = T_1 - T_0$*,* $T_0 \le T_1$*, let* Y^ϑ *be as in* (4.27)*, let* $F(T_0, T_1) = \exp\left(\delta_\vartheta Y^\vartheta_{T_0}\right)$*, and let* $u = (\nu + \mathrm{i}\lambda)u_\vartheta$*, for* $\nu > 1$*,* $\lambda \in \mathbb{R}$*. It holds*

$$\begin{aligned}&\mathbb{E}_{\mathbb{Q}}[(F(T_0,T_1) - K)^+ \mid \mathcal{F}_t] \\ &\quad = \frac{1}{2\pi}\int_{\mathbb{R}} g(\lambda)\,\mathrm{e}^{-\phi^\vartheta(T_0-t,u,0)+\langle \psi_1^\vartheta(T_0-t,u,0),Y_t^\vartheta\rangle_{H_w^{\mathbb{C}}}+\langle \psi_2^\vartheta(T_0-t,u,0),X_t\rangle_{\mathcal{H}^{\mathbb{C}}}}\,\mathrm{d}\lambda\,,\end{aligned}$$

where g *is as in* (4.26) *and* $(\phi^\vartheta(\cdot,u,0), \psi_1^\vartheta(\cdot,u,0), \psi_2^\vartheta(\cdot,u,0))$ *is a solution of*

$$\begin{cases} \dfrac{\partial \phi^\vartheta}{\partial t}(t,u) = \varphi_L(\psi_2^\vartheta(t,u)), \quad 0 < t \le T, \quad \phi^\vartheta(0,u) = 0, & (4.28a) \\ \psi_1^\vartheta(t,u) = u + \mathcal{A}^*\displaystyle\int_0^t \psi_1^\vartheta(s,u)\,\mathrm{d}s, \quad 0 < t \le T, \quad \psi_1^\vartheta(0,u) = u\,, & (4.28b) \\ \dfrac{\partial \psi_2^\vartheta}{\partial t}(t,u) = R_\Upsilon(\psi_1^\vartheta(t,u), \psi_2^\vartheta(t,u)), \quad 0 < t \le T, \quad \psi_2^\vartheta(0,u) = 0. & (4.28c) \end{cases}$$

with $R_\Upsilon(h,u) = B^*(u) - \frac{1}{2}(D^{1/2}h)^{\otimes 2} - D^{1/2}h \otimes D^{1/2}\Upsilon^\vartheta$ *and* $\varphi_L(u) = \langle b, u\rangle - \int_{\mathcal{H}^+\setminus\{0\}}\left(\mathrm{e}^{-\langle \xi, u\rangle} - 1\right) m(\mathrm{d}\xi)$.

4.4 Concluding remarks

In this chapter we studied option-pricing in infinite-dimensional affine stochastic covariance models. As a particular example we considered a geometric affine model for forward curve dynamics formulated in the HJMM framework and derived quasi-explicit formulas for plain vanilla call options on forwards in terms of the solutions to a class of extended generalized Riccati equations. The derivations in this chapter demonstrated that the virtues of the affine class for option-pricing based on Fourier methods is maintained in this infinite-dimensional setting.

- ***On the numerical analysis***

This chapter is part of a working paper in which we also conduct numerical experiments for option-pricing in affine stochastic covariance models. In particular, the numerical feasibility of the pricing formulas (4.25) and the respective complex generalized Riccati equations (4.10) will be analyzed for a number of examples, including such affine models that admit for state-dependent jump intensity in the instantaneous covariance process, i.e. consider also examples that go beyond the discussed cases in Examples 4.8 and 4.10.

CHAPTER 5

Stationary Covariance Regime for Infinite-Dimensional Affine Stochastic Covariance Models

Abstract of the chapter In this chapter we study the long-time behavior of affine processes on positive self-adjoint Hilbert-Schmidt operators. More precisely, for subcritical affine processes we prove the existence of a unique invariant distribution and construct the corresponding stationary affine process. Moreover, we derive explicit rates of the convergence of the underlying transition kernels to the limit distribution in the Wasserstein distance of order $p \in [1, 2]$ and provide explicit formulas for the first two moments of the limit distribution. We apply our results to the study of infinite-dimensional affine stochastic covariance models and introduce the so-called *stationary covariance regime*. In the stationary covariance regime we use a stationary affine process to model the instantaneous covariance process in an infinite-dimensional affine stochastic covariance model. In this context we investigate the behavior of the implied forward volatility smile for large forward dates in the geometric affine stochastic covariance model for forward curve dynamics as introduced in Chapter 4.

This chapter is based on the preprint [67]:
Friesen, M., and Karbach, S.
Stationary Covariance Regime for Affine Stochastic Covariance Models in Hilbert Spaces, 2022, DOI: arXiv.2203.14750.

5.1 Introduction

In this chapter we are concerned with the long-time behavior of affine processes on positive Hilbert-Schmidt operators. The long-time behavior of instantaneous covariance processes play an important role in the calibration of stochastic covariance models, see [1, 105]. Moreover, inspired by the finite-dimensional case in [94], the existence of stationary affine processes on $\mathcal{H}^+$ allows us to introduce infinite-dimensional affine stochastic covariance models in the *stationary covariance regime*. As an application, we study the forward volatility smile of forward-start options for large forward-start dates and relate this to the pricing of a European call option on forwards modeled in the stationary covariance regime. In the following we give a more detailed introduction to the theoretic and applied aspects of studying the long-time behavior of affine processes.

On the long-time behavior of affine processes The long-time behavior of affine processes on the finite dimensional state spaces $(\mathbb{R}_+)^d \times \mathbb{R}^n$ and $\mathbb{S}_d^+$ for integers $d, n \in \mathbb{N}$, is now mostly well-understood. More precisely, based on the representation by strong solutions of stochastic differential equations, ergodicity was studied in [66] for different Wasserstein distances. By using regularity of transition densities with respect to the Lebesgue measure combined with the Meyn-and-Tweedie stability theory the ergodicity in total variation distances has been studied in [6, 63, 85, 84, 115, 65]. Finally, coupling techniques for affine processes are studied in [145, 110]. Unfortunately, these methods implicitly use the dimension of the state-space and hence do not allow for an immediate extension to infinite-dimensional settings. Indeed, for general affine processes on $\mathcal{H}^+$ there does not exist, so far, a pathwise construction. The absence of an infinite-dimensional Lebesgue measure prevents us to effectively use the Meyn-and-Tweedie stability theory (in terms of estimates on the density). Although there exist some extensions of the coupling techniques to infinite dimensional state-spaces (see [109] for measure-valued branching processes), these methods seem to be closely related to the measure-valued structure of the process and hence not suitable for our Hilbert space framework.

The most promising method to study the long-time behavior for affine processes in infinite-dimensional settings is therefore based on the convergence of Fourier-Laplace transforms. The latter one requires, in view of the affine-transform formula (1.3), to study the long-time behavior of the solutions to the generalized Riccati equations ϕ and ψ. For finite-dimensional state spaces, these ideas have been developed in [70, 94, 96, 86, 126, 64]. In these works, the existence of an invariant distribution (as well as weak convergence of transition probabilities) is obtained from Lévy's continuity theorem.

Unfortunately, analogs of Lévy's continuity theorem on Hilbert spaces of infinite dimensions require an additional tightness condition on the transition probabilities (to obtain the existence of a limit distribution and hence invariant measure). For Ornstein-Uhlenbeck processes on Hilbert spaces, this problem can be avoided by taking advantage of their infinite divisibility, see [36]. We treat the class of OU processes on positive Hilbert-Schmidt operators in Example 5.8 below. Apart from this, the long-time behavior of affine processes in infinite-dimensions has not been investigated in a systematic way. This chapter provides a first general treatment of long-time behavior for affine processes on infinite-dimensional Hilbert spaces as state space.

Our methodology for the proof of our main result builds on the ideas taken from [64], where the long-time behavior of affine processes on $\mathbb{S}_d^+$ was studied. Namely, we 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 $u\in\mathcal{H}^+$ and hence the Fourier-Laplace transform of the process (see (1.3)) converges when $t\to\infty$. To overcome the difficulty related to the absence of a full analogue of Lévy's continuity theorem, we utilize the generalized Feller semigroup approach for the process (see Section 2.4.1). More precisely, we provide uniform bounds on the operator norm of the transition semigroup which allows us to prove that $\lim_{t\to\infty}P_t f =: \ell(f)$ has a limit 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 generalized Feller semigroups to show that ℓ has representation $\ell(f)=\int_{\mathcal{H}^+} f(y)\pi(\mathrm{d}y)$. The measure π is the desired unique invariant probability measure. As a byproduct we also obtain weak convergence of transition probabilities in the weak topology on $\mathcal{H}^+$. In the second step we strengthen this convergence by proving estimates on the Wasserstein distance of order $p\in[1,2]$ of the transition probabilities to the invariant measure. In contrast to the finite-dimensional results in [66, 64], our new bounds are dimension-free and explicit. As a consequence, we conclude that the transition probabilities converge weakly to the invariant measure in the norm topology on $\mathcal{H}^+$. Finally, we show that the invariant measure has finite second moments and compute them explicitly.

The Stationary covariance regime Our main motivation for studying the long-time behavior of affine processes on positive self-adjoint Hilbert-Schmidt operators comes from infinite-dimensional stochastic covariance modeling. Motivated by the univariate case in [94] we introduce an affine stochastic covariance model (Y,X) of the form in Definition 3.8 under the *stationary covariance regime*. This is done by replacing a subcritical affine instantaneous covariance process X by its stationary version, i.e. the unique stationary affine process associated with the same admissible parameter set as X. The existence of this stationary process is guaranteed by Corollary 5.5 below.

In Proposition 5.10 we show that the affine stochastic covariance model in the stationary covariance regime again satisfies an affine transform formula, which makes it a tractable model for, e.g. pricing options written on forwards as demonstrated in Section 4.3. As an example, we derive the characteristic function of the operator valued Barndorff–Nielsen-Shepard model from [21] in the stationary covariance regime. This complements the literature on operator-valued BNS type models by their long-time behavior, which was already studied in finite dimensions in [10, 126].
In Section 5.4.2 we consider the geometric affine stochastic covariance model for commodity forward curve dynamics from Section 4.3 and study the pricing of forward-start options in this model. More precisely, we are concerned with the implied forward volatility smile in geometric affine models and show in Proposition 5.12, that the smile converges to the implied spot volatility of a European call option written on a forward contract with dynamics given in the stationary covariance regime. This extends a result in [94, Proposition 5.2] for forward-start options in (univariate) affine stochastic volatility models to an infinite-dimensional setting.

5.1.1 Layout of the chapter

We begin with Section 5.2 where we recall some preliminaries and introduce our setting. In Section 5.3 we present our main results on the existence of a unique invariant measure and the ergodicity in Wasserstein distance. Afterwards, in Section 5.4, we discuss applications of our results in the context of affine stochastic covariance models in Hilbert spaces. The proofs are contained in Section 5.5, which is subdivided into several subsections: We consider the long-time behavior of the solutions of the generalized Riccati equations in Section 5.5.1, prove the existence of a unique invariant measure in Section 5.5.2, derive the explicit convergence rates in Section 5.5.3, show existence of stationary affine processes in Section 5.5.4 and prove the moment formulas of the invariant measure in Section 5.5.5. In Section 5.6 we show a version of Kolmogorov's extension theorem.

5.2 Setting and preliminaries

The effective drifts $\hat{b}$ and $\hat{B}$ We assume that (b, B, m, μ) is an admissible parameter set according to Definition 2.3 and we define the constant and linear *effective drift* terms $\hat{b}$ and $\hat{B}(u)$, for all $u \in \mathcal{H}^+$, as follows:

$$\hat{b} := b + \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \xi \, m(\mathrm{d}\xi), \quad \hat{B}(u) := B^*(u) + \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \langle \xi, u \rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}. \tag{5.1}$$

Note that $\hat{b} \in \mathcal{H}$ and $\hat{B} \in \mathcal{L}(\mathcal{H})$ are well-defined, since by part (a) of Definition 2.3 i) we have

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

i.e. the integral $\int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \xi \, m(\mathrm{d}\xi)$ is well-defined in the Bochner sense. Similarly, it can be see 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 $|\langle \mu(\mathrm{d}\xi), v \rangle| \coloneqq \langle \mu(\mathrm{d}\xi), v^+ \rangle + \langle \mu(\mathrm{d}\xi), v^- \rangle$ and see that $|\langle \mu(\mathrm{d}\xi), v \rangle|$ is a positive measure for all $v \in \mathcal{H}$. We thus have:

$$\begin{aligned} \Big\langle \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \langle \xi, u \rangle \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}, v \Big\rangle &\leq \|u\| \Big(\int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \|\xi\|^{-1} |\langle \mu(\mathrm{d}\xi), v \rangle| \Big) \\ &\leq \|u\| |\langle \mu(\mathcal{H}^+ \cap \{\|\xi\| > 1\}), v \rangle|, \end{aligned}$$

taking the supremum over all $v \in \mathcal{H}$ with $\|v\| = 1$ on both sides proves the boundedness of the map. Note that $\|\mu(\mathcal{H}^+ \cap \{\|\xi\| > 1\})\| < \infty$, since by Definition 2.3 iii) for every $A \in \mathcal{B}(\mathcal{H}^+ \setminus \{0\})$ we have $\mu(A) \in \mathcal{H}^+$ and hence $\|\mu(A)\| < \infty$.

A reminder on the generalized Feller property The results in the previous Chapters 3 and 4, for the most part, built on the semimartingale property of càdlàg affine processes on $\mathcal{H}^+$. In this chapter we are working in the spirit of Chapter 2 again and rely more on the Markovian structure of the affine class (in particular its generalized Feller property). For this reason we briefly recall some important aspects and introduce the setting: Let (b, B, m, μ) be an admissible parameter set and denote by X the unique associated affine process from Theorem 2.8. Moreover, for $x \in \mathcal{H}^+$ we denote by $(p_t(x, \cdot))_{t \geq 0}$ the transition kernels of X and note that for all $t \geq 0$ and $u, x \in \mathcal{H}^+$ the affine transform formula (1.3) can be written as

$$\int_{\mathcal{H}^+} \mathrm{e}^{-\langle \xi, u \rangle} \, p_t(x, \mathrm{d}\xi) = \mathrm{e}^{-\phi(t,u) - \langle x, \psi(t,u) \rangle}, \tag{5.2}$$

for $(\phi(\cdot, u), \psi(\cdot, u))$ the unique solution to the *generalized Riccati equations* (2.8). In Chapter 2, we considered the transition semigroup $(P_t)_{t \geq 0}$ of X given by

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

for bounded measurable functions $f \colon \mathcal{H}^+ \longrightarrow \mathbb{R}$. Note that $(P_t)_{t \geq 0}$ is also well-defined on $B_\rho(\mathcal{H}^+)$ and satisfies for each $f \in B_\rho(\mathcal{H}^+)$ the growth-bound

$$|P_t f(x)| \leq \|f\|_\rho \int_{\mathcal{H}^+} \rho(y) p_t(x, \mathrm{d}y) \leq \|f\|_\rho (1 + K) e^{\omega t} \rho(x), \qquad x \in \mathcal{H}^+. \tag{5.3}$$

This means that $(P_t)_{t\geq 0}$ leaves $B_\rho(\mathcal{H}^+)$ invariant, see also Section 2.4. As before we denote by $\mathcal{H}^+_w$ the space $\mathcal{H}^+$ equipped with the weak topology, and let $C_b(\mathcal{H}^+_w)$ be the space of all bounded and weakly continuous functions $f : \mathcal{H}^+ \longrightarrow \mathbb{R}$. Finally let $\mathcal{B}_\rho(\mathcal{H}^+)$ be the closure of $C_b(\mathcal{H}^+_w)$ in $B_\rho(\mathcal{H}^+)$. It follows from Chapter 2, that $(P_t)_{t\geq 0}$ is positive, leaves $\mathcal{B}_\rho(\mathcal{H}^+_w)$ invariant and satisfies $\lim_{t\to 0+} P_t f(x) = f(x)$ for all $f \in \mathcal{B}_\rho(\mathcal{H}^+_w)$ and $x \in \mathcal{H}^+_w$. From (5.3) we then obtain

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

Which reminds us, that $(P_t)_{t\geq 0}$ is a generalized Feller semigroup on $\mathcal{B}_\rho(\mathcal{H}^+_w)$.

Remark 5.1. 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 [45, 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 $X_t \colon \Omega \to \mathcal{H}$, given by $X_t(\omega) = \omega(t)$ for $\omega \in \Omega$, are measurable. 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 $\mathbb{E}_x[\cdot]$. In Section 5.6 below, we prove a version of the extension theorem that copes with initial distributions beyond delta-distributions.

In the following proposition, we give explicit formulas for the first two moments of the affine process $(X_t)_{t\geq 0}$ with admissible parameter set (b, B, m, μ). We need those explicit versions for Theorem 5.3 and Proposition 5.6 below. We obtain the formulas simply by inserting (2.29)-(2.32) into the formulas (2.102) and (2.103).

Proposition 5.2. *Let $(X_t)_{t\geq 0}$ be the affine process associated with the admissible parameter set (b, B, m, μ). Then for all $v, w \in \mathcal{H}^+$ the following formulas hold true:*

$$\mathbb{E}_x\left[\langle X_t, v\rangle\right] = \int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} v\rangle \,\mathrm{d}s + \langle x, \mathrm{e}^{t\hat{B}} v\rangle \tag{5.4}$$

and

$$\begin{aligned}\mathbb{E}_x\left[\langle X_t, v\rangle\langle X_t, w\rangle\right] = &\Big(\int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} v\rangle \,\mathrm{d}s + \big\langle x, \mathrm{e}^{t\hat{B}} v\big\rangle\Big)\Big(\int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} w\rangle \,\mathrm{d}s + \big\langle x, \mathrm{e}^{t\hat{B}} w\big\rangle\Big)\\ &+ \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, \mathrm{e}^{s\hat{B}} v\rangle\langle \xi, \mathrm{e}^{s\hat{B}} w\rangle \, m(\mathrm{d}\xi)\,\mathrm{d}s\\ &+ \int_0^t \int_0^s \Big\langle \hat{b}, \mathrm{e}^{(s-u)\hat{B}} \int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, \mathrm{e}^{u\hat{B}} v\rangle\langle \xi, \mathrm{e}^{u\hat{B}} w\rangle \,\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big\rangle \,\mathrm{d}u\,\mathrm{d}s\\ &+ \int_0^t \Big\langle x, \mathrm{e}^{(t-s)\hat{B}} \int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, \mathrm{e}^{s\hat{B}} w\rangle\langle \xi, \mathrm{e}^{s\hat{B}} v\rangle \,\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big\rangle \,\mathrm{d}s.\end{aligned} \tag{5.5}$$

5.3 Stationary affine processes and ergodicity in Wasserstein distance

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

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

For $\nu_1,\ \nu_2 \in \mathcal{M}(\mathcal{H}^+)$ we call a probability measure G, defined 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 every $p \in [1,\infty)$ 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

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

For an introduction to Wasserstein distances we refer to [143, Section 6].
Now, let (b,B,m,μ) be an admissible parameter set and denote by $\sigma(\hat{B})$ the spectrum of the operator $\hat{B}$ in (5.1). We introduce the following crucial assumption:

Assumption $\mathcal{D}$. The spectral bound $s(\hat{B}) \coloneqq \sup\left\{\Re(\lambda)\colon \lambda \in \sigma(\hat{B})\right\}$ of $\hat{B}$ is strictly negative, i.e. $s(\hat{B}) < 0$.

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 $\mathcal{D}$ 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 the series representation $\mathrm{e}^{t\hat{B}} \coloneqq \sum_{n=0}^{\infty} \frac{(t\hat{B})^n}{n!}$, where the convergence 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 [55, Chapter I, Section 3], and hence it follows that the spectral bound $s(\hat{B})$ coincides with the *growth bound* of $(\mathrm{e}^{t\hat{B}})_{t\geq 0}$, see [55, Corollary 4.2.4], i.e.

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

Therefore, whenever Assumption $\mathcal{D}$ is satisfied, there exists a $M \geq 1$ and $\delta > 0$ such that

$$\|\mathrm{e}^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})} \leq M\, \mathrm{e}^{-\delta t}, \tag{5.6}$$

in particular we could choose $\delta = -s(\hat{B})$. In the following theorem we assert the existence of a unique invariant measure π of $(p_t(x,\cdot))_{t\geq 0}$, the transition kernels of a subcritical affine process on $\mathcal{H}^+$. Moreover, we derive explicit convergence rates for $p_t(x,\cdot) \to \pi$ as $t \to \infty$ in Wasserstein distance of order $p \in [1,2]$.

Theorem 5.3. *Let (b, B, m, μ) be an admissible parameter set such that Assumption $\mathcal{D}$ is satisfied. 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 holds 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

$$\int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,x\rangle}\, \pi(\mathrm{d}x) = \exp\Big(-\int_0^\infty F(\psi(s,u))\, \mathrm{d}s\Big), \quad u \in \mathcal{H}^+, \tag{5.7}$$

where F and $\psi(s,u)$ are as in (2.6a) and (2.8b). 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

$$W_p(p_t(x,\cdot),\pi) \leq C_1\, \mathrm{e}^{-\delta t}\Big(\|x\| + \big(\int_{\mathcal{H}^+} \|y\|^p\, \pi(\mathrm{d}y)\big)^{1/p}\Big) \tag{5.8}$$

$$+ C_2\, \mathrm{e}^{-\delta/2t}\Big(\|x\|^{1/2} + \big(\int_{\mathcal{H}^+} \|y\|^{p/2}\, \pi(\mathrm{d}y)\big)^{1/p}\Big), \tag{5.9}$$

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 (5.6). In particular, we have $p_t(x,\cdot) \Rightarrow \pi$ as $t \to \infty$.

Remark 5.4. i) Note that for a locally compact and second countable Hausdorff space, in particular for finite-dimensional normed spaces, every probability measure defined on the Borel-σ-algebra is regular.

ii) The last assertion in Theorem 5.3 i) 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, albeit $\mathcal{H}^+_{\mathrm{w}}$ in the infinite-dimensional case is not locally compact. We see that the inner-regularity is a non-trivial property of the invariant measure and we actually use it in the proof of Corollary 5.5 below.

iii) For the case $p = 1$ we can compare the convergence rates obtained in Theorem 5.3 ii) with the ones in [64, Theorem 2.9] for the state space $\mathbb{S}^+_d$ ($d \in \mathbb{N}$) i.e. $H = \mathbb{R}^d$ in our case. We see that instead of the square-root of the dimension $d \in \mathbb{N}$ as it appears in the convergence rate in [64, equation 2.12], we have the additional term (5.9) which also converges exponentially fast as $t \to \infty$, but with the exponential factor $-\delta/2$ instead of $-\delta$. However, the convergence rates here do not depend on the dimension of the state-space, in particular they hold true in infinite dimensions.

As a corollary from Theorem 5.3 i), which ensures the existence of an invariant inner-regular measure π on $\mathcal{B}(\mathcal{H}^+_{\mathrm{w}})$, we assert the existence of a stationary process with stationary measure π. The only assertion that is left to prove here is, that we can start an affine process with transition kernels $p_t(x,\cdot)$ at distribution π instead of δ_x:

Corollary 5.5. *There exists a process $(X^\pi_t)_{t\geq 0}$ on $\mathcal{H}^+$ with transition kernels $(p_t(x,\cdot))_{t\geq 0}$ such that the distribution of X^π_t equals π for all $t \geq 0$.*

Note here that the p-th absolute moment of π shows up in the convergence rate in (5.8), where we implicitly assumed that these terms are finite. That this is indeed the case is part of the next proposition, where we also give explicit formulas for the first two moments of the invariant measures π.

Proposition 5.6. *Under the same conditions as in Theorem 5.3 and by denoting the unique invariant measure of $(p_t(x,\cdot))_{t\geq 0}$ by π we have $\int_{\mathcal{H}^+} \|y\|^2\,\pi(\mathrm{d}y) < \infty$,*

$$\lim_{t\to\infty} \mathbb{E}_x\left[X_t\right] = \int_{\mathcal{H}^+} y\,\pi(\mathrm{d}y) = \int_0^\infty \mathrm{e}^{s\hat{B}}\Big(b + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \xi\,m(\mathrm{d}\xi)\Big)\,\mathrm{d}s, \tag{5.10}$$

and

$$\begin{aligned}\lim_{t\to\infty} \mathbb{E}_x\left[X_t\otimes X_t\right] &= \int_{\mathcal{H}^+} y\otimes y\,\pi(\mathrm{d}y)\\ &= \int_0^\infty \big(\mathrm{e}^{s\hat{B}^*}\hat{b}\big)^{\otimes 2}\,\mathrm{d}s + \int_0^\infty\int_{\mathcal{H}^+\setminus\{0\}} \big(\mathrm{e}^{s\hat{B}^*}\xi\big)^{\otimes 2}\,m(\mathrm{d}\xi)\,\mathrm{d}s\\ &\quad + \int_0^\infty\int_0^s\int_{\mathcal{H}^+\setminus\{0\}} \big(\mathrm{e}^{u\hat{B}^*}\xi\big)^{\otimes 2}\langle\hat{b}, \mathrm{e}^{(s-u)\hat{B}}\,\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\rangle\,\mathrm{d}u\,\mathrm{d}s.\end{aligned} \tag{5.11}$$

Remark 5.7. It is well known that convergence in Wasserstein distance of order $p \in [1,\infty)$ implies weak convergence and the convergence of the p-th absolute moment, see [143, Theorem 6.9]. However, we note that the proof of (5.10), given in Section 5.5.5, does not depend on the convergence in Wasserstein distance of order $p = 2$ as established by Theorem 5.3 ii). Instead, we only use the generalized Feller property of the transition semigroups $(P_t)_{t\geq 0}$ together with Proposition 5.2.

Example 5.8 (Lévy driven Ornstein-Uhlenbeck processes)**.** Let m be a Lévy measure on $\mathcal{B}(\mathcal{H}^+ \setminus \{0\})$ with finite second moment and $b \in \mathcal{H}^+$ such that Definition 2.3 ii) 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 Definition 2.3 iv) 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 $t \geq 0$. Hence, [106, Theorem 1] implies that B satisfies Definition 2.3 iv). Thus, the tuple $(b, m, B, 0)$ is an admissible parameter set according to Definition 2.3 and the associated affine process $(X_t)_{t\geq 0}$ becomes an Ornstein-Uhlenbeck process driven by a $\mathcal{H}^+$-valued Lévy process $(L_t)_{t\geq 0}$ with characteristics $(b, 0, m)$, i.e.

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

see Lemma 3.4. Since $\sigma(B) = \sigma(G) + \sigma(G)$, see, e.g., [129], 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 $\mathcal{D}$ is satisfied. This provides an explicit and simple sufficient criterion for the OU process $(X_t)_{t\geq 0}$ to be subcritical. By Theorem 5.3 there exists a unique invariant measure π such that

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

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

$$\varphi_L(u) = \langle b, u\rangle - \int_{\mathcal{H}^+\setminus\{0\}} \big(\mathrm{e}^{-\langle \xi,u\rangle} - 1 + \langle \chi(\xi), u\rangle\big)\,m(\mathrm{d}\xi), \quad u \in \mathcal{H}^+. \tag{5.12}$$

Existence and uniqueness of invariant measures for Ornstein-Uhlenbeck processes were studied in [36], where a similar result follows under the weaker log-moment condition on the Lévy measure m. Following Proposition 5.6 the stronger second moment assumption in our case allows us to deduce explicit formulas for the first and second moments of π. Indeed, setting $\mu = 0$ and $B(u) = Gu + uG^*$ in (5.10) and (5.11) gives

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

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

5.4 The stationary covariance regime

In this section we discuss applications of our results on the long-time behavior of affine processes on positive Hilbert-Schmidt operators in the context of affine stochastic covariance models. In Section 5.4.1 we introduce infinite-dimensional affine stochastic covariance model in the *stationary covariance regime* and derive a stationary version of the affine transform formula. In Section 5.4.2 we come back to the geometric affine stochastic covariance model from Section 4.3 and show that in this model the implied volatility of forward-start options written on forwards can be related to the implied volatility of plain vanilla options on forwards with dynamics modeled under the stationary covariance regime.

5.4.1 Affine models in the stationary covariance regime

Let $(X_t)_{t\geq 0}$ be an affine process on $\mathcal{H}^+$ with initial value $X_0 = x$ and associated admissible parameter set (b, B, m, μ) such that Assumptions $\mathcal{A}$ is satisfied. We fix $y, \Upsilon \in H$, $x \in \mathcal{H}^+$, and let $D \in \mathcal{L}_2(H)$ be self-adjoint and positive. As before, we assume that $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H. Following Chapter 4, we consider the process $(Y, X, \Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, the stochastically weak solution of (4.1), with parameters $(b, B, m, \nu, D, \Upsilon, \mathcal{A})$ and initial value $(Y_0, X_0) = (x, y)$. For notational brevity we define $G\colon \mathcal{H} \to H$ as $G(x) \coloneqq D^{1/2}xD^{1/2}\Upsilon$, i.e. $(Y_t)_{t\geq 0}$ is the unique (mild) solution to the following stochastic differential equation on some Hilbert space $(H, \langle\cdot,\cdot\rangle_H)$:

$$Y_t = S(t)y + \int_0^t S(t-s)G(X_s)\,\mathrm{d}s + \int_0^t S(t-s)D^{1/2}X_s^{1/2}\,\mathrm{d}W_s, \quad t \geq 0, \tag{5.13}$$

where $(W_t)_{t\geq 0}$ is a cylindrical Brownian-motion, independent of $(X_t)_{t\geq 0}$. We consider the joint process $(Y_t, X_t)_{t\geq 0}$ as a stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}_x) \coloneqq (\Omega^1 \times \Omega^2, (\mathcal{F}^1 \otimes \mathcal{F}^2), (\mathcal{F}_t^1 \otimes \mathcal{F}_t^2)_{t\geq 0}, \mathbb{Q} \otimes \mathbb{P}_x)$, where $(\Omega^2, \mathcal{F}^2, (\mathcal{F}_t^2)_{t\geq 0}, \mathbb{P}_x)$ denotes the filtered probability space accommodating the affine process $(X_t)_{t\geq 0}$, see also Remark 5.1, and $(\Omega^1, \mathcal{F}^1, (\mathcal{F}_t^1)_{t\geq 0}, \mathbb{Q})$ is another filtered probability space, that carries a cylindrical Brownian Motion $W\colon [0,\infty) \times \Omega \to H$ and the solution process $(Y_t)_{t\geq 0}$ such that $\mathbb{Q}(Y_0 = y) = 1$, see Section 3.2.2.
From now on, we write $(Y_t^y)_{t\geq 0}$ where the superscript y indicates the initial value of the process $(Y_t)_{t\geq 0}$ (not to be confused with the notation Y^ϑ in (4.24)). Moreover, we denote the expectation with respect to the product measure $\mathbb{Q}_x$ by $\mathbb{E}_x[\cdot]$ and denote by π the unique invariant measure π for $(p_t(x,\cdot))_{t\geq 0}$, the existence of which is guaranteed by Theorem 5.3. Corollary 5.5 ensures the existence of the unique stationary affine process$(X_t^\pi)_{t\geq 0}$ associated with (b, B, m, μ) and, inspired by [94, Section 3], we introduce the following terminology.

Definition 5.9. If there exists a mild solution $(\tilde{Y}_t)_{t\geq 0}$ of (5.13) for $y=0$ and with $(X_t)_{t\geq 0}$ replaced by the process $(X_t^{\pi})_{t\geq 0}$, then we call the joint process $(\tilde{Y}_t, X_t^{\pi})_{t\geq 0}$, defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}\otimes\mathbb{P}_{\pi})$ an affine stochastic covariance model on H in the *stationary covariance regime.* Moreover, we write $\mathbb{Q}_{\pi}=\mathbb{Q}\otimes\mathbb{P}_{\pi}$ and denote the expectation with respect to $\mathbb{Q}_{\pi}$ by $\mathbb{E}_{\pi}[\cdot]$.

Let $(Y_t^y, X_t)_{t\geq 0}$ be as in (5.13) and for simplicity set G to zero. It follows from Theorem 3.14 that the stochastic covariance model $(Y_t^y, X_t)_{t\geq 0}$ satisfies the affine transform formula (3.27). In the following proposition we give an analogous affine transform formula for this model in the stationary covariance regime:

Proposition 5.10. *Assume that $(Y_t^y, X_t)_{t\geq 0}$ is an affine stochastic covariance model satisfying the assumptions above (with $G=0$) 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 $u=(u_1,u_2)\in \mathrm{i}H\times\mathcal{H}^+$ we have*

$$\mathbb{E}_{\pi}\left[\mathrm{e}^{\langle \tilde{Y}_t, u_1\rangle-\langle X_t^{\pi}, u_2\rangle}\right]=\mathrm{e}^{-\Phi(t,u_1,u_2)-\int_0^{\infty}F(\psi_2(s,0,\psi_2(t,u_1,u_2)))\,\mathrm{d}s},\ t\in[0,T], \quad (5.14)$$

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

$$\begin{cases} \dfrac{\partial\Phi}{\partial t}(t,u)=F(\psi_2(t,u)), & \Phi(0,u)=0, \quad (5.15\text{a})\\ \psi_1(t,u)=u_1-\mathrm{i}\mathcal{A}^*\left(\mathrm{i}\displaystyle\int_0^t\psi_1(s,u)\,\mathrm{d}s\right), & \psi_1(0,u)=u_1, \quad (5.15\text{b})\\ \dfrac{\partial\psi_2}{\partial t}(t,u)=R(\psi_1(t,u),\psi_2(t,u)), & \psi_2(0,u)=u_2, \quad (5.15\text{c})\end{cases}$$

where F and R are as in (2.6a) and (3.23), and $(\mathcal{A}^*, D(\mathcal{A}^*))$ denotes the adjoint operator of the generator $(\mathcal{A}, D(\mathcal{A}))$ of $(S(t))_{t\geq 0}$.

Proof. Let $T\geq 0$ and let $(Y_t^y)_{t\in[0,T]}$ be the mild solution to (5.13) on $[0,T]$ satisfying the assumptions above. From Theorem 3.14 in Chapter 3 we recall the following affine transform formula for the mixed Fourier-Laplace transform of (Y_t^y, X_t) for $t\in[0,T]$ and $u=(u_1,u_2)\in\mathrm{i}H\times\mathcal{H}^+$:

$$\mathbb{E}_x\left[\mathrm{e}^{\langle Y_t^y,u_1\rangle-\langle X_t,u_2\rangle}\right]=\mathrm{e}^{-\Phi(t,u_1,u_2)-\langle x,\psi_2(t,u_1,u_2)\rangle},\quad x\in\mathcal{H}^+, \quad (5.16)$$

where $\Phi(\cdot,u_1,u_2)$ and $\psi_2(\cdot,u_1,u_2)$ are the unique strong solutions to (5.15a) and (5.15c), respectively, and $\psi_1(\cdot,u_1,u_2)$ is the unique mild solution of (5.15b). Note that for every $\mathcal{F}^2$-measurable and bounded function f we have

$$\int_{\Omega_2} f(\omega_2)\ \mathrm{d}\mathbb{P}_x(\omega_2)=\int_{\mathcal{H}^+}\left(\int_{\Omega_2}f(\omega_2)\ \mathrm{d}\mathbb{P}_x(\omega_2)\right)\pi(\mathrm{d}x).$$

From this and (5.16) we conclude

$$\begin{aligned}\mathbb{E}_\pi\left[\mathrm{e}^{\langle \tilde{Y}_t,u_1\rangle-\langle X_t^\pi,u_2\rangle}\right] &= \int_{\Omega_2}\Big(\int_{\Omega_1}\mathrm{e}^{\langle \tilde{Y}_t(\omega_1,\omega_2),u_1\rangle_H-\langle X_t^\pi(\omega_2),u_2\rangle}\,\mathrm{d}\mathbb{Q}(\omega_1)\Big)\,\mathrm{d}\mathbb{P}_\pi(\omega_2)\\ &= \int_{\mathcal{H}^+}\mathbb{E}_x\left[\mathrm{e}^{\langle Y_t,u_1\rangle_H-\langle X_t,u_2\rangle}\right]\pi(\mathrm{d}x)\\ &= \int_{\mathcal{H}^+}\mathrm{e}^{-\Phi(t,u_1,u_2)-\langle x,\psi_2(t,u_1,u_2)\rangle}\,\pi(\mathrm{d}x). \qquad (5.17)\end{aligned}$$

From Theorem 5.3 i) it then follows that

$$\int_{\mathcal{H}^+}\mathrm{e}^{-\langle x,\psi_2(t,u_1,u_2)\rangle}\,\pi(\mathrm{d}x)=\exp\left(-\int_0^\infty F\big(\psi(s,\psi_2(t,u_1,u_2))\big)\,\mathrm{d}s\right), \qquad (5.18)$$

where $\psi(\cdot,u)$ denotes the unique solution of (2.8b). From (5.15c) and the definition of R in (3.23) we see that $\psi_2(t,0,u_2)=\psi(t,u_2)$ for every $u_2\in\mathcal{H}^+$, hence multiplying both sides of (5.18) by $\mathrm{e}^{-\Phi(t,u_1,u_2)}$ together with (5.17) yields the desired formula (5.14). □

Recall the operator-valued BNS stochastic covariance model in Section 3.4.1, i.e. (Y,X) is given by (3.36) with B and $(L_t)_{t\geq 0}$ being as in Example 5.8. The operator-valued BNS model falls into the class (5.13) with $G=0$ and we note that for the $\mathcal{H}^+$-valued Ornstein-Uhlenbeck process $(X_t)_{t\geq 0}$ we already proved the existence of a unique invariant measure π of $(X_t)_{t\geq 0}$ in Example 5.8. Hence, we may consider the *operator-valued BNS model in the stationary covariance regime*, denote it by $(\tilde{Y}_t,X_t^\pi)_{t\geq 0}$ and obtain the following example and application of Proposition 5.10:

Example 5.11. The operator-valued BNS model in the stochastic covariance regime satisfies

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

for every $(u_1,u_2)\in \mathrm{i}H\times\mathcal{H}^+$, where φ_L is given by (5.12) and $\psi(t,u_1,u_2)$ is explicitly given by

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

Note that if $G\neq 0$, then $\psi(t,u_1,u_2)$ would admit an additional inhomogeneous term of the form $\int_0^t\mathrm{e}^{(s-\tau)B^*}G^*(S^*(\tau)u_1)\,\mathrm{d}\tau$, where G^* denotes the adjoint of G.

5.4.2 Forward-start options on forwards

In this section we follow up on Section 4.3 and consider so called *forward start options* written on commodity forwards. For this type of options we study the long-time behavior of the implied *forward volatility smile* in the geometric affine stochastic covariance model from Section 4.3 and prove an intimate connection to the pricing of plain vanilla options on forwards modeled under the stationary covariance regime in Proposition 5.12 below.

We define a *forward-start option* with forward-start date $\tau \geq 0$, forward maturity T and strike K written on a forward with maturity date $\tau + \hat{T}$, denoted by $F(\cdot, \tau + \hat{T})$, as an European option with pay-off at time $\tau + T$ given by

$$\left(\frac{F(\tau + T, \tau + \hat{T})}{F(\tau, \tau + \hat{T})} - K \right)^{+}, \tag{5.19}$$

see, e.g. [41, equation 11]. Forward-start options are contracts on the relative (or absolute) price difference of a forward contract at two times, τ and $\tau + T$, in the future. In practice, it is used to price future volatility of the underlying asset. Forward-start options are common in commodity forward markets and more complex derivatives such as *Cliquet options* are building up on these see [41]. Forward-start options on stocks are discussed in, e.g. [81, 82, 101, 94].

Let $0 \leq T \leq \hat{T}$ and as above, we denote by $F(T, \hat{T})$ the forward price at time T with maturity date $\hat{T}$ of some underlying spot commodity. Let $(Y_t^y, X_t)_{t \geq 0}$ be an affine stochastic covariance as in (5.13) on the Hilbert space H_β (see Section 4.3) and let $(S(t))_{t \geq 0}$ be the left-shift semigroup on H_β. Moreover, we assume that the drift G is given by $G = -\frac{1}{2} D^{1/2} x D^{1/2} S^*(\cdot) u_0$. For $0 \leq T \leq \hat{T}$ we set

$$F(T, \hat{T}) := \exp\left(\delta_{\hat{T}-T} Y_T^y\right) = \exp\left(\langle Y_T^{y, \hat{T}-T}, u_{\hat{T}-T} \rangle_\beta\right), \tag{5.20}$$

where we use the notations of Section 4.3. Note that by this choice for G the drift condition (4.22) is satisfied and hence we model directly under the risk-neutral measure $\tilde{\mathbb{Q}}$, i.e. $(F(T, \hat{T}))_{0 \leq T \leq \hat{T}}$ is a $\tilde{\mathbb{Q}}$-martingale, see Lemma 4.6.

Next, we define the *implied forward volatility smile* of the model (5.20). We denote the price of a forward-start option with pay-off (5.19) by $C_{\text{fwd}}(\tau, T, \hat{T}, K)$. Next, as a reference model for the forward prices $F(T, \hat{T})$ we choose *Black's model* in [26] and denote the forward prices in this model by $F^{\text{B}}(T, \hat{T})$. We assume that the following spot-forward relation holds:

$$F^{\text{B}}(T, \hat{T}) = s_T \, \mathrm{e}^{r(\hat{T}-T)}, \quad 0 \leq T \leq \hat{T}, \tag{5.21}$$

where $r \geq 0$ denotes the risk-free interest rate and $(s_t)_{t \geq 0}$ denotes the spot price process of the underlying commodity, which is given by a geometric Brownian motion with volatility parameter σ.

We denote by $C^{\mathrm{B}}_{\mathrm{fws}}(\tau, T, \hat{T}, K, \sigma)$ the price of a forward-start option with identical pay-off function as in (5.19) in Black's model and define the *implied forward volatility smile* $\sigma(\tau, T, \hat{T}, K)$ as the solution of $C^{\mathrm{B}}_{\mathrm{fws}}(\tau, T, \hat{T}, K, \sigma(\tau, T, \hat{T}, K)) = C_{\mathrm{fwd}}(\tau, T, \hat{T}, K)$. In the following proposition we show that $\sigma(\tau, T, \hat{T}, K)$ exists for all $\tau, K \geq 0$ and study its long-time behavior as $\tau \to \infty$:

Proposition 5.12. *Let $0 \leq T \leq \hat{T}$ and $\tau \geq 0$. Denote by $F(\tau + T, \tau + \hat{T})$ the forward price at time $\tau + T$ with maturity date $\tau + \hat{T}$ given by* (5.20) *with $(Y^y_t, X_t)_{t \geq 0}$ specified as above. We write $(\tilde{Y}_t, X^\pi_t)_{t \geq 0}$ for the model in the stationary covariance regime, defined in Definition 5.9, and set*

$$\tilde{F}(T, \hat{T}) \coloneqq \delta_{\hat{T}-T} \exp\big(\tilde{Y}_T\big) = \exp\big(\langle \tilde{Y}^{\hat{T}-T}_T, u_{\hat{T}-T} \rangle_\beta\big), \quad 0 \leq T \leq \hat{T}.$$

Then for all $\tau, K \geq 0$ the implied forward volatility smile $\sigma(\tau, T, \hat{T}, K)$ exists and we have

$$\lim_{\tau \to \infty} \sigma(\tau, T, \hat{T}, K) = \tilde{\sigma}(T, \hat{T}, K), \tag{5.22}$$

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

Proof. First, we show a specific relation between the price of a European call option $C^{\mathrm{B}}\big(T, \hat{T}, K, \sigma\big)$, where $0 \leq T \leq \hat{T}$ and $K \geq 0$, and the forward-start call option $C^{\mathrm{B}}_{\mathrm{fws}}\big(\tau, T, \hat{T}, K, \sigma\big)$ with forward start date $\tau \geq 0$ in Black's model. Namely, let $\mathbb{Q}$ denote the unique risk-neutral measure in Black's model and recall $C^{\mathrm{B}}_{\mathrm{fws}}\big(\tau, T, K, \sigma\big)$, the price of a forward-start option at time zero with forward-start date τ written on the forward with maturity $\hat{T}$. Inserting (5.21) into the pay-off function and by risk-neutral pricing we have

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

It is known, see [94], that in the Black-Scholes model the forward-start call option and the European call option satisfy

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

We set $K' = K + rT$ and write $C^{\mathrm{BS}}\left(T, K', \sigma\right) \coloneqq \mathrm{e}^{-r(\tau+T)}\,\mathrm{e}^{-rT} \mathbb{E}_{\mathbb{Q}}\left[\big(s_T - \mathrm{e}^{K'}\big)^+\right]$, whenever the underlying spot process $(s_t)_{t \geq 0}$ is given by the Black-Scholes model.

We then see that

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

From this and the definition of the implied forward volatility smile $\sigma(\tau,T,\hat{T},K)$ we have

$$\begin{aligned} C^{\mathrm{BS}}\left(T,K',\sigma(\tau,T,\hat{T},K)\right) &= \mathrm{e}^{r(\tau+T)}C^{\mathrm{B}}_{\mathrm{fws}}\left(\tau,T,\hat{T},K,\sigma(\tau,T,\hat{T},K)\right) \\ &= \mathrm{e}^{r(\tau+T)}C_{\mathrm{fws}}\left(\tau,T,\hat{T},K\right). \end{aligned} \tag{5.23}$$

Next, we compute the right-hand side of (5.23). 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 $u_t\in H_\beta$, see (4.17). Moreover, we denote the expectation with respect to the pricing measure $\tilde{\mathbb{Q}}$ by $\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\cdot\right]$ (here we suppress the initial value x compared to $\mathbb{Q}_x$ above).
The payoff function of the forward-start option is given by (5.19) and by inserting our model (5.20), we obtain

$$\frac{F(\tau+T,\tau+\hat{T})}{F(\tau,\tau+\hat{T})} = \exp\left(\langle Y^{y,\hat{T}-T}_{\tau+T},u_{\hat{T}-T}\rangle_\beta - \langle Y^{y,\hat{T}}_{\tau},u_{\hat{T}}\rangle_\beta\right). \tag{5.24}$$

First, we note that left-shift $S(T)$ satisfies $\langle S(T)Y^{y,\hat{T}-T}_\tau,u_{\hat{T}-T}\rangle_\beta=\langle Y^{y,\hat{T}}_\tau,u_{\hat{T}}\rangle_\beta$. Second, writing $G^{\hat{T}-T}(x)=-\frac{1}{2}D^{1/2}xD^{1/2}S^*(\hat{T}-T)u_0$ for $x\in\mathcal{H}^+$ we see that

$$\begin{aligned} \langle S(T)Y^{y,\hat{T}-T}_\tau,u_{\hat{T}-T}\rangle_\beta &= \langle S(T+\tau)y,u_{\hat{T}-T}\rangle_\beta \\ &\quad+\Big\langle\int_0^\tau S(T+\tau-s)G^{\hat{T}-T}(X_s)\,\mathrm{d}s,u_{\hat{T}-T}\Big\rangle_\beta \\ &\quad+\Big\langle\int_0^\tau S(T+\tau-s)X_s^{1/2}\,\mathrm{d}\tilde{W}_s,u_{\hat{T}-T}\Big\rangle_\beta, \end{aligned}$$

Thus, the difference in the exponent on the right-hand side of (5.24) satisfies

$$\begin{aligned} \langle Y^y_{\tau+T},u_{\hat{T}-T}\rangle_\beta-\langle Y^y_\tau,u_{\hat{T}}\rangle_\beta &= \Big\langle\int_\tau^{T+\tau} S(T+\tau-s)G^{\hat{T}-T}(X_s)\,\mathrm{d}s,u_{\hat{T}-T}\Big\rangle_\beta \\ &\quad+\Big\langle\int_\tau^{T+\tau} S(T+\tau-s)X_s^{1/2}\,\mathrm{d}\tilde{W}_s,u_{\hat{T}-T}\Big\rangle_\beta. \end{aligned} \tag{5.25}$$

Now, we see that by the independent increments property and the Markov property of $(X_t)_{t\geq 0}$ the sum of the integrals inside the inner products on the right-hand side of (5.25) has the same distribution as Y_T^0 given X_τ, i.e. the same as $Y_T^0 = \int_0^T S(T-s)G(X_{\tau+s})\,\mathrm{d}s + \int_0^T S(T-s)X_{\tau+s}^{1/2}\,\mathrm{d}\tilde{W}_s$. We conclude that also

$$\left(\mathrm{e}^{\langle Y^y_{\tau+T}, u_{\hat{T}-T}\rangle_\beta - \langle Y^y_\tau, u_{\hat{T}}\rangle_\beta} - \mathrm{e}^K\right)^+ \quad \text{and} \quad \mathbb{E}_{\tilde{\mathbb{Q}}}\left[\left(\mathrm{e}^{\langle Y^0_T, u_{\hat{T}-T}\rangle_\beta} - \mathrm{e}^K\right)^+ |\, X_\tau\right],$$

have the same distribution. Moreover, we note that $T \mapsto \mathrm{e}^{\langle Y^0_T, u_{\hat{T}-T}\rangle_\beta}$ given X_τ is a martingale as well, since the drift condition (4.22) still holds for X_t replaced by $X_{t+\tau}$. Hence by risk-neutral pricing, we obtain

$$\begin{aligned} C_{\mathrm{fws}}\left(\tau, T, \hat{T}, K\right) &= \mathrm{e}^{-r(\tau+T)}\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\left(\frac{F(\tau+T, \tau+\hat{T})}{F(\tau, \tau+\hat{T})} - \mathrm{e}^K\right)^+\right] \\ &= \mathrm{e}^{-r(\tau+T)}\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\left(\mathrm{e}^{\langle Y^y_{\tau+T}, u_{\hat{T}-T}\rangle_\beta - \langle Y^y_\tau, u_{\hat{T}}\rangle_\beta} - \mathrm{e}^K\right)^+\right] \\ &= \mathrm{e}^{-r(\tau+T)}\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\left(\mathrm{e}^{\langle Y^0_T, u_{\hat{T}-T}\rangle_\beta} - \mathrm{e}^K\right)^+ |\, X_\tau\right]\right], \end{aligned} \tag{5.26}$$

and we therefore conclude that the left-hand side of (5.23) is given by

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

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

$$\begin{aligned} \lim_{\tau\to\infty} C^{\mathrm{BS}}\left(T, K', \sigma(\tau, T, \hat{T}, K)\right) &= \mathbb{E}_{\tilde{\mathbb{Q}}}\left[(\mathrm{e}^{\langle \tilde{Y}_T, u_{\hat{T}-T}\rangle_\beta} - \mathrm{e}^K)^+\right] \\ &= \mathbb{E}_{\tilde{\mathbb{Q}}}\left[\left(\tilde{F}(T, \hat{T}) - \mathrm{e}^K\right)^+\right]. \end{aligned} \tag{5.27}$$

The term $\mathbb{E}_{\tilde{\mathbb{Q}}}\left[\left(\tilde{F}(T, \hat{T}) - \mathrm{e}^K\right)^+\right]$ on the right-hand side of (5.27) is precisely the price of an European call option remunerated by e^{rT} and we have

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

from which we conclude (5.22), since equation (5.27) has a unique solution in terms of the Black-Scholes implied volatility. □

5.5 Proof: Stationarity and ergodicitiy

Throughout this section we assume that (b, B, m, μ) is an admissible parameter set according to Definition 2.3. We denote by $(X_t)_{t\geq 0}$ the unique affine process associated with (b, B, m, μ) given by Theorem 2.8 and the associated family of transition kernels by $(p_t(x,\cdot))_{t\geq 0}$. Moreover, we set $P_t f := \int_{\mathcal{H}} f(\xi)\, p_t(\cdot, \mathrm{d}\xi)$ for all measurable functions f such that the integral exists, see also the reminder in Section 5.2.

5.5.1 Some properties of the generalized Riccati equations

In this section we consider the long-time behavior of the solutions $\phi(\cdot, u)$ and $\psi(\cdot, u)$ of the generalized Riccati equations in (2.8a)-(2.8b). We recall from Section 2.3 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 (2.8b). Given $\psi(\cdot, u)$ we solve the first equation (2.8a) by mere integration and obtain $\phi(\cdot, u) \in C^1(\mathbb{R}^+, \mathbb{R}^+)$ given by

$$\phi(t,u) = \int_0^t F(\psi(s,u))\ \mathrm{d}s.$$

This means that we can write the affine transform formula (5.2) as

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

Moreover, we recall that the unique solution $\psi(\cdot, u)$ to (2.8b) satisfies the flow equation:

$$\psi(t+s, u) = \psi(t, \psi(s,u)). \tag{5.28}$$

In the next lemma we show that F and R are continuous functions on $\mathcal{H}^+$ and grow at most quadratically, which is a slight adaption of Lemma 2.6:

Lemma 5.13. *Let (b, B, m, μ) be an admissible parameter set according to Definition 2.3 and let F and R be given by* (2.6a) *and* (2.6b), *respectively. Then F and R are continuous on $\mathcal{H}^+$ and for all $u \in \mathcal{H}^+$ we have*

$$|F(u)| \leq \Big(\|b\| + \int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^2\, m(\mathrm{d}\xi)\Big)\left(\|u\| + \|u\|^2\right), \tag{5.29}$$

and

$$\|R(u)\| \leq \left(\|B\|_{\mathcal{L}(\mathcal{H})} + \|\mu(\mathcal{H}^+ \setminus \{0\})\|\right)\left(\|u\| + \|u\|^2\right). \tag{5.30}$$

Proof. Note that for all $\xi, u \in \mathcal{H}^+$ we have

$$\begin{aligned}\left|\mathrm{e}^{-\langle \xi, u\rangle} - 1 + \langle \chi(\xi), u\rangle\right| &\leq \frac{1}{2}|\langle \xi, u\rangle|^2 \mathbf{1}_{\{\|\xi\|\leq 1\}} + |\langle \xi, u\rangle| \mathbf{1}_{\{\|\xi\|>1\}} \\ &\leq \frac{1}{2}\|\xi\|^2\|u\|^2 \mathbf{1}_{\{\|\xi\|\leq 1\}} + \|\xi\|\|u\| \mathbf{1}_{\{\|\xi\|>1\}},\end{aligned} \tag{5.31}$$

from which (5.29), (5.30), and the continuity of F, R readily follows (by dominated convergence). □

Assumption $\mathcal{D}$ implies that the semigroup $(\mathrm{e}^{t\hat{B}})_{t\geq 0}$ satisfies (5.6), that is $(\mathrm{e}^{t\hat{B}})_{t\geq 0}$ is uniformly exponential stable. This has the following consequence on the solution $\psi(\cdot, u)$ of the generalized Riccati equation (2.8b):

Lemma 5.14. *Let (b, B, m, μ) be an admissible parameter set according to Definition 2.3 and for $u \in \mathcal{H}^+$ let $\psi(\cdot, u)$ be the unique solution to* (2.8b)*. Then*

$$\|\psi(t, u)\| \leq \|\mathrm{e}^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})}\|u\|, \quad \forall t \geq 0. \tag{5.32}$$

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

Proof. First note that whenever (b, B, m, μ) is an admissible parameter set according to Definition 2.3, then so is $(0, B, 0, \mu)$. Therefore, the existence of an affine Markov process $(Y_t)_{t\geq 0}$ associated to $(0, B, 0, \mu)$ and initial value $Y_0 = x$ is guaranteed by Theorem 2.8. Note that the unique solution $\psi(\cdot, u)$ to (2.8b) is $\mathcal{H}^+$-valued. Let $u \in \mathcal{H}^+$ and note that due to the convexity of the exponential function and Jensen's inequality we have

$$\mathrm{e}^{-\mathbb{E}_x[\langle u, Y_t\rangle]} \leq \mathbb{E}_x\left[\mathrm{e}^{-\langle u, Y_t\rangle}\right] = \exp(-\langle \psi(t, u), x\rangle).$$

Hence, we find that for all $u, x \in \mathcal{H}^+$:

$$\langle \psi(t, u), x\rangle \leq \mathbb{E}_x\left[\langle u, Y_t\rangle\right] = \langle x, \mathrm{e}^{t\hat{B}}u\rangle \leq \|x\|\|u\|\|\mathrm{e}^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})}.$$

For fixed $t \geq 0$ and $u \in \mathcal{H}^+$ we choose $x = \psi(t, u) \in \mathcal{H}^+$ and obtain

$$\|\psi(t, u)\|^2 \leq \|\psi(t, u)\|\|u\|\|\mathrm{e}^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})},$$

which proves the first statement. If Assumption $\mathcal{D}$ is satisfied, then from (5.32) and (5.6) it follows that $\|\psi(t, u)\| \leq M\,\mathrm{e}^{-\delta t}\|u\|$ and hence $\lim_{t\to\infty} \psi(t, u) = 0$. □

5.5.2 Invariant measure 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 lemma we give an important convolution property of the transition kernels $p_t(x, \cdot)$.

Lemma 5.15. *Let $(Y_t)_{t\geq 0}$ be the unique affine process associated with the admissible parameter set $(0, B, 0, \mu)$ and denote its transition kernels by $(q_t(x, \cdot))_{t\geq 0}$. Then for every $t \geq 0$ and $x \in \mathcal{H}^+$ we have*

$$p_t(x, \cdot) = p_t(0, \cdot) * q_t(x, \cdot). \tag{5.33}$$

Proof. Since $b = 0$ and $m = 0$ the function F in (2.8a) vanishes, see also (2.6a), and thus $\phi(t, u) = 0$ for all $t \geq 0$. Hence, for every $t \geq 0$ the affine-transform formula (5.2) for Y_t takes the form

$$\int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,\xi\rangle}\, q_t(x, \mathrm{d}\xi) = \exp\big(- \langle \psi(t, u), x\rangle\big), \quad \text{for } u \in \mathcal{H}^+. \tag{5.34}$$

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 $t \geq 0$ arbitrary and $u \in \mathcal{H}^+$, then

$$\begin{aligned}
\int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,\xi\rangle}\, p_t(0, \cdot) * q_t(x, \cdot)(\mathrm{d}\xi) &= \int_{\mathcal{H}^+} \Big(\int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,\xi_1+\xi_2\rangle}\, p_t(0, \mathrm{d}\xi_1)\Big)\, q_t(x, \mathrm{d}\xi_2) \\
&= \mathrm{e}^{-\phi(t,u)} \int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,\xi_2\rangle}\, q_t(x, \mathrm{d}\xi_2) \\
&= \mathrm{e}^{-\phi(t,u)}\, \mathrm{e}^{-\langle \psi(t,u),x\rangle},
\end{aligned}$$

which completes the proof thanks to (5.2) and the fact that the functions $x \mapsto \mathrm{e}^{-\langle u,x\rangle}$ characterize measures, see Lemma 3.19 in Chapter 3. □

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

Proposition 5.16. *Let $(X_t)_{t\geq}$ be an affine process associated with the admissible parameter set (b, B, m, μ) satisfying Assumption $\mathcal{D}$. Then for all $u \in \mathcal{H}^+$ and for all $x \in \mathcal{H}^+$ we have*

$$\lim_{t\to\infty} \mathbb{E}_x \left[\mathrm{e}^{-\langle u,X_t\rangle}\right] = \exp\left(- \int_0^\infty F(\psi(s, u))\, \mathrm{d}s\right) \in [0, \infty). \tag{5.35}$$

Proof. Let $u \in \mathcal{H}^+$ and $x \in \mathcal{H}^+$, then by Lemma 5.14 and (5.6) we have

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

Lemma 5.13 gives

$$|F(\psi(t,u))| \leq C\left(\|\psi(t,u)\| + \|\psi(t,u)\|^2\right) \leq CM^2\,\mathrm{e}^{-\delta s}(\|u\| + \|u\|^2), \tag{5.36}$$

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

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

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, the continuity of the exponential function and the fact that by Lemma 5.14 $\langle \psi(t,u), x\rangle \to 0$ for all $x, u \in \mathcal{H}^+$ as $t \to \infty$ imply (5.35). □

The next lemma 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 5.17. *Let $(X_t)_{t\geq}$ be an affine process associated with the admissible parameter set (b, B, m, μ) satisfying Assumption $\mathcal{D}$ and denote its transition semigroup by $(P_t)_{t\geq 0}$. Then we have*

$$\sup_{t\geq 0} \|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}}))} < \infty. \tag{5.37}$$

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 $\sup_{t\geq 0}\|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}}))} \leq \sup_{t\geq 0}\|P_t\rho\|_{\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})}$. Let $(e_i)_{i\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$ and recall that by Remark 2.27 we have $P_t\rho(x) = \mathbb{E}_x\left[\rho(X_t)\right]$ for all $t \geq 0$. Hence, by Parseval's identity we conclude

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

Using (5.5) with $v = w = e_i$ for $i \in \mathbb{N}$ we find

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

We thus see that

$$\sum_{i=1}^{\infty} \mathbb{E}_x \left[\langle X_t, e_i\rangle^2\right] \leq 2\|\int_0^t \mathrm{e}^{s\hat{B}^*}\hat{b}\ \mathrm{d}s\|^2 + 2\|\mathrm{e}^{t\hat{B}^*}x\|^2 \tag{5.38a}$$

$$+ \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} \|\mathrm{e}^{s\hat{B}^*}\xi\|^2\, m(\mathrm{d}\xi)\ \mathrm{d}s \tag{5.38b}$$

$$+ \int_0^t \int_0^s \int_{\mathcal{H}^+\setminus\{0\}} \|\mathrm{e}^{u\hat{B}^*}\xi\|^2 \big\langle \hat{b}, \mathrm{e}^{(s-u)\hat{B}} \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\big\rangle\, \mathrm{d}u\, \mathrm{d}s \tag{5.38c}$$

$$+ \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} \|\mathrm{e}^{s\hat{B}^*}\xi\|^2 \big\langle x, \mathrm{e}^{(t-s)\hat{B}}\ \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\big\rangle\ \mathrm{d}s. \tag{5.38d}$$

In the following we show that every term on the right-hand side of (5.38a)-(5.38d) converges as $t \to \infty$ uniformly in x, which then yields (5.37).
Note first that the adjoint semigroup $(\mathrm{e}^{t\hat{B}^*})_{t\geq 0}$ generated by $\hat{B}^*$, the adjoint of $\hat{B}$, is also uniformly stable as $\|\mathrm{e}^{t\hat{B}}\|_{\mathcal{L}(\mathcal{H})} = \|\mathrm{e}^{t\hat{B}*}\|_{\mathcal{L}(\mathcal{H})}$ for all $t \geq 0$. For the first term on the right-hand side of (5.38a) we have $\int_0^t \left\|\mathrm{e}^{s\hat{B}^*}\hat{b}\right\|\ \mathrm{d}s \leq \frac{M}{\delta}\|\hat{b}\|$. The second term in (5.38a) vanishes as $t \to \infty$, since $(\mathrm{e}^{t\hat{B}^*})_{t\geq 0}$ is uniformly stable. Note that $s \mapsto M^2\,\mathrm{e}^{-2\delta s}\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\, m(\mathrm{d}\xi)$ is an integrable majorant for the term in (5.38b) and thus the integral converges for $t \to \infty$. For (5.38c) note that $\langle \hat{b}, \mathrm{e}^{(s-u)\hat{B}}\ \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\rangle \geq 0$ for every $s, u \in \mathbb{R}^+$, which follows from the admissible parameter conditions, which imply that $\hat{b} \in \mathcal{H}^+$ and $\mathrm{e}^{(s-u)\hat{B}}(\mathcal{H}^+) \subseteq \mathcal{H}^+$, whenever $s \geq u$. Hence, we have

$$\int_0^\infty \int_0^s \int_{\mathcal{H}^+\setminus\{0\}} \|\mathrm{e}^{u\hat{B}^*}\xi\|^2 \big\langle \hat{b}, \mathrm{e}^{(s-u)\hat{B}}\ \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\big\rangle\ \mathrm{d}u\ \mathrm{d}s \leq \frac{3}{2}\frac{M^3}{\delta^2}\|\hat{b}\|\|\mu(\mathcal{H}^+\setminus\{0\})\|.$$

Finally, note that $\int_0^t \mathrm{e}^{-2\delta s}\,\mathrm{e}^{-\delta(t-s)}\ \mathrm{d}s = \frac{1}{\delta}(\mathrm{e}^{-\delta t} - \mathrm{e}^{-2\delta t})$ and hence the last term in (5.38d) vanishes as $t \to \infty$, which can be seen from

$$\begin{aligned}
&\int_0^t \int_{\mathcal{H}^+\setminus\{0\}} \|\mathrm{e}^{s\hat{B}^*}\xi\|^2 \big\langle x, \mathrm{e}^{(t-s)\hat{B}}\ \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\big\rangle\ \mathrm{d}s \\
&\quad\leq M^3\|\mu(\mathcal{H}^+\setminus\{0\})\|\|x\|\int_0^t \mathrm{e}^{-2\delta s}\,\mathrm{e}^{-\delta(t-s)}\ \mathrm{d}s \\
&\quad\leq \frac{M^3}{\delta}\|\mu(\mathcal{H}^+\setminus\{0\})\|\|x\|(\mathrm{e}^{-\delta t} - \mathrm{e}^{-2\delta t}).
\end{aligned} \tag{5.39}$$

Thus, we proved that $\sup_{t\geq 0}\sup_{x\in\mathcal{H}^+}\mathbb{E}_x\left[\rho(X_t)\right] < \infty$, which proves the statement. □

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

Proposition 5.18. *For all $f \in \mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+)$ the limit $\lim_{t\to\infty} P_t f$ exists in $\mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+)$ and $\pi(f) \coloneqq \lim_{t\to\infty} P_t f(x)$ defines a continuous linear functional on $\mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+)$.*

Proof. By Proposition 5.16 we know that for every $u \in \mathcal{H}^+$

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

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

$$\begin{aligned}
\|P_t\, \mathrm{e}^{-\langle u,\cdot\rangle} - \pi_u\|_\rho &= \sup_{x\in\mathcal{H}^+} \frac{\left|\left(\mathrm{e}^{-\int_0^t F(\psi(s,u))\,\mathrm{d}s - \langle \psi(t,u),x\rangle} - \mathrm{e}^{-\int_0^\infty F(\psi(s,u))\,\mathrm{d}s}\right)\right|}{\rho(x)} \\
&\leq \sup_{x\in\mathcal{H}^+} \frac{\left|\int_t^\infty F(\psi(s,u))\,\mathrm{d}s - \langle \psi(t,u),x\rangle\right|}{\rho(x)} \\
&\leq \int_t^\infty |F(\psi(s,u))|\,\mathrm{d}s + \|\psi(t,u)\| \sup_{x\in\mathcal{H}^+} \frac{\|x\|}{\rho(x)}.
\end{aligned}$$

where we have used $\rho(x) = 1 + \|x\|^2$. The first term converges to zero due to (5.36), while the second term tends to zero by Lemma 5.14.
Let $\mathcal{D} \coloneqq \operatorname{lin}\{\mathrm{e}^{-\langle u,\cdot\rangle} \colon u \in \mathcal{H}^+\}$ 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 5.17 we have $\sup_{t\geq 0}\|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+))} < \infty$ and hence

$$|\pi(f)| \leq \sup_{t\geq 0}\|P_t\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+))}\|f\|_{\mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+)},$$

i.e. π is bounded on $\mathcal{D}$. Since $\mathcal{D}$ is dense in $\mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+)$, see Lemma 2.28, 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 thus proved the existence of $\pi \in \mathcal{L}(\mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+), \mathbb{R})$ and it is only left to show that $P_t f \to \pi(f)$ as $t \to \infty$ for all $f \in \mathcal{B}_\rho(\mathcal{H}_\mathrm{w}^+)$. The latter one 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 conclude the assertion. □

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

Lemma 5.19. *Let π denote the continuous linear functional in Proposition 5.18. Then there exists a unique probability measure ν on $\mathcal{H}^+$ such that*

$$\pi(f) = \int_{\mathcal{H}^+} f(\xi)\,\nu(\mathrm{d}\xi) \quad \text{for all } f \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}}), \tag{5.40}$$

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 [45, Theorem 2.4] there exists a unique finite signed Radon measure ν on $\mathcal{B}(\mathcal{H}^+)$ such that (5.40) and

$$\int_{\mathcal{H}^+} \big(1 + \|x\|^2\big)\,|\nu|(\mathrm{d}\xi) = \|\pi\|_{\mathcal{L}(\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}}),\mathbb{R})} \tag{5.41}$$

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 is left 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 non-negative $f \in C_b(\mathcal{H}^+_{\mathrm{w}})$ 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 non-negative $f \in C_b(\mathcal{H}^+_{\mathrm{w}})$, which implies that the measure ν is also non-negative and hence it is a probability measure on $\mathcal{B}(\mathcal{H}^+)$. □

In the following we identify the linear functional π with the measure ν given by Lemma 5.19 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 5.20. *Let (b, B, m, μ) be an admissible parameter set such that Assumption $\mathcal{D}$ is satisfied and denote the associated subcritical affine Markov process on $\mathcal{H}^+$ by $(X_t)_{t\geq 0}$ and its transition kernels by $(p_t(x,\mathrm{d}\xi))_{t\geq 0}$. 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*

$$\lim_{t\to\infty} \int_{\mathcal{H}^+} f(\xi)\,p_t(x,\mathrm{d}\xi) \to \int_{\mathcal{H}^+} f(\xi)\,\pi(\mathrm{d}\xi), \quad \forall f \in C_b(\mathcal{H}^+_{\mathrm{w}}), \tag{5.42}$$

and the Laplace transform of π is given by (5.7)*.*

Proof. In Proposition 5.18 and the subsequent arguments, we have already shown the existence of the Borel measure π such that (5.42) holds. It is left to show that π is the unique invariant measure. We have

$$\begin{aligned}\int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,\xi\rangle}\Big(\int_{\mathcal{H}^+} p_t(x,\mathrm{d}\xi)\,\pi(\mathrm{d}x)\Big) &= \int_{\mathcal{H}^+}\Big(\int_{\mathcal{H}^+} \mathrm{e}^{-\langle u,\xi\rangle}\,p_t(x,\mathrm{d}\xi)\Big)\,\pi(\mathrm{d}x)\\ &= \mathrm{e}^{-\phi(t,u)}\int_{\mathcal{H}^+} \mathrm{e}^{-\langle x,\psi(t,u)\rangle}\,\pi(\mathrm{d}x).\end{aligned}$$

Note that by (5.28) we have $\psi(t{+}s,u) = \psi(t,\psi(s,u))$ and hence for every $u \in \mathcal{H}^+$ we have

$$\begin{aligned}
\mathrm{e}^{-\phi(t,u)} \int_{\mathcal{H}^+} \mathrm{e}^{-\langle x,\psi(t,u)\rangle}\, \pi(\mathrm{d}x) &= \mathrm{e}^{-\phi(t,u)}\, \mathrm{e}^{-\int_0^\infty F(\psi(s,\psi(t,u)))\,\mathrm{d}s} \\
&= \mathrm{e}^{-\phi(t,u)}\, \mathrm{e}^{-\int_0^\infty F(\psi(t+s,u))\,\mathrm{d}s} \\
&= \mathrm{e}^{-\phi(t,u)}\, \mathrm{e}^{-\int_t^\infty F(\psi(s,u))\,\mathrm{d}s} \\
&= \mathrm{e}^{-\int_0^\infty F(\psi(s,u))\,\mathrm{d}s} \\
&= \int_{\mathcal{H}^+} \mathrm{e}^{-\langle x,u\rangle}\, \pi(\mathrm{d}x).
\end{aligned}$$

This proves the invariance of π. Next, we prove that π is the unique invariant measure. Suppose there exists a $\pi' \in \mathcal{M}(\mathcal{H}^+)$ which is invariant for $p_t(x,\mathrm{d}\xi)$, then for every $u \in \mathcal{H}^+$ and $t \geq 0$ we have:

$$\begin{aligned}
\int_{\mathcal{H}^+} \mathrm{e}^{-\langle x,u\rangle}\, \pi'(\mathrm{d}x) &= \int_{\mathcal{H}^+\setminus\{0\}} \mathrm{e}^{-\langle u,\xi\rangle} \Big(\int_{\mathcal{H}^+\setminus\{0\}} p_t(x,\mathrm{d}\xi)\, \pi'(\mathrm{d}x) \Big) \\
&= \int_{\mathcal{H}^+\setminus\{0\}} \mathrm{e}^{-\phi(t,u)-\langle x,\psi(t,u)\rangle}\, \pi'(\mathrm{d}x),
\end{aligned}$$

now by letting $t \to \infty$ we find that

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

The Laplace transform is measure determining for measures on $\mathcal{B}(\mathcal{H}^+)$ and hence $\pi = \pi'$. □

Remark 5.21. The convergence in (5.42) is weak convergence of $p_t(x,\cdot)$ to π as $t \to \infty$ in the weak topology on $\mathcal{H}$. Even though the Borel 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 $P_t f(x) \to \int_{\mathcal{H}^+} f(\xi)\,\pi(\mathrm{d}\xi)$ for all $f \in C_b(\mathcal{H}^+_{\mathrm{w}})$.
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 [116, Theorem 1 and 2] we know that weak convergence in the weak topology together with

$$\lim_{N\to\infty} \sup_{n\in\mathbb{N}} p_t(x, A_{N,n}) = 0, \quad \text{for all } \epsilon > 0,$$

where $A_{N,n} \coloneqq \left\{ \sum_{i=N}^\infty \langle x,e_i\rangle^2 \geq \epsilon \right\}$ for $N,n \in \mathbb{N}$, implies $p_t(x,\cdot) \Rightarrow \pi$ as $t \to \infty$. Note that in our main Theorem 5.3 we assert weak convergence in the *strong topology*, which will be shown below.

5.5.3 Proof of Theorem 5.3

Proposition 5.20 ensures the existence of a unique invariant measure π of the family $(p_t(x,\cdot))_{t\geq 0}$ with Laplace transform (5.7). We also proved weak convergence of $(p_t(x,\cdot))_{t\geq 0}$ to π as $t\to\infty$ in the weak topology. What is left to show is the convergence rates in Wasserstein distance of order p for $p\in[1,2]$ as in (5.8). Then convergence in Wasserstein distance of some order $p\in[1,\infty)$ implies weak convergence (in the strong topology) and convergence of the p-th absolute moment, see [143, Theorem 6.9]. This implies the last assertion of Theorem 5.3. In the remainder we prove the convergence rates (5.8). We need the following lemma on the convolution property of the Wasserstein distance on $\mathcal{H}^+$:

Lemma 5.22. *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 let G' be any coupling of (ρ,ρ). For each $f,g:\mathcal{H}\longrightarrow\mathbb{R}_+$ we find that

$$\begin{aligned}&\int_{\mathcal{H}^+\times\mathcal{H}^+}(f(x)+g(y))\,(G'*G)(\mathrm{d}x,\mathrm{d}y)\\&=\int_{\mathcal{H}^+\times\mathcal{H}^+}\int_{\mathcal{H}^+\times\mathcal{H}^+}(f(x+z)+g(y+z'))\,G'(\mathrm{d}z,\mathrm{d}z')G(\mathrm{d}x,\mathrm{d}y)\\&=\int_{\mathcal{H}^+\times\mathcal{H}^+}f(x+z)\rho(\mathrm{d}z)\mu(\mathrm{d}x)+\int_{\mathcal{H}^+\times\mathcal{H}^+}g(y+z')\rho(\mathrm{d}z')\nu(\mathrm{d}y),\end{aligned}$$

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

$$\begin{aligned}&W_2(\rho*\mu,\rho*\nu)^2\\&\leq\int_{\mathcal{H}^+\times\mathcal{H}^+}\|x-y\|^2(G'*G)(\mathrm{d}x,\mathrm{d}y)\\&=\int_{\mathcal{H}^+\times\mathcal{H}^+}\int_{\mathcal{H}^+\times\mathcal{H}^+}\|(x+z)-(y+z')\|^2G'(\mathrm{d}z,\mathrm{d}z')G(\mathrm{d}x,\mathrm{d}y)\\&=\iint_{(\mathcal{H}^+\times\mathcal{H}^+)^2}\left(\|x-y\|^2+2\langle x-y,z-z'\rangle+\|z-z'\|^2\right)G'(\mathrm{d}z,\mathrm{d}z')G(\mathrm{d}x,\mathrm{d}y)\\&=\int_{\mathcal{H}^+\times\mathcal{H}^+}\|x-y\|^2G(\mathrm{d},\mathrm{d}y)+\int_{\mathcal{H}^+\times\mathcal{H}^+}\|z-z'\|^2G'(\mathrm{d}z,\mathrm{d}z'),\end{aligned}$$

where the last inequality 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 $\int_{\mathcal{H}^+\times\mathcal{H}^+}\|z-z'\|^2G'(\mathrm{d}z,\mathrm{d}z')=0$. Since G was arbitrary, the assertion is proved. □

Let $p \in [1,2]$ and as before we denote by $q_t(x, \mathrm{d}\xi)$ the transition kernel 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 [143, Theorem 4.8] and since $W_p \leq W_2$ for $p \in [1,2]$ we have

$$\begin{aligned} W_p\big(p_t(x, \cdot), \pi\big) &= W_p\left(\int_{\mathcal{H}^+} p_t(z, \cdot)\, \delta_x(\mathrm{d}z), \int_{\mathcal{H}^+} p_t(y, \cdot)\, \pi(\mathrm{d}y)\right) \\ &\leq \left(\int_{\mathcal{H}^+ \times \mathcal{H}^+} W_2\big(p_t(z, \cdot), p_t(y, \cdot)\big)^p\, G(\mathrm{d}z, \mathrm{d}y)\right)^{1/p}. \end{aligned} \tag{5.43}$$

By Lemma 5.15 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 5.22 that

$$\begin{aligned} W_2\big(p_t(z, \cdot), p_t(y, \cdot)\big)^p &= W_2\big(q_t(z, \cdot) * p_t(0, \cdot), q_t(y, \cdot) * p_t(0, \cdot)\big)^p \\ &\leq W_2\big(q_t(z, \cdot), q_t(y, \cdot)\big)^p \\ &\leq \left(\int_{\mathcal{H}^+ \times \mathcal{H}^+} \|\tilde{x} - \tilde{y}\|^2\, H(\mathrm{d}\tilde{x}, \mathrm{d}\tilde{y})\right)^{p/2} \\ &\leq \left(2 \int_{\mathcal{H}^+ \times \mathcal{H}^+} \big(\|\tilde{x}\|^2 + \|\tilde{y}\|^2\big)\, H(\mathrm{d}\tilde{x}, \mathrm{d}\tilde{y})\right)^{p/2} \\ &= \left(2 \int_{\mathcal{H}^+ \times \mathcal{H}^+} \|\tilde{x}\|^2\, q_t(z, \mathrm{d}\tilde{x}) + 2 \int_{\mathcal{H}^+ \times \mathcal{H}^+} \|\tilde{y}\|^2\, q_t(y, \mathrm{d}\tilde{y})\right)^{p/2}. \end{aligned} \tag{5.44}$$

Now, recall from (5.38a) that

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

while all the other terms vanish as $\hat{b} = 0$ and $m = 0$.

By the same estimations as in (5.39) we conclude that

$$\int_{\mathcal{H}^+\times\mathcal{H}^+} \|\tilde{x}\|^2 \, q_t(z,\mathrm{d}\tilde{x}) \leq 2\|\mathrm{e}^{t\hat{B}^*}z\|^2 + \int_0^t \int_{\mathcal{H}^+\setminus\{0\}} \|\mathrm{e}^{s\hat{B}^*}\xi\|^2 \big\langle z, \mathrm{e}^{(t-s)\hat{B}}\, \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\big\rangle \, \mathrm{d}s$$
$$\leq 2M^2\,\mathrm{e}^{-2\delta t}\|z\|^2 + \frac{M^3}{\delta}\|\mu(\mathcal{H}^+\setminus\{0\})\|\,\mathrm{e}^{-\delta t}\|z\|.$$

Inserting this back into (5.44) and using the sub-additivity of $x \mapsto x^{p/2}$ for $p \in [1,2]$, we obtain

$$W_2\big(p_t(z,\cdot),p_t(y,\cdot)\big)^p \leq \big(C_1\,\mathrm{e}^{-\delta t}\|z\|\big)^p + \big(C_2\,\mathrm{e}^{-\delta/2}\|z\|^{1/2}\big)^p$$
$$+ \big(C_1\,\mathrm{e}^{-\delta t}\|y\|\big)^p + \big(C_2\,\mathrm{e}^{-\delta/2}\|y\|^{1/2}\big)^p, \tag{5.45}$$

for $C_1 = 2M$ and $C_2 = 2^{1/2}M^{3/2}\delta^{-1/2}\|\mu(\mathcal{H}^+\setminus\{0\})\|^{1/2}$. Now, plugging (5.45) back into (5.43) and again by the sub-additivity of $x \mapsto x^{1/p}$, we obtain the desired (5.8).

5.5.4 Proof of Corollary 5.5

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 2.8. Moreover, let π be the unique invariant distribution of $(p_t(x,\cdot))_{t\geq 0}$ (which is independent of $x \in \mathcal{H}^+$). From Theorem 5.3 i) we know that π is inner-regular. Thus, from Proposition 5.23 below, we conclude the existence of a unique Markov process $(X^\pi)_{t\geq 0}$ such that for all $f \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ we have $\mathbb{E}_\pi\left[f(X_t)\right] = \int_{\mathcal{H}^+} P_t f(x)\,\pi(\mathrm{d}x)$. Moreover, since π is the invariant measure we have for each $t \geq 0$

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

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

5.5.5 Proof of Proposition 5.6

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 (5.41) 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, which implies

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

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 distribution. 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

$$K_{\mathrm{w}}^R \coloneqq \left\{x \in \mathcal{H}^+ \colon \; \|x\|^2 + 1 \leq R\right\},$$

is compact in $\mathcal{H}^+$ equipped with the weak topology and

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

which by [51, Theorem 2.7] implies $\langle u, \cdot\rangle \in \mathcal{B}_\rho(\mathcal{H}_{\mathrm{w}}^+)$ for all $u \in \mathcal{H}$. By Proposition 5.18 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 (5.4) for $u = e_i$ for $i \in \mathbb{N}$ we have

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

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

$$\begin{aligned}\lim_{t\to\infty} \int_{\mathcal{H}^+} \xi\, p_t(x, \mathrm{d}\xi) &= \lim_{t\to\infty} \sum_{i=1}^\infty \int_{\mathcal{H}^+} \langle \xi, e_i\rangle e_i\, p_t(x, \mathrm{d}\xi) \\ &= \sum_{i=1}^\infty \Big(\int_0^\infty \langle \hat{b}, \mathrm{e}^{s\hat{B}} e_i\rangle \; \mathrm{d}s\Big) e_i \\ &= \int_0^\infty \mathrm{e}^{s\hat{B}^*} \hat{b} \; \mathrm{d}s,\end{aligned}$$

where we have used that

$$\lim_{t\to\infty} \sum_{i=1}^\infty \Big(\int_0^t \langle \hat{b}, \mathrm{e}^{s\hat{B}} e_i\rangle \; \mathrm{d}s\Big) e_i = \sum_{i=1}^\infty \Big(\int_0^\infty \langle \hat{b}, \mathrm{e}^{s\hat{B}} e_i\rangle \; \mathrm{d}s\Big) e_i,$$

which is justified if $\lim_{N\to\infty} \sup_{t\geq 0} \sum_{i=N}^\infty \Big\| \int_0^t \langle e^{s\widehat{B}^*} \widehat{b}, e_i\rangle e_i \,\mathrm{d}s\Big\| = 0$. The latter one follows from

$$\sup_{t\geq 0} \sum_{i=N}^\infty \Big\| \int_0^t \langle e^{s\widehat{B}^*} \widehat{b}, e_i\rangle e_i \; \mathrm{d}s\Big\| \leq \int_0^\infty \sum_{i=1}^N |\langle e^{s\widehat{B}^*} \widehat{b}, e_i\rangle| \; \mathrm{d}s$$

and

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

Thus, recalling that $\hat{b} = b + \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \xi\, m(\mathrm{d}\xi)$ yields (5.10).
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 (5.5) and analogous arguments as we used in Lemma 5.17 (to show that the integrals on the right-hand side of (5.46) below exists and are finite), we find that

$$\begin{aligned}\lim_{t\to\infty} P_t g^{i,j}(x) &= \Big(\int_0^\infty \langle \hat{b}, \mathrm{e}^{s\hat{B}} e_i\rangle \,\mathrm{d}s\Big)\Big(\int_0^\infty \langle \hat{b}, \mathrm{e}^{s\hat{B}} e_j\rangle \,\mathrm{d}s\Big) \\ &\quad + \int_0^\infty \int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, \mathrm{e}^{s\hat{B}} e_i\rangle\langle \xi, \mathrm{e}^{s\hat{B}} e_j\rangle \, m(\mathrm{d}\xi)\,\mathrm{d}s \\ &\quad + \int_0^\infty \int_0^s \Big\langle \hat{b}, \mathrm{e}^{(s-u)\hat{B}} \int_{\mathcal{H}^+\setminus\{0\}} \langle \xi, \mathrm{e}^{u\hat{B}} e_i\rangle\langle \xi, \mathrm{e}^{u\hat{B}} e_j\rangle \,\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big\rangle \,\mathrm{d}u\,\mathrm{d}s.\end{aligned} \tag{5.46}$$

holds for all $i, j \in \mathbb{N}$. The second moment formula (5.11) 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

$$\lim_{t\to\infty} P_t g^{i,j}(x) = \int_{\mathcal{H}^+} g^{i,j}(y)\,\pi(\mathrm{d}y), \qquad i, j \in \mathbb{N}. \tag{5.47}$$

Since the function $g^{i,j}$ does not belong to $\mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$, we cannot obtain (5.47) directly from Proposition 5.18. However, since $P_t\langle \cdot, e_i\rangle\langle \cdot, e_j\rangle(x) \leq P_t\rho(x) < \infty$ 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 $g^{i,j}_n := g^{i,j} \wedge n$ for $n \in \mathbb{N}$. Then $g^{i,j}_n \in \mathcal{B}_\rho(\mathcal{H}^+_{\mathrm{w}})$ and we find that

$$\begin{aligned}\Big| P_t g^{i,j}(x) - \int_{\mathcal{H}^+} g^{i,j}(x)\,\pi(\mathrm{d}x)\Big| &\leq |P_t g^{i,j}(x) - P_t g^{i,j}_n(x)| \\ &\quad + \Big| P_t g^{i,j}_n(x) - \int_{\mathcal{H}^+} g^{i,j}_n(x)\,\pi(\mathrm{d}x)\Big| \\ &\quad + \Big| \int_{\mathcal{H}^+} g^{i,j}_n(x)\,\pi(\mathrm{d}x) - \int_{\mathcal{H}^+} g^{i,j}(x)\,\pi(\mathrm{d}x)\Big|.\end{aligned}$$

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

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

where the last identity follows from the characterization of convergence in the Wasserstein distance (see [143, Section 6]). Hence we find n large enough such that $|P_t g^{i,j}(x) - P_t g^{i,j}_n(x)| < \varepsilon$ holds uniformly in $t \geq 0$.

Finally, for this fixed choice of n, we may choose in view of Proposition 5.18 t large enough so that $\left|P_t g^{i,j}(x) - \int_{\mathcal{H}^+} g^{i,j}(x)\,\pi(\mathrm{d}x)\right| < \varepsilon$. Combining all these estimates proves (5.47), which completes the proof of Proposition 5.6.

5.6 A version of Kolmogorov's extension theorem for inner-regular distributions

In the proof of Corollary 5.5 we made use of the following adapted version of Kolmogorov's extensions theorem (Theorem 2.26, which is taken from [45, Theorem 2.11]).

Proposition 5.23. *Let (Y, ρ) be a weighted space and let $(P_t)_{t\geq 0}$ be a generalized 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 exists a probability space $(\Omega, \mathcal{F}, \mathbb{P}_\nu)$, filtered by a right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$, and a Markov process $(X_t)_{t\geq 0}$ with values in Y such that $\mathbb{P}_\nu(X_0 \in A) = \nu(A)$ for every $\mathcal{B}(Y)$ and*

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

Proof. In [45, Theorem 2.11] this version of the Kolmogorov extension theorem is proven for $\nu = \delta_x$ and $x \in Y$. In the proof of [45, 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

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

form a Kolmogorov consistent family according to [2, Section 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

$$\nu(A) = \sup\left\{\mu(K)\colon K \subseteq A\colon\ A \in \mathcal{K} \text{ and } A \in \mathcal{B}(Y) \cap \mathcal{K}\right\}, \tag{5.48}$$

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 (5.48) and hence the assertion of the proposition follows from this, Kolmogorov extension theorem [2, Theorem 15.23] and analogous arguments as in [45, Theorem 2.11]. □

5.7 Concluding remarks

In this chapter we studied the long-time behavior 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 provide explicit formulas for the first and second moment of the invariant distribution. Moreover, we proved ergodicity of the affine processes and established explicit and dimension-free convergence rates for the convergence of the transition kernels to the invariant measure in Wasserstein distance of order $p \in [1, 2]$. From a theoretical viewpoint, this chapter provides the first systematic study of the long-time behavior of affine processes in a Hilbert space setting, in particular for affine processes admitting state-dependent jumps. We suppose that our techniques, e.g. the use of the generalized Feller framework, can be used effectively to study the long-time behavior of affine processes in general Hilbert spaces.

From an application point of view, we use 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 *stationary covariance regime* by using the stationary affine process to model the instantaneous covariance. In this context we considered the geometric affine stochastic covariance model for forward curve dynamics in commodity markets, introduced in Chapter 4, and studied the long-time behavior of the implied forward volatility smile for large forward-start dates which can be combined with the option-pricing formulas for plain-vanilla options from Proposition 4.7.

CHAPTER 6

Finite-rank Approximation of Affine Stochastic Covariance Models

Abstract of the chapter In this chapter we study finite-dimensional approximations of instantaneous covariance processes in infinite-dimensional affine stochastic covariance models. More specifically, we construct sequences of finite-rank operator-valued affine processes converging weakly to the class of affine pure-jump processes on positive Hilbert-Schmidt operators introduced in Chapter 2. In addition to that, we present explicit convergence rates for the Laplace transforms in terms of a Galerkin approximation of the associated generalized Riccati equations. The relevance of this chapter is at least twofold: First, it provides an alternative proof for the existence of affine pure-jump processes on positive Hilbert-Schmidt operators, this time by exploiting the connection to their finite-dimensional versions. Second, it gives a numerically feasible approximation scheme for this class of infinite-dimensional affine processes and quantifies the approximation error by establishing explicit error bounds for finite-rank approximations. As a byproduct of this approach, we prove the existence of càdlàg versions of affine pure-jump processes. Moreover, we apply our findings to stochastic covariance modeling and study weakly convergent finite-rank approximation of affine stochastic covariance models.

This chapter is based on the working paper:
Karbach, S.
Finite-rank approximation of affine processes on Hilbert-Schmidt operators.

6.1 Introduction

The leitmotif of the affine class is its good tractability entailed by the *affine transform formula* (1.3), i.e. that computations of the Fourier-Laplace transform reduce to mere evaluations of ϕ and ψ, provided that both functions are (approximately) known. However, a priori it is not clear that this assumption is satisfied, especially since in our case the associated generalized Riccati equations are, in general, non-linear and infinite-dimensional.
This issue motivates us to analyze feasible numerical approximations of affine processes on positive Hilbert-Schmidt operators and their associated generalized Riccati equations. More specifically, we present a Galerkin type approximation of the generalized Riccati equations (2.8) and construct an associated sequence of finite-rank operator-valued affine processes that converge weakly to the infinite-dimensional affine process associated with the generalized Riccati equations. In addition, we provide explicit convergence rates of the corresponding Laplace transforms and quantify the approximation error. As a convenient byproduct of our method, we prove that the considered class of affine processes on $\mathcal{H}^+$ always admits versions with càdlàg paths and in this respect qualify as the instantaneous covariance process in stochastic covariance models of the form discussed in Chapter 3 (see also Section 6.5 below). Moreover, we want to emphasize that our approximation method is also constructive and gives an alternative proof for the existence of affine pure-jump processes on $\mathcal{H}^+$.
As noted before, we view affine processes on the cone $\mathcal{H}^+$ as the natural infinite-dimensional analog of affine processes on $\mathbb{S}_d^+$. In fact, our findings in this chapter prove that affine processes on $\mathcal{H}^+$ are actually the limits of properly chosen affine $\mathbb{S}_d^+$-valued affine processes studied in [42]. In the following paragraphs we describe and structure the main contributions of this chapter:

Galerkin approximation of generalized Riccati equations In Proposition 6.1, we construct Galerkin type approximations of the functions ϕ and ψ, the unique solutions of the generalized Riccati equations (2.8) modulated by an admissible parameter set (b, B, m, μ). The Galerkin approximations are defined on positive finite-rank operators and we prove their convergence to ϕ and ψ uniformly on compact time intervals. Moreover, denoting the sequence of Galerkin approximations by $(\phi_d)_{d\in\mathbb{N}}$ and $(\psi_d)_{d\in\mathbb{N}}$, we present explicit bounds, in terms of (b, B, m, μ) and the initial value $u \in \mathcal{H}^+$ of ψ, for the approximation error

$$\sup_{t\in[0,T]} \big(|\phi_d(t, \mathbf{P}_d(u)) - \phi(t, u)| + \|\psi_d(t, \mathbf{P}_d(u)) - \psi(t, u)\|_{\mathcal{H}}\big),$$

where $\mathbf{P}_d$ denotes the orthogonal projection onto $\mathcal{H}_d$, the space of operators on H with rank at most $d \in \mathbb{N}$ spanned by a set of basis vectors $\{\mathbf{e}_{i,j} : 1 \leq i \leq j \leq d\}$.

Finite-rank operator-valued affine processes For every rank $d \in \mathbb{N}$ we show in Proposition 6.2 the existence of an affine process on $\mathcal{H}_d^+$, the cone of positive operators in $\mathcal{H}_d$, associated with the Galerkin approximations ϕ_d and ψ_d. More precisely, we prove the existence of a unique finite-variation Markov process $(X^d, (\mathbb{P}_x^d)_{x\in\mathcal{H}^+})$ on $\mathcal{H}_d^+$ such that for all initial values $X_0^d = \mathbf{P}_d(x) \in \mathcal{H}_d^+$ the following affine transform formula holds true:

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\mathrm{e}^{-\langle X_t^d, u_d\rangle}\right] = \mathrm{e}^{-\phi_d(t,u_d)-\langle \mathbf{P}_d(x), \psi_d(t,u_d)\rangle}, \quad t \geq 0,\, u_d \in \mathcal{H}_d^+.$$

Existence and weak convergence In Theorem 6.4 we show that the sequence of *finite-rank operator-valued affine processes* $(X^d)_{d\in\mathbb{N}}$ from above is *tight* on $D(\mathbb{R}^+, \mathcal{H}^+)$, the Skorohod space of all càdlàg paths from $\mathbb{R}^+$ into $\mathcal{H}^+$. Moreover, we prove that the processes $(X^d)_{d\in\mathbb{N}}$ solve the *martingale problems* for an associated sequence of Kolmogorov-type operators. From this, convergence of the operators and tightness, we derive the weak convergence of $(X^d)_{d\in\mathbb{N}}$ to a unique càdlàg Markov process X with values in $\mathcal{H}^+$ satisfying the affine transform formula (1.3). In addition to that, we present explicit convergence rates for the associated Laplace transforms and give a class of examples for affine process on $\mathcal{H}^+$ with jumps of infinite-variation and describe their finite-rank approximations.

6.1.1 Layout of the chapter

This chapter is structured as follows: In Section 6.2 we introduce some additional notation, in particular the notion of *finite-rank projection schemes*. Section 6.3 is devoted to the presentation of our main results. More specifically, in Section 6.3.1 we introduce Galerkin approximations of the generalized Riccati equations and provide the convergence rates. In Section 6.3.2 we state our results on the existence of finite-rank operator-valued affine processes associated with the Galerkin approximations. In Section 6.3.3 we present our main result on the existence and approximation of affine processes on positive Hilbert-Schmidt operators. To illustrate our main findings, we give in Section 6.4 a concrete example of an affine process on positive Hilbert-Schmidt operators of infinite-variation and specify its finite-rank approximations. Applications of our results in the context of infinite-dimensional stochastic covariance modeling are discussed in Section 6.5. All the proofs are contained in the subsequent four chapters. More precisely, in Section 6.6 we prove the existence and convergence of the Galerkin approximations. In Section 6.7 we construct associated sequences of finite-rank operator-valued affine processes. In Section 6.8 we prove weak convergence of the sequences of finite-rank processes to a class of affine processes with values in the cone of positive Hilbert-Schmidt operators. In Section 6.9 we prove our results concerning the finite-rank approximations of stochastic covariance models.

6.2 Notation and preliminaries

Throughout this chapter we let $(H, \langle\cdot,\cdot\rangle_H)$ be a real separable Hilbert space and denote its norm by $\|\cdot\|_H$. Moreover, let $(V, (\cdot,\cdot)_V)$ be a second separable Hilbert space with norm $\|\cdot\|_V$. For a subspace $V \subseteq H$ we say that V is *continuously embedded* in H, if there exists a constant C such that $\|u\|_H \leq C\|u\|_V$ for all $u \in V$. If in addition the embedding operator of V into H is compact, then we say that V is *compactly embedded* into H and write $V \subset\subset H$.
Let $(B, \|\cdot\|_B)$ be a real separable Banach space. Then for any $F \subseteq B$, we denote by $D(\mathbb{R}^+, F)$ the space of all càdlàg path from $\mathbb{R}^+$ into F equipped with the Skorohod topology, see e.g. [83].

Finite-rank projection schemes in Hilbert-Schmidt operator space Let $(e_i)_{i\in\mathbb{N}}$ be an orthonormal basis of H which can be chosen arbitrarily, but is fixed throughout the section. For every $d \in \mathbb{N}$ we denote by H_d the d-dimensional subspace of H spanned by the first d basis vectors, i.e.

$$H_d \coloneqq \operatorname{lin}\{e_i \colon i = 1, \ldots, d\}$$

We denote the orthogonal projection of H onto H_d, with respect to the inner product $\langle\cdot,\cdot\rangle_H$, by P_d. For every $i \in \mathbb{N}$ we set $\mathbf{e}_{i,i} \coloneqq e_i \otimes e_i$ and for all $i \neq j$ set $\mathbf{e}_{i,j} \coloneqq \frac{1}{\sqrt{2}}(e_i \otimes e_j + e_j \otimes e_i)$. Note that $\|\mathbf{e}_{i,j}\| = 1$, $\mathbf{e}_{i,j} = \mathbf{e}_{j,i}$ for every $i, j \in \mathbb{N}$ and it can be seen that the family $\{\mathbf{e}_{i,j}\}_{i\leq j\in\mathbb{N}} \coloneqq \{\mathbf{e}_{i,j} \colon i, j \in \mathbb{N}, i \leq j\}$ is an orthonormal basis of $\mathcal{H}$. For every $d \in \mathbb{N}$, we let $\mathcal{H}_d$ stand for the finite-dimensional subspace of $\mathcal{H}$ spanned by the family $\{\mathbf{e}_{i,j} \colon 1 \leq i \leq j \leq d\}$, i.e.

$$\mathcal{H}_d \coloneqq \operatorname{lin}\{\mathbf{e}_{i,j} \colon 1 \leq i \leq j \leq d\}.$$

We denote the orthogonal projection of $\mathcal{H}$ onto $\mathcal{H}_d$, with respect to the inner product $\langle\cdot,\cdot\rangle$, by $\mathbf{P}_d$ and note that for every $d \in \mathbb{N}$ and $u \in \mathcal{H}$ we have $\mathbf{P}_d(u) = \mathsf{P}_d u \mathsf{P}_d$. Moreover, every operator in $\mathcal{H}_d$ is self-adjoint and of rank at most d. We write $\mathbf{P}_d^{\perp}(u) \coloneqq u - \mathbf{P}_d(u)$ and note that $\lim_{d\to\infty}\|\mathbf{P}_d^{\perp}x\| = 0$ for all $x \in \mathcal{H}$. In addition, it can be seen that $\mathcal{H}_d = \{u\mathsf{P}_d \colon u \in \mathcal{L}(H_d),\, u = u^*\}$ and for the cone of all positive self-adjoint operators in $\mathcal{H}_d$, denoted by $\mathcal{H}_d^+$, we have

$$\mathcal{H}_d^+ \coloneqq \{u\mathsf{P}_d \colon u \in \mathcal{L}(H_d),\, u = u^*,\, \langle uh, h\rangle_H \geq 0\ \forall h \in H_d\}.$$

Note further that $\mathcal{H}_d^+ \subseteq \mathcal{H}_{d+1}^+ \subseteq \mathcal{H}^+$ for all $d \in \mathbb{N}$. For more details on the subspace of finite-rank operators in the ambient space of all Hilbert-Schmidt operators see [128]. As in [75], we call a sequence $(\mathcal{H}_d, \mathbf{P}_d)_{d\in\mathbb{N}}$ defined as above a *projection scheme* in $\mathcal{H}$ (with respect to the orthonormal basis $\{\mathbf{e}_{i,j}\}_{i\leq j\in\mathbb{N}}$).

6.3 Finite-rank approximation of affine processes on positive Hilbert-Schmidt operators

Let $(H, \langle\cdot,\cdot\rangle_H)$ and $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be as in Section 6.2 and let $(\mathcal{H}_d, \mathbf{P}_d)_{d\in\mathbb{N}}$ be a projection scheme in $\mathcal{H}$ with respect to an orthonormal basis $\{\mathbf{e}_{i,j}\}_{i\leq j\in\mathbb{N}}$ of $\mathcal{H}$. In the following Sections 6.3.1, 6.3.2 and 6.3.3 we give comprehensive versions of our main results concerning the finite-rank approximation of affine processes on Hilbert-Schmidt operators.

6.3.1 Galerkin approximation of the generalized Riccati equation

Recall the two functions F and R from (2.6). Then for every $d \in \mathbb{N}$ we define $R_d\colon \mathcal{H}_d^+ \to \mathcal{H}_d$ and $F_d\colon \mathcal{H}_d^+ \to \mathbb{R}$ as $R_d(u_d) \coloneqq \mathbf{P}_d(R(u_d))$ and $F_d(u_d) \coloneqq F(u_d)$ for $u_d \in \mathcal{H}_d^+$. In particular, for every $u \in \mathcal{H}^+$ we have $F_d(\mathbf{P}_d(u)) = F(\mathbf{P}_d(u))$ and $R_d(\mathbf{P}_d(u)) = \mathbf{P}_d(R(\mathbf{P}_d(u)))$. The next proposition asserts the convergence of Galerkin type approximations $\big(\phi_d(\cdot,\mathbf{P}_d(u)), \psi_d(\cdot,\mathbf{P}_d(u))\big)$, with respect to the projection scheme $(\mathcal{H}_d, \mathbf{P}_d)_{d\in\mathbb{N}}$, to the unique solution $(\phi(\cdot,u), \psi(\cdot,u))$ of (2.8). Moreover, we provide explicit convergence rates that hold uniformly on compact time intervals.

Proposition 6.1. *Let (b, B, m, μ) be an admissible parameter set as in Definition 2.3 and for every $u \in \mathcal{H}^+$ denote by $(\phi(\cdot,u), \psi(\cdot,u))$ the unique solution of (2.8). Then for every $d \in \mathbb{N}$, $T > 0$ and $u \in \mathcal{H}^+$ there exists a unique solution $\big(\phi_d(\cdot,\mathbf{P}_d(u)), \psi_d(\cdot,\mathbf{P}_d(u))\big)$ of*

$$\begin{cases} \dfrac{\partial \phi_d(t,\mathbf{P}_d(u))}{\partial t} = F_d\big(\psi_d(t,\mathbf{P}_d(u))\big), \quad \phi_d\big(0,\mathbf{P}_d(u)\big) = 0, & (6.1a) \\ \dfrac{\partial \psi_d(t,\mathbf{P}_d(u))}{\partial t} = R_d\big(\psi_d(t,\mathbf{P}_d(u))\big), \quad \psi_d\big(0,\mathbf{P}_d(u)\big) = \mathbf{P}_d(u), & (6.1b) \end{cases}$$

*such that $\phi_d(\cdot,\mathbf{P}_d(u)) \in C^1(\mathbb{R}^+,\mathbb{R}^+)$, $\psi_d(\cdot,\mathbf{P}_d(u)) \in C^1(\mathbb{R}^+,\mathcal{H}_d^+)$ and to which we refer as the d^{th}-*Galerkin approximation *of $(\phi(\cdot,u), \psi(\cdot,u))$. Moreover, there exists a constant $K \geq 0$, independent of $d \in \mathbb{N}$, such that*

$$\sup_{t\in[0,T]} \big(|\phi_d(t,\mathbf{P}_d(u)) - \phi(t,u)| + \|\psi_d(t,\mathbf{P}_d(u)) - \psi(t,u)\|\big) \leq K C_{T,d}, \tag{6.2}$$

where $C_{T,d}$ is given by

$$C_{T,d} = \sup_{t\in[0,T]} \big(\|\mathbf{P}_d^{\perp}(\mathrm{e}^{tB^*}u)\| + \|\mathbf{P}_d^{\perp}(\mathrm{e}^{tB^*}\mu(\mathcal{H}^+\setminus\{0\}))\|\big). \tag{6.3}$$

In particular, for every $u \in \mathcal{H}^+$ the sequence $\big(\phi_d(t,\mathbf{P}_d(u)), \psi_d(t,\mathbf{P}_d(u))\big)_{d\in\mathbb{N}}$ converges to $(\phi(t,u), \psi(t,u))$ uniformly on compact sets in time.

6.3.2 Finite-rank operator-valued affine processes

For every $d \in \mathbb{N}$ we define the set $\mathcal{D}^d := \{\mathrm{e}^{-\langle \cdot, u\rangle} \colon u \in \mathcal{H}_d^+\} \subseteq C(\mathcal{H}^+, \mathbb{R})$ and the operator $\mathcal{G}^d \colon \mathcal{D}^d \to C(\mathcal{H}^+, \mathbb{R})$ as

$$\mathcal{G}^d\, \mathrm{e}^{-\langle \,\cdot\,, u\rangle}(x) := \big(- F_d(u) - \langle x, R_d(u)\rangle\big)\, \mathrm{e}^{-\langle x, u\rangle}, \quad x \in \mathcal{H}^+, \tag{6.4}$$

where F_d and R_d are defined as in Section 6.3.1 above. The next proposition asserts that the d^{th}-Galerkin approximation $\big(\phi_d(\cdot, \mathbf{P}_d(u)), \psi_d(\cdot, \mathbf{P}_d(u))\big)$ gives rise to an affine Markov process X^d with values in $\mathcal{H}_d^+$ that solves the martingale problem for $\mathcal{G}^d$ with $X_0^d = \mathbf{P}_d(x)$ on a suitable stochastic basis. The proof of this proposition is based on [42], i.e. we essentially use the matrix-valued theory and a subsequent transformation to finite-rank operator-valued processes to establish the existence (see Section 6.7).

Proposition 6.2. *Let the assumptions of Proposition 6.1 hold. Then for every $d \in \mathbb{N}$ the following holds true:*

i) There exists a unique Markov process $(X^d, (\mathbb{P}_x^d)_{x\in\mathcal{H}^+})$ on $\mathcal{H}^+$, where $\mathbb{P}_x^d$ denotes the law of X^d given $X_0^d = \mathbf{P}_d(x)$, with paths in $D(\mathbb{R}^+, \mathcal{H}^+)$ and such that for every $x \in \mathcal{H}^+$ we have $\mathbb{P}_x^d(\{X_t^d \in \mathcal{H}_d^+ \colon t \geq 0\}) = 1$ and the following affine transform formula holds true:

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\mathrm{e}^{-\langle X_t^d, \mathbf{P}_d(u)\rangle}\right] = \mathrm{e}^{-\phi_d(t,\mathbf{P}_d(u)) - \langle \mathbf{P}_d(x), \psi_d(t, \mathbf{P}_d(u))\rangle}, \; t \geq 0, \, u \in \mathcal{H}^+, \tag{6.5}$$

for $\big(\phi_d(\cdot, \mathbf{P}_d(u)), \psi_d(\cdot, \mathbf{P}_d(u))\big)$ the unique solution of (6.1).

ii) For every $x \in \mathcal{H}^+$ and every $u \in \mathcal{H}^+$ the process

$$\Big(\mathrm{e}^{-\langle X_t^d, \mathbf{P}_d(u)\rangle} - \mathrm{e}^{-\langle \mathbf{P}_d(x), \mathbf{P}_d(u)\rangle} - \int_0^t (\mathcal{G}^d\, \mathrm{e}^{-\langle\,\cdot\,, \mathbf{P}_d(u)\rangle})(X_s^d)\, \mathrm{d}s\Big)_{t\geq 0}, \tag{6.6}$$

is a martingale with respect to the stochastic basis $(\Omega, \bar{\mathcal{F}}^d, \bar{\mathbb{F}}^d, \mathbb{P}_x^d)$, where $\Omega = D(\mathbb{R}^+, \mathcal{H}^+)$ and $\bar{\mathbb{F}}^d = (\bar{\mathcal{F}}_t^d)_{t\geq 0}$ denotes the augmentation of the natural filtration $(\mathcal{F}_t^d)_{t\geq 0}$ of X^d with respect to the measure $\mathbb{P}_x^d$ from part i).

Remark 6.3. In Proposition 6.22 and 6.23 below we give some additional properties of the processes $(X^d)_{d\in\mathbb{N}}$, namely: We present a semimartingale representation of $(X_t^d)_{t\geq 0}$, give a more detailed description of the operator $\mathcal{G}^d$ and show that all the processes $(X^d)_{d\in\mathbb{N}}$ must be of finite-variation.

6.3.3 Existence and weak convergence

In the following we recall the setting of [139], namely: We let $(H, \langle\cdot,\cdot\rangle_H)$ be a separable (infinite-dimensional) Hilbert space and assume that $(V, (\cdot,\cdot)_V)$ is a second separable Hilbert space such that $V \subseteq H$. We assume the following:

Assumption $\mathcal{E}$. $(V, (\cdot,\cdot)_V)$ is densely and compactly embedded in $(H, \langle\cdot,\cdot\rangle_H)$.

We denote by V^* the Hilbert space dual of V with respect to the inner-product $\langle\cdot,\cdot\rangle_H$ and by identifying H with its dual space H^* we obtain the *Gelfand triple*: $V \hookrightarrow H \hookrightarrow V^*$. Moreover, we define the space $\mathcal{V}$ by

$$\mathcal{V} \coloneqq \mathcal{L}_2(V^*, H) \cap \mathcal{L}_2(H, V), \tag{6.7}$$

equip $\mathcal{V}$ with the inner-product $\langle\cdot,\cdot\rangle_{\mathcal{V}} \coloneqq \langle\cdot,\cdot\rangle_{\mathcal{L}_2(V^*,H)} + \langle\cdot,\cdot\rangle_{\mathcal{L}_2(H,V)}$ and denote the induced norm by $\|\cdot\|_{\mathcal{V}}$. Note that $\mathcal{V} \subseteq \mathcal{L}_2(H)$ and $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}})$ is a Hilbert space, which is itself densely and compactly embedded in $(\mathcal{L}_2(H), \langle\cdot,\cdot\rangle)$, see [139, Proposition 2.1]. We define $\mathcal{V}_0 \subseteq \mathcal{V}$ to be the subspace of all self-adjoint operators (with respect to $\langle\cdot,\cdot\rangle$), i.e. $\mathcal{V}_0 \coloneqq \mathcal{V} \cap \mathcal{H}$.
We now proceed with our main result on the existence and approximation of càdlàg affine pure-jump processes on positive Hilbert-Schmidt operators.

Theorem 6.4. *Let (b, B, m, μ) be an admissible parameter set as in Definition 2.3 and let Assumption $\mathcal{E}$ be satisfied. Then the following holds true:*

i) There exists a unique affine process $(X, (\mathbb{P}_x)_{x\in\mathcal{H}^+})$ on $\mathcal{H}^+$, where $\mathbb{P}_x$ denotes the law of X given $X_0 = x$, with paths in $D(\mathbb{R}^+, \mathcal{H}^+)$ and such that for every $x \in \mathcal{H}^+$ we have

$$\mathbb{E}_{\mathbb{P}_x}\left[\mathrm{e}^{-\langle X_t, u\rangle}\right] = \mathrm{e}^{-\phi(t,u)-\langle x,\psi(t,u)\rangle}, \quad t \geq 0,\ u \in \mathcal{H}^+, \tag{6.8}$$

for $(\phi(\cdot,u), \psi(\cdot,u))$ the unique solution of (2.8).

ii) Moreover, let $(X^d)_{d\in\mathbb{N}}$ be the sequence of finite-rank operator-valued affine processes from Proposition 6.2. Then $(X^d)_{d\in\mathbb{N}}$ converges weakly to X on $D(\mathbb{R}^+, \mathcal{H}^+)$ equipped with the Skorohod topology, i.e. for all functions $f \in C(D(\mathbb{R}^+, \mathcal{H}^+), \mathbb{R})$ we have

$$\mathbb{E}_{\mathbb{P}_x^d}\left[f(X^d)\right] \to \mathbb{E}_{\mathbb{P}_x}\left[f(X)\right], \quad \text{as } d \to \infty.$$

If, in addition, we have $\|\mu(\mathcal{H}^+ \setminus \{0\})\|_{\mathcal{V}} < \infty$ and $B^*(\mathcal{V}_0) \subseteq \mathcal{V}_0$, then for every $T > 0$ and $u \in \mathcal{H}^+$ with $\|u\|_{\mathcal{V}} \leq 1$ there exists a constant $C_T > 0$ such that for all $d \in \mathbb{N}$:

$$\sup_{t\in[0,T]}\left|\mathbb{E}_{\mathbb{P}_x}\left[\mathrm{e}^{-\langle u, X_t\rangle}\right] - \mathbb{E}_{\mathbb{P}_x^d}\left[\mathrm{e}^{-\langle u, X_t^d\rangle}\right]\right| \leq C_T \|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}(1 + \|x\|). \tag{6.9}$$

Remark 6.5. i) A detailed version of the constant C_T in (6.9) is given in Corollary 6.12 below. Note that if the conditions $\|\mu(\mathcal{H}^+ \setminus \{0\})\|_{\mathcal{V}} < \infty$ and $B^*(\mathcal{V}_0) \subseteq \mathcal{V}_0$ do not hold, then there still exists a constant $\tilde{K}$, independent of $d \in \mathbb{N}$, such that (6.9) holds with right-hand side $\tilde{K} C_{T,d}(1 + \|x\|)$ where $C_{T,d}$ is as in (6.3).

ii) Also, note that the error bound C_T depends on the norm of $u \in \mathcal{H}^+$. Indeed, in the the proof of Proposition 6.1 and Theorem 6.4 ii) we make this dependence precise, which makes the dependence of the convergence rates on the initial values more explicit.

iii) Analogous to Theorem 6.2, we will show in Proposition 6.27 that X is the (unique) solution to the martingale problem for the operator $\mathcal{G}$ given by $\mathcal{G}\, \mathrm{e}^{-\langle \cdot\, u \rangle} = (-F(u) - \langle \cdot\, , R(u) \rangle)\, \mathrm{e}^{-\langle \cdot, u \rangle}$ on the set $\mathcal{D} := \big\{ \mathrm{e}^{-\langle \cdot, u \rangle} \colon u \in \mathcal{H}^+ \big\}$. To keep this section reasonably concise and focused on our main results, we relegate this (and other) side results to the latter Sections 6.7 and 6.8.

In the following proposition we assert that the affine process $(X, \mathbb{P}_x)$ from Theorem 6.4 i) is a semimartingale and we specify its characteristic triplet, see Definition 3.1. This proposition is analogous to Proposition 3.5, but this time we can drop Assumption $\mathcal{A}$.

Proposition 6.6. *For every $x \in \mathcal{H}^+$ the process X from Theorem 6.4 i) is a square-integrable semimartingale with respect to the stochastic basis $(\Omega, \bar{\mathcal{F}}, \bar{\mathbb{F}}, \mathbb{P}_x)$, where $\bar{\mathbb{F}} = (\bar{\mathcal{F}}_t)_{t \geq 0}$ denotes the augmentation of the natural filtration $(\mathcal{F}_t)_{t \geq 0}$ of X with respect to the measure $\mathbb{P}_x$. The characteristic triplet (A, C, ν^X) of X, with respect to χ, is given by:*

$$A_t = \int_0^t \big(b + B(X_s) \big)\, \mathrm{d}s, \quad t \geq 0, \tag{6.10}$$

$$C_t = 0, \quad t \geq 0, \tag{6.11}$$

$$\nu^X(\mathrm{d}t, \mathrm{d}\xi) = \big(m(\mathrm{d}\xi) + M(X_t, \mathrm{d}\xi) \big)\, \mathrm{d}t, \tag{6.12}$$

and for every $t \geq 0$ the following representation holds true

$$X_t = x + \int_0^t \Big(b + B(X_s) + \int_{\mathcal{H}^+ \cap \{\|\xi\| > 1\}} \xi\, \nu(X_s, \mathrm{d}\xi)) \Big)\ \mathrm{d}s + \bar{J}_t, \tag{6.13}$$

where $\nu(x, \mathrm{d}\xi) = m(\mathrm{d}\xi) + M(x, \mathrm{d}\xi)$ and $(\bar{J}_t)_{t \geq 0}$ is a purely discontinuous square-integrable martingale of the $\nu(\mathrm{d}t, \mathrm{d}\xi)$-compensated jumps of X .

Proof. Follows immediately from Proposition 3.5 and Proposition 6.27. □

6.4 Example: An affine process with jumps of infinite-variation

In this section we follow up on Section 3.4 and give another example of an affine process on positive Hilbert-Schmidt operators by specifying a suitable admissible parameter set (b, B, m, μ). The novelty here is, that this example admits jumps of infinite-variation (and still has càdlàg paths). Similarly, to the instantaneous covariance process that we constructed in the *general affine stochastic covariance model with state-dependent jumps* example in Section 3.4.4, this process is driven by a jump process supported on the diagonal elements $(\mathbf{e}_{n,n})_{n\in\mathbb{N}}$. Indeed, we specify the parameter set (b, B, m, μ) as follows:

i) We set $m(\mathrm{d}\xi) = \|\xi\|^{-2}\eta(\mathrm{d}\xi)$ for $\eta\colon \mathcal{B}(\mathcal{H}^+ \setminus \{0\}) \to [0, \infty]$ given by

$$\eta(A) = \sum_{n\in\mathbb{N}} \eta_n(\{\lambda \in (0, \infty)\colon \lambda\mathbf{e}_{n,n} \in A\}), \quad A \in \mathcal{B}(\mathcal{H}^+ \setminus \{0\}), \tag{6.14}$$

where $(\eta_n)_{n\in\mathbb{N}}$ is a sequence of finite measures on $\mathcal{B}((0, \infty))$ such that for all $u \in \mathcal{H}$ we have

$$\sum_{n\in\mathbb{N}} \Big(\int_0^1 \lambda^{-1}\eta_n(\mathrm{d}\lambda)\Big)\langle u, \mathbf{e}_{n,n}\rangle < \infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \eta_n((0, \infty)) < \infty. \tag{6.15}$$

ii) We let $\tilde{b} \in \mathcal{H}^+$ be arbitrary and let $I_m \in \mathcal{H}$ be such that for all $u \in \mathcal{H}$ we have

$$\langle I_m, u\rangle = \sum_{n\in\mathbb{N}} \Big(\int_0^1 \lambda^{-1}\eta_n(\mathrm{d}\lambda)\Big)\langle u, \mathbf{e}_{n,n}\rangle,$$

and define $b := \tilde{b} + I_m$.

iii) We let $(g_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}^+$ and define $\mu(\mathrm{d}\xi)\colon \mathcal{B}(\mathcal{H}^+ \setminus \{0\}) \to \mathcal{H}^+$ by

$$\mu(A) = \sum_{n\in\mathbb{N}} g_n\mu_n(\{\lambda \in (0, \infty)\colon \lambda\mathbf{e}_{n,n} \in A\}), \tag{6.16}$$

where $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of finite-measures on $\mathcal{B}((0, \infty))$ such that for all $x \in \mathcal{H}^+$ and $u \in \mathcal{H}$ we have

$$\sum_{n\in\mathbb{N}} \Big(\int_0^1 \lambda^{-1}\mu_n(\mathrm{d}\lambda)\Big)\langle g_n, x\rangle\langle u, \mathbf{e}_{n,n}\rangle < \infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} g_n\mu_n((0, \infty)) \in \mathcal{H}^+. \tag{6.17}$$

Moreover, for every $x \in \mathcal{H}^+$ we set $M(x, \mathrm{d}\xi) := \|\xi\|^{-2}\langle x, \mu(\mathrm{d}\xi)\rangle$.

iv) Finally, let C be a bounded linear operator on H and let $\Gamma \in \mathcal{L}(\mathcal{H})$ be such that for all $u, x \in \mathcal{H}$ we have

$$\langle \Gamma(x), u\rangle = \sum_{n\in\mathbb{N}} \Big(\int_0^1 \lambda^{-1}\mu_n(\mathrm{d}\lambda)\Big)\langle g_n, x\rangle\langle u, \mathbf{e}_{n,n}\rangle$$

Then, we define $B \in \mathcal{L}(\mathcal{H})$ by $B(u) := Cu + uC^* + \Gamma(u)$.

Similarly to Section 3.4 it can be seen that the parameter set (b, B, m, μ) is correctly set up to satisfy the conditions of Definition 2.3. Moreover, we introduce the following finite-rank operator-valued approximations of the set (b, B, m, μ): For every $d \in \mathbb{N}$ and $x \in \mathcal{H}_d^+$ set

i) $m_d(A) := \sum_{n=1}^{d} m_n(\{\lambda \in (0,\infty) \colon \lambda\mathbf{e}_{n,n} \in A\})$ for $A \in \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\})$;

ii) $c_d := \tilde{b}_d + \sum_{n=1}^{d} \big(\int_0^1 \lambda^{-1}\eta_n(\mathrm{d}\lambda)\big)\mathbf{e}_{n,n}$ for $\tilde{b}_d = \mathbf{P}_d(\tilde{b})$;

iii) $\mu_d(A) := \sum_{n=1}^{d} \mathbf{P}_d(g_n)\mu_n(\{\lambda \in (0,\infty) \colon \lambda\mathbf{e}_{n,n} \in A\})$ for $A \in \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\})$;

iv) $D_d(x) := C_d x + xC_d^* + \sum_{n=1}^{d} \big(\int_0^1 \lambda^{-1}\mu_n(\mathrm{d}\lambda)\big)\langle g_n, x\rangle\langle u, \mathbf{e}_{n,n}\rangle$ for $C_d := \mathsf{P}_d C$.

Proposition 6.22 below guarantees the existence of an affine process $(X_t^d)_{t\geq 0}$ on $\mathcal{H}_d^+$ associated with (b_d, B_d, m_d, μ_d) admitting the following representation:

$$X_t^d = X_0^d + \int_0^t \Big(\tilde{b}_d + C_d X_s^d + X_s^d C_d^*\Big)\,\mathrm{d}s + \int_0^t \int_{\mathcal{H}_d^+} \xi\, \mu^{X^d}(\mathrm{d}t, \mathrm{d}\xi), \quad t \geq 0, \tag{6.18}$$

where $\mu^{X^d}(\mathrm{d}t, \mathrm{d}\xi)$ denotes the random jump-measure of X^d with compensator $\nu^{X^d}(\mathrm{d}t, \mathrm{d}\xi) = (m_d(\mathrm{d}\xi) + \|\xi\|^{-2}\langle X_t^d, \mu_d(\mathrm{d}\xi)\rangle)\,\mathrm{d}t$. Apparently, we see that the process X^d given by (6.18) is the affine process asserted by Theorem 6.2. Therefore, it follows from Theorem 6.4 and Proposition 6.6 that the sequence $(X^d)_{d\in\mathbb{N}}$ converges weakly to an affine process $X = (X_t)_{t\geq 0}$ on $\mathcal{H}^+$ with representation

$$\begin{aligned} X_t = X_0 &+ \int_0^t \Big(\tilde{b} + I_m + CX_s + X_s C^* + \Gamma(X_s)\Big)\,\mathrm{d}s \\ &+ \int_0^t \Big(\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \xi\,(m(\mathrm{d}\xi) + M(X_s, \mathrm{d}\xi)\Big)\,\mathrm{d}s + \bar{J}_t, \quad t \geq 0, \end{aligned} \tag{6.19}$$

where $M(x, \mathrm{d}\xi)$ is as in (2.3) and the process $(\bar{J}_t)_{t\geq 0}$ is a purely-discontinuous square-integrable martingale.

Note that in (6.18) the $\mathcal{H}_d$-projections of the two drift terms I_m and Γ are killed by the compensator of the jump-process $\big(\int_0^t \int_{\mathcal{H}_d^+} \xi\, \mu^{X^d}(\mathrm{d}t, \mathrm{d}\xi)\big)_{t\geq 0}$, as for every $d \in \mathbb{N}$ the jumps are of finite-variation. In the limit case (6.19), however, I_m and Γ must occur in the drift (and jump-part) again as the driving jump-process possibly converges to a process of infinite-variation. That this is indeed possible is shown in the following paragraph:

Can the jumps be really of infinite-variation?

We now concretize the parameter specification to show that affine processes with jumps of infinite-variation are contained in the class of processes given by Theorem 6.4. For every $n \in \mathbb{N}$ let $g_n = g$ for some $g \in \mathcal{H}^+$ and define

$$\mu(\mathrm{d}\xi) := g \sum_{n\in\mathbb{N}} \frac{1}{n^2} \delta_{\frac{1}{n}\mathbf{e}_{n,n}}(\mathrm{d}\xi).$$

First, we observe that

$$\mu(\mathcal{H}^+ \setminus \{0\}) = g \sum_{n\in\mathbb{N}} \frac{1}{n^2} \delta_{\frac{1}{n}\mathbf{e}_{n,n}}(\mathcal{H}^+ \setminus \{0\}) = \frac{\pi}{6}\, g \in \mathcal{H}^+ .$$

Next, we show that the measure $M(x, \mathrm{d}\xi)$ is of infinite variation. Indeed, we have

$$\begin{aligned}
\int_{0<\|\xi\|\leq 1} \|\xi\| M(x, \mathrm{d}\xi) &= \int_{0<\|\xi\|\leq 1} \frac{1}{\|\xi\|} \sum_{n\in\mathbb{N}} \frac{1}{n^2} \langle x, g\rangle \delta_{n^{-1}\mathbf{e}_{n,n}}(\mathrm{d}\xi) \\
&= \sum_{n\in\mathbb{N}} \frac{1}{\|n^{-1}\mathbf{e}_{n,n}\|} \frac{1}{n^2} \langle x, g\rangle \\
&= \langle x, g\rangle \sum_{n\in\mathbb{N}} \frac{1}{n} = \infty,
\end{aligned}$$

which implies that also the measure ν in (6.12) is of infinite variation and thus also the associated affine process X. However, note that for every $u \in \mathcal{H}^+$ we have

$$\begin{aligned}
\int_{0<\|\xi\|\leq 1} \langle \xi, u\rangle M(x, \mathrm{d}\xi) &= \sum_{n\in\mathbb{N}} \langle n^{-1}\mathbf{e}_{n,n}, u\rangle \frac{1}{\|n^{-1}\mathbf{e}_{n,n}\|^2} \frac{1}{n^2} \langle x, g\rangle \\
&= \langle x, g\rangle \sum_{n\in\mathbb{N}} \frac{1}{n} \langle u, \mathbf{e}_{n,n}\rangle \\
&\leq \langle x, g\rangle \big(\sum_{n\in\mathbb{N}} \frac{1}{n^2}\big)^{1/2} \big(\sum_{n\in\mathbb{N}} \langle u, \mathbf{e}_{n,n}\rangle^2\big)^{1/2} < \infty,
\end{aligned}$$

which proves that the admissibility condition in equation (2.4) is satisfied although $M(x, \mathrm{d}\xi)$ is of infinite variation.

6.5 Finite-rank stochastic covariance models

In this section we study finite-rank approximation of affine stochastic covariance models of the form in Definition 3.8 and Remark 3.10. More precisely, let X be an affine process on $\mathcal{H}^+$ associated with an admissible parameter set (b, B, m, μ) and let $(Y_t)_{t\geq 0}$ be given by

$$\begin{cases} \mathrm{d}Y_t = \mathcal{A}Y_t \,\mathrm{d}t + D^{1/2}X_t^{1/2}\,\mathrm{d}W_t, \\ Y_0 = y, \end{cases} \tag{6.20}$$

where we assume that $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on H, $D \in \mathcal{H}^+$ and $(W_t)_{t\geq 0}$ denotes a H-cylindrical Brownian motion, which is independent of the process X.
It follows from Theorem 6.4 that every affine process X has a version with càdlàg paths and thus, by Lemma 3.7 and Theorem 3.14, we see that the stochastic covariance model (Y, X) is well-defined and has the affine property. In particular, to ensure the existence of càdlàg versions we are not relying on Lemma 3.4 anymore, which allows us to give a great variety of examples for instantaneous covariance processes that go beyond such with Lévy noise or finite-activity jumps, as for example the class of examples in Section 6.4.
Moreover, Theorem 6.4 ii) provides us with a convenient approximation scheme for affine processes on positive Hilbert-Schmidt operators. In the context of affine stochastic covariance models, this allows us to approximate the (infinite-dimensional) instantaneous covariance process by finite-rank operator-valued processes. In Section 6.5.1, we show that this approach yields tractable finite-rank approximation of affine stochastic covariance models and we prove that the weak convergence of the instantaneous covariance processes implies the weak convergence of the finite-rank stochastic covariance models (Y^d, X^d) to (Y, X), see Proposition 6.7 below. Moreover, we provide convergence rates for the associated characteristic functions which allow us to quantify the approximation error, which is relevant for applications in, e.g., option-pricing, see Section 4.3.

6.5.1 Finite-rank approximation of affine stochastic covariance models

Let $(V, (\cdot,\cdot)_V)$ and $(H, \langle\cdot,\cdot\rangle_H)$ be two separable Hilbert spaces such that Assumption $\mathcal{E}$ is satisfied. Let $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ and $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}})$ be as in Section 6.3.3 and let (Y, X) be an affine stochastic covariance model on H given by (6.20), with X being an affine Markovian semimartingale on $\mathcal{H}^+$ with admissible parameters (b, B, m, μ) and the model is defined on the probability basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ as in Section 3.2.2. We prove the following assertions on finite-rank approximation of infinite-dimensional stochastic covariance models.

Proposition 6.7. *Let (Y, X) be given by (6.20) with X as in Theorem 6.4. Moreover, for every $d \in \mathbb{N}$ let X^d be as in Proposition 6.2 and denote by (Y^d, X^d) the stochastic covariance model with X replaced by X^d such that $X_0^d = \mathbf{P}_d(x)$. In addition, assume that $(\mathcal{A}, \mathrm{dom}(\mathcal{A}))$ generates a strongly continuous semigroup on the space V and $D^{1/2} \in \mathcal{L}_2(H, V)$. Then the following holds true:*

i) For every $y \in H$, $x \in \mathcal{H}^+$ the process (Y^d, X^d) with $(Y_0^d, X_0^d) = (y, \mathbf{P}_d(x))$ is well-defined and for every $u = (u_1, u_2) \in \mathrm{i}H \times \mathcal{H}_d^+$ we have

$$\mathbb{E}\left[\mathrm{e}^{\langle Y_t^d, u_1\rangle_H - \langle X_t^d, u_2\rangle}\right] = \mathrm{e}^{-\Phi_d(t,u) + \langle y, \psi_1(t,u)\rangle_H - \langle \mathbf{P}_d(x), \psi_{2,d}(t,u)\rangle}, \quad t \geq 0, \tag{6.21}$$

for $(\Phi_d(\cdot, u), (\psi_1(\cdot, u), \psi_{2,d}(\cdot, u)))$ the unique solution of (3.24) with F and R in (3.24a) and (3.24c) replaced by $F \circ \mathbf{P}_d$ and $\mathbf{P}_d \circ R \circ \mathbf{P}_d$, respectively.

ii) The sequence $(Y^d, X^d)_{d\in\mathbb{N}}$ converges weakly to (Y, X) as the rank $d \in \mathbb{N}$ tends to infinity, i.e. for all functions $f \in C(D(\mathbb{R}^+, H \times \mathcal{H}^+), \mathbb{R})$ we have

$$\mathbb{E}\left[f(Y^d, X^d)\right] \to \mathbb{E}\left[f(Y, X)\right], \quad \text{as } d \to \infty.$$

If, in addition, $\|\mu(\mathcal{H}^+ \setminus \{0\})\|_{\mathcal{V}} < \infty$ and $B^*(\mathcal{V}_0) \subseteq \mathcal{V}_0$, then for every $T > 0$ and $\tilde{u} \in H$ such that $\|\tilde{u}\|_V \leq 1$ there exists a constant $\tilde{C}_T$ such that for all $d \in \mathbb{N}$ we have

$$\sup_{t\in[0,T]} \left|\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle Y_t, \tilde{u}\rangle_H}\right] - \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle Y_t^d, \tilde{u}\rangle_H}\right]\right| \leq \tilde{C}_T \|\mathbf{P}_d^\perp\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}(1 + \|x\|). \tag{6.22}$$

Remark 6.8. If, e.g., the ONB $(e_n)_{n\in\mathbb{N}}$ consists of eigenvectors of $\mathcal{A}$, then we could also approximate the process Y^d by H_d-valued processes $(\tilde{Y}^d)_{d\in\mathbb{N}}$ solving (6.20) with $\mathcal{A}$ replaced by $\mathcal{A}^d := \mathsf{P}_d \mathcal{A} \mathsf{P}_d$ and X by X^d. In this case it can be seen that an analog of Proposition 6.7 holds true with (Y^d, X^d) replaced by $(\tilde{Y}^d, X^d)$. However, in applications we usually have $\mathcal{A} = \frac{\partial}{\partial x}$, see e.g. Section 4.3, which does not admit an ONB of eigenvectors on, e.g., the space H_β.

6.5.2 Example: Finite-rank approximation of geometric affine forward curve models

In this section we discuss finite-rank approximation of geometric affine stochastic covariance models for forward curve dynamics, as introduced in Section 4.3.

Let $H = L^2(\mathbb{R}^+, \mathrm{e}^{\gamma x}\,\mathrm{d}x) \oplus \mathbb{R}$, for some $\gamma > 0$, $V = H_\beta$ as in Section 4.3 with $0 < \gamma < \beta$ and let $\mathcal{V}$ and $\mathcal{H}$ be as in Section 6.3.3 accordingly. It then follows from [138, Theorem 6], that $H_\beta \subset\subset L^2(\mathbb{R}^+, \mathrm{e}^{\gamma x}\,\mathrm{d}x) \oplus \mathbb{R}$ whenever $0 < \gamma < \beta$, which means that Assumption $\mathcal{E}$ is satisfied and we have $\mathcal{V} \subset\subset \mathcal{H}$.

Assume moreover that (Y, X) is given by (4.1) on $H = L^2(\mathbb{R}^+, \mathrm{e}^{\gamma x}\,\mathrm{d}x) \oplus \mathbb{R}$ with X a càdlàg affine process on $\mathcal{H}^+$ associated with (b, B, m, μ) satisfying Assumption $\mathcal{C}$ and where $(y, x) \in V \times \mathcal{H}^+$, $\mathcal{A} = \frac{\partial}{\partial x}$, $D \in \mathcal{L}_2(H, V)$, W is a H-cylindrical Brownian motion and $\Upsilon \in H$ satisfies (4.23). Note that the first-derivative operator $\mathcal{A}$ generates a strongly continuous semigroup on H_β, hence for this parameter choice the conditions of Proposition 6.7 are satisfied, up to the inclusion of the additional drift term $D^{1/2}X_tD^{1/2}\Upsilon$ in (4.1a) compared to the dynamics (6.20). This does not affect the weak convergence in Proposition 6.7 ii), but only the exact form of (6.21) and the rates in (6.22), see Section 4.2.2.
We model the *arbitrage-free* forward price under the HJMM framework as in equation (4.19) for Y as specified as above. Note that in contrast to the geometric affine model in Section 4.3, the cylindrical Brownian motion W takes values in the larger space $L^2(\mathbb{R}^+, \mathrm{e}^{\gamma x}\,\mathrm{d}x) \oplus \mathbb{R}$. However, note that by assumption $D \in \mathcal{L}_2(H, V)$, $y \in V$ and moreover $(S(t))_{t\geq 0}$, the shift-semigroup generated by $\mathcal{A}$, is strongly continuous on V. From this, it follows that the logarithmic forward curve dynamics Y assume values in the space $V = H_\beta$ again. Note that the weighted Lebesgue space $L^2(\mathbb{R}^+, \mathrm{e}^{\gamma x}\,\mathrm{d}x) \oplus \mathbb{R}$ was also used as a state-space for forward curve dynamics in, e.g., [130, 13].
It follows from Proposition 6.7 and Section 4.3 that the finite-rank stochastic covariance models $(Y^d, X^d)_{d\in\mathbb{N}}$ converge weakly to the (infinite-rank) affine stochastic covariance model (Y, X) as $d \to \infty$. Since the exponential and the evaluation maps are both continuous functions on V, we conclude that also the forward curve dynamics $(f^d)_{d\in\mathbb{N}}$, for every $d \in \mathbb{N}$ given by

$$f_t^d(x) \coloneqq \exp\left(\delta_x(Y_t^d)\right) = \exp\left(\langle Y_t^d, u_x\rangle_\beta\right), \quad t \geq 0,\ x \geq 0, \tag{6.23}$$

converge weakly. More precisely, we have the following result.

Corollary 6.9. *Assume the setting above. Then the sequence of finite-rank forward price curve dynamics $(f^d)_{d\in\mathbb{N}}$ given by (6.23) converges weakly to the forward curve dynamics $(f_t)_{t\geq 0}$ in (4.19) for $(Y_t)_{t\geq 0}$ specified as above, i.e. for all functions $g \in C\big(C(\mathbb{R}^+, H_\beta), \mathbb{R}\big)$ we have $\mathbb{E}\left[g(f^d)\right] \to \mathbb{E}\left[g(f)\right]$ as $d \to \infty$.*

Remark 6.10. Since $H_\beta \subset\subset L^2(\mathbb{R}^+, \mathrm{e}^{\gamma x}\,\mathrm{d}x) \oplus \mathbb{R}$ we could, in principle, approximate the forward curve dynamics $(f_t)_{t\geq 0}$ as in [138] by a finite-dimensional process $(U_d(f_t))_{t\geq 0}$ in the space $L^2(\mathbb{R}^+, \mathrm{e}^{\gamma x}\,\mathrm{d}x) \oplus \mathbb{R}$, where $(U_d)_{d\in\mathbb{N}}$, for $d \in \mathbb{N}$, is defined as the finite-rank operator $U_d v = \sum_{k=1}^d s_k \langle v, h_k^{(d)}\rangle_V e_k$ for some $(s_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}^+$, $(h_k^{(d)})_{k\leq d} \subseteq \operatorname{dom}(\mathcal{A}^*)$. However, this approach would lead to dynamics of $(f_t)_{t\geq 0}$ that are not exponential affine, since projections of affine processes are in general not affine. Thus, $(U_d(f_t))_{t\geq 0}$, although living in a finite-dimensional subspace, would loose its tractability. Instead, we propose to approximate the forward curve dynamics (4.19) by *geometric finite-rank affine stochastic covariance models* $(f^d)_{d\in\mathbb{N}}$, which for every $d \in \mathbb{N}$, are affine and have finite-rank noise.

6.6 Proof: Galerkin approximations of generalized Riccati equations

This section is devoted to the proof of Proposition 6.1. If Assumption $\mathcal{E}$ is satisfied, then Corollary 6.12 below makes the convergence rate (6.3) more specific. We begin this section with a short lemma on the local Lipschitz continuity of the functions F, R, $(F_d)_{d\in\mathbb{N}}$ and $(R_d)_{d\in\mathbb{N}}$.

Lemma 6.11. *Let $M>0$ and $d\in\mathbb{N}$. Then for every $u,v\in\mathcal{H}_d^+$ with $\|u\|\vee\|v\|\leq M$ we have*

$$|F_d(u)-F_d(v)|\leq\Big(\|b\|+(M+1)\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\,m(\mathrm{d}\xi)\Big)\|u-v\|, \tag{6.24}$$

$$\|R_d(u)-R_d(v)\|\leq\Big(\|B\|_{\mathcal{L}(\mathcal{H})}+(M+1)\|\mu(\mathcal{H}^+\setminus\{0\})\|\Big)\|u-v\|. \tag{6.25}$$

Moreover, for every $u,v\in\mathcal{H}^+$ with $\|u\|\vee\|v\|\leq M$ we can replace F_d by F and R_d by R, respectively, and the inequalities (6.24) *and* (6.25) *continue to hold with the same local Lipschitz constants.*

Proof. We prove the inequalities for F and R first. Let $M>0$ and $u,v\in\mathcal{H}^+$ such that $\|u\|\vee\|v\|\leq M$ and note that for all $\xi\in\mathcal{H}^+$ we have

$$|\mathrm{e}^{-\langle\xi,u\rangle}-\mathrm{e}^{-\langle\xi,v\rangle}+\langle\xi,u-v\rangle|\leq\|\xi\|^2(\|u\|\vee\|v\|)\|u-v\|,$$

and $|\,\mathrm{e}^{-\langle\xi,u\rangle}-\mathrm{e}^{-\langle\xi,v\rangle}|\leq|\langle\xi,u-v\rangle|$, see also Remark 2.13, it thus follows that

$$\begin{aligned}|F(u)-F(v)|&\leq|\langle b,u-v\rangle|+\int_{\mathcal{H}^+\setminus\{0\}}|\,\mathrm{e}^{-\langle\xi,u\rangle}-\mathrm{e}^{-\langle\xi,u\rangle}+\langle\chi(\xi),u-v\rangle|\,m(\mathrm{d}\xi)\\&\leq\|b\|\|u-v\|+M\Big(\int_{\mathcal{H}^+\cap\{0<\|\xi\|\leq1\}}\|\xi\|^2\,m(\mathrm{d}\xi)\Big)\|u-v\|\\&\quad+\Big(\int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\|\xi\|\,m(\mathrm{d}\xi)\Big)\|u-v\|\\&\leq\Big(\|b\|+(M+1)\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\,m(\mathrm{d}\xi)\Big)\|u-v\|,\end{aligned} \tag{6.26}$$

which proves inequality (6.24) for F_d replaced with F. For every $d\in\mathbb{N}$ it is then obvious that also the function F_d is Lipschitz continuous on the set $\mathcal{H}_d^+\cap\{u\in\mathcal{H}^+\colon\|u\|\leq M\}$ with the same Lipschitz constant that F in (6.26) admits.

For the second inequality (6.25), again at first for R replacing R_d, we note that by the monotonicity of the cone $\mathcal{H}^+$ we have

$$\begin{aligned}\|R(u)-R(v)\| &\leq \|B^*\|_{\mathcal{L}(\mathcal{H})}\|u-v\| + \Big\| \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}} \big|\, \mathrm{e}^{-\langle \xi,u\rangle} - \mathrm{e}^{-\langle \xi,u\rangle} \big| \, \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \Big\| \\ &\quad + \Big\| \int_{\mathcal{H}^+\cap\{0<\|\xi\|\leq 1\}} \big|\, \mathrm{e}^{-\langle \xi,u\rangle} - \mathrm{e}^{-\langle \xi,u\rangle} + \langle \xi, u-v\rangle \big| \, \frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} \Big\| \\ &\leq \big(\|B\|_{\mathcal{L}(\mathcal{H})} + (M+1)\|\mu(\mathcal{H}^+\setminus\{0\})\|\big)\, \|u-v\|. \qquad (6.27)\end{aligned}$$

We see that inequality (6.27) also holds for every $(R_d)_{d\in\mathbb{N}}$ with the same local Lipschitz constant on $\mathcal{H}_d^+ \cap \{u\in\mathcal{H}^+ \colon \|u\|\leq M\}$ for all $d\in\mathbb{N}$, given by (6.27). In particular, we find Lipschitz constants for R_d and F_d on $\big\{u\in\mathcal{H}_d^+\colon \|u\|\leq M\big\}$ that hold uniformly for all $d\in\mathbb{N}$. □

With this lemma at hand we can now prove Proposition 6.1.

Proof of Proposition 6.1. Let $u\in\mathcal{H}^+$, $T>0$ and $d\in\mathbb{N}$. We begin with showing the existence and uniqueness of the solution of (6.1) on the interval $[0,T]$. From (2.4) it follows that for every $x,v\in\mathcal{H}_d^+$ such that $\langle x,v\rangle=0$ we have

$$\langle R_d(v),x\rangle = \langle \mathbf{P}_d(R(v)),x\rangle = \langle R(v),x\rangle \geq 0,$$

which implies that R_d is *quasi-monotone increasing* with respect to the cone $\mathcal{H}_d^+$, see also Definition 2.10. This and the Lipschitz continuity of R_d on the sets $\big\{v\in\mathcal{H}_d^+\colon \|v\|\leq M\big\}$, for every $M>0$, see Lemma 6.11, implies the existence and uniqueness of a continuously differentiable function $\psi_d(\cdot,\mathbf{P}_d(u))$ on $[0,T]$ that solves (6.1b), see also the proof of Proposition 2.16. The existence and uniqueness of a continuously differentiable function $\phi_d(\cdot,\mathbf{P}_d(u))$ on $[0,T]$ solving (6.1a) then follows immediately from the continuity of F_d and mere integration of both sides of (6.1a).

Next, we prove the inequality (6.2). For this let us fix $M>0$ and note that by Lemma 6.11 we find a Lipschitz constant of R_d and R on $\big\{v\in\mathcal{H}_d^+\colon \|v\|\leq M\big\}$ which does not depend on $d\in\mathbb{N}$. It thus follows from (2.21) (see also the proof of Proposition 2.16) that for all $t\in[0,T]$ and all $u\in\{u\in\mathcal{H}^+\colon \|u\|\leq M\}$ we have

$$\|\psi(t,u)\| \vee \|\psi_d(t,\mathbf{P}_d(u))\| \leq M\exp\big((\|B\|_{\mathcal{L}(\mathcal{H})} + 2\|\mu(\mathcal{H}^+\setminus\{0\})\|)T\big).$$

Let us set $H_M \coloneqq M\exp\big((\|B\|_{\mathcal{L}(\mathcal{H})} + 2\|\mu(\mathcal{H}^+\setminus\{0\})\|)T\big)$ and note that for every $t\in[0,T]$ and $u\in\mathcal{H}^+$ we have

$$\|\psi(t,u)-\psi_d(t,\mathbf{P}_d(u))\| \leq \|\psi(t,u)-\mathbf{P}_d(\psi(t,u))\| + \|\mathbf{P}_d(\psi(t,u))-\psi_d(t,\mathbf{P}_d(u))\|.$$

From (6.25) we observe that the second term, for all $u \in \mathcal{H}^+$ with $\|u\| \leq M$, satisfies

$$\begin{aligned}\|\mathbf{P}_d(\psi(t,u)) - \psi_d(t,\mathbf{P}_d(u))\| &\leq \int_0^t \|\mathbf{P}_d R(\psi(s,u)) - R_d(\psi_d(s,\mathbf{P}_d(u)))\| \,\mathrm{d}s \\ &\leq L_M^{(1)} \int_0^t \|\psi(s,u) - \psi_d(s,\mathbf{P}_d(u))\| \,\mathrm{d}s, \qquad (6.28)\end{aligned}$$

with $L_M^{(1)} \coloneqq \|B\|_{\mathcal{L}(\mathcal{H})} + (H_M + 1)\|\mu(\mathcal{H}^+ \setminus \{0\})\|$. Moreover, for all $u, \xi \in \mathcal{H}$ we set $K_u(\xi) \coloneqq \mathrm{e}^{-\langle \xi, u\rangle} - 1 + \langle \chi(\xi), u\rangle$ and recall that by the variation-of-constant formula the solution $\psi(\cdot, u)$ satisfies

$$\psi(t,u) = \mathrm{e}^{tB^*}u + \int_0^t \mathrm{e}^{(t-s)B^*}\Big(\int_{\mathcal{H}^+\setminus\{0\}} K_{\psi(s,u)}(\xi)\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big)\,\mathrm{d}s, \quad t \in [0,T].$$

From this and writing $\|\psi(t,u) - \mathbf{P}_d(\psi(t,u))\| = \|\mathbf{P}_d^{\perp}(\psi(t,u))\|$ we obtain

$$\begin{aligned}\|\mathbf{P}_d^{\perp}(\psi(t,u))\| &\leq \|\mathbf{P}_d^{\perp}(\mathrm{e}^{tB^*}u)\| + \int_0^t \Big\|\mathbf{P}_d^{\perp}\,\mathrm{e}^{(t-s)B^*}\int_{\mathcal{H}^+\setminus\{0\}} K_{\psi(s,u)}(\xi)\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big\|\,\mathrm{d}s \\ &\leq \|\mathbf{P}_d^{\perp}(\mathrm{e}^{tB^*}u)\| + tH_M^2 \sup_{s\in[0,t]}\|\mathbf{P}^{\perp}\,\mathrm{e}_d^{sB^*}(\mu(\mathcal{H}^+ \setminus \{0\}))\|, \qquad (6.29)\end{aligned}$$

where in the last line of (6.29) we used that

$$|\mathrm{e}^{-\langle \xi,\psi(s,u)\rangle} - 1 + \langle \chi(\xi), \psi(s,u)\rangle| \leq \frac{\|\xi\|^2}{2}\|\psi(s,u)\|^2\mathbf{1}_{\{\|\xi\|\leq 1\}} + \|\xi\|\|\psi(s,u)\|\mathbf{1}_{\{\|\xi\|>1\}},$$

$\sup_{s\in[0,t]}\|\psi(s,u)\| \leq H_M$ for all $t \in [0,T]$ and the monotonicity of the integral, which implies that for every $s \in [0,t]$ we have

$$\Big\|\int_{\mathcal{H}^+\setminus\{0\}} K_{\psi(s,u)}(\xi)\frac{\mathbf{P}_d^{\perp}\,\mathrm{e}^{(t-s)B^*}\mu(\mathrm{d}\xi)}{\|\xi\|^2}\Big\| \leq H_M^2\|\mathbf{P}_d^{\perp}\,\mathrm{e}^{(t-s)B^*}\mu(\mathcal{H}^+ \setminus \{0\})\|.$$

Let us denote the right-hand side of (6.29) by K_t^d, then for every $t \in [0,T]$ we see from (6.28) and (6.29) that

$$\|\psi(t,u) - \psi_d(t,\mathbf{P}_d(u))\| \leq K_t^d + L_M^{(1)}\int_0^t \|\psi(s,u) - \psi_d(s,\mathbf{P}_d(u))\|\,\mathrm{d}s.$$

This, the fact that K_t^d is non-decreasing in t and an application of Gronwall's inequality yields

$$\|\psi(t,u) - \psi_d(t,\mathbf{P}_d(u))\| \leq K_t^d \exp(L_M^{(1)}t), \quad t \in [0,T], \qquad (6.30)$$

and we note that $\sup_{t\in[0,T]} K_{t,d} = K_T^d$.

Hence, taking the supremum over all $t \in [0,T]$ on both sides of (6.30) yields

$$\sup_{t\in[0,T]} \|\psi(t,u) - \psi_d(t,\mathbf{P}_d(u))\| \leq K_T^d \exp(L_M^{(1)}T).$$

Similarly, for the error term in (6.2) involving $\phi(\cdot,u)$, we note that $F_d(u) = F(u)$ for all $u \in \mathcal{H}_d^+$ and hence by using (6.24) we obtain

$$\begin{aligned}|\phi(t,u) - \phi_d(t,\mathbf{P}_d(u))| &\leq \int_0^t |F(\psi(s,u)) - F_d(\psi_d(s,\mathbf{P}_d(u)))| \,\mathrm{d}s \\ &\leq \Big(\|b\| + (H_M+1)\int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^2 \, m(\mathrm{d}\xi)\Big)\int_0^t K_{s,d}\, \mathrm{e}^{L_M^{(1)}s}\,\mathrm{d}s \\ &\leq L_M^{(2)} t \sup_{s\in[0,t]} K_s^d \exp(L_M^{(1)}s),\end{aligned}$$

with $L_M^{(2)} := \|b\| + (H_M+1)\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\, m(\mathrm{d}\xi)$. Moreover, we conclude that the left-hand side in (6.2) is bounded by $\mathrm{e}^{L_M^{(1)}T}\big(1+L_M^{(2)}T\big)K_T^d$. We note that $L_M^{(1)}$ and $L_M^{(2)}$ are independent of $d \in \mathbb{N}$ and thus setting $K := \mathrm{e}^{L_M^{(1)}T}\big(1+L_M^{(2)}T\big)(1+TH_M^2)$ yields (6.2) with $C_{T,d}$ given by (6.3).
Let us prove that $C_{T,d}$ vanishes when d tends to infinity. Indeed, note first that the map $t \mapsto \mathrm{e}^{tB^*}$ is continuous and thus maps compact sets to compact sets. In particular, for every $v \in \mathcal{H}^+$ we see that the set $\{\mathrm{e}^{tB^*}v \colon t \in [0,T]\}$ is compact in $\mathcal{H}^+$ and since for every $d \in \mathbb{N}$ the operators $\mathbf{P}_d^\perp$ converge uniformly on compact sets, we conclude that $\sup_{t\in[0,T]}\|\mathbf{P}_d^\perp(\mathrm{e}^{tB^*}v)\| \to 0$ as $d \to \infty$. Applying this to $v = u$ and $v = \mu(\mathcal{H}^+ \setminus \{0\})$ accordingly, implies that the left-hand side in (6.2) converges to zero uniformly on compact sets in time as d tends to infinity. □

We end this section with a corollary of Proposition 6.1 providing more specific convergence rates under Assumption $\mathcal{E}$. This convergence rate appears again in (6.9). Let V, H and $\mathcal{V}$ be as in Section 6.3.3, then the following corollary holds true:

Corollary 6.12. *Let the assumptions of Proposition 6.1 hold and assume in addition that Assumption $\mathcal{E}$ is satisfied. If moreover $\|\mu(\mathcal{H}^+ \setminus \{0\})\|_{\mathcal{V}} < \infty$ and $B^*(\mathcal{V}_0) \subseteq \mathcal{V}_0$, then for all $T > 0$ we have*

$$\sup_{t\in[0,T],\, \|u\|_{\mathcal{V}}\leq 1} \big(|\phi_d(t,\mathbf{P}_d(u)) - \phi(t,u)| + \|\psi_d(t,\mathbf{P}_d(u)) - \psi(t,u)\|\big) \leq C_T \|\mathbf{P}_d^\perp\|_{\mathcal{L}(\mathcal{V},\mathcal{H})},$$

with $C_T = \big(\mathrm{e}^{L_C^{(1)}}(1+L_C^{(2)}T)(1+TH_C^2)\big)\big(\mathrm{e}^{T\|B\|_{\mathcal{L}(\mathcal{H})}}(1+\|\mu(\mathcal{H}^+\setminus\{0\})\|_{\mathcal{V}})\big)$ for H_C, $L_C^{(1)}$ and $L_C^{(2)}$ being as in the proof of Proposition 6.1 with $M = C$.

Proof. Let K and $C_{T,d}$ be as in Proposition 6.1 and note that $\sup_{\|u\|_{\mathcal{V}}\leq 1} KC_{T,d} \leq C_T \|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}$. Moreover, since $B^*(\mathcal{V}_0) \subseteq \mathcal{V}_0$, we have $\mathrm{e}^{tB^*}(\mathcal{V}_0) \subseteq \mathcal{V}_0$ for all $t \geq 0$, which together with $\|\mu(\mathcal{H}^+ \setminus \{0\})\|_{\mathcal{V}} < \infty$ implies

$$\|\mathbf{P}_d^{\perp}\, \mathrm{e}^{tB^*} \mu(\mathcal{H}^+ \setminus \{0\})\| \leq \|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}\, \mathrm{e}^{t\|B\|_{\mathcal{L}(\mathcal{H})}} \|\mu(\mathcal{H}^+ \setminus \{0\})\|_{\mathcal{V}}.$$

Similarly, for every $u \in \mathcal{H}^+$ with $\|u\|_{\mathcal{V}} \leq 1$ we see that

$$\begin{aligned}\sup_{t\in[0,T]} \|\mathbf{P}_d^{\perp}\, \mathrm{e}^{tB^*} u\| \leq \sup_{t\in[0,T]} \|\mathbf{P}_d^{\perp}\, \mathrm{e}^{tB^*}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})} &\leq \|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})} \sup_{t\in[0,T]} \|\mathrm{e}^{tB^*}\|_{\mathcal{L}(\mathcal{H})} \\ &\leq \|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}\, \mathrm{e}^{T\|B\|_{\mathcal{L}(\mathcal{H})}}.\end{aligned}$$

Note further that $\|u\| \leq C\|u\|_{\mathcal{V}} \leq C$ and hence from the proof of Proposition 6.1 we see that for $M = C$ we can take $K = \mathrm{e}^{L_C^{(1)}}(1 + L_C^{(2)}T)(1 + TH_C)$ where H_C is such that $\|\psi_d(t,u)|\| \vee \|\psi(t,u)|\| \leq H_C$ for all $t \in [0,T]$ and $\|u\| \leq C$, $L_C^{(1)} = \|B\|_{\mathcal{L}(\mathcal{H})} + 2(H_C+1)\|\mu(\mathcal{H}^+ \setminus \{0\})\|$ and $L_C^{(2)} = \|b\| + 2(H_C+1)\int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^2\, m(\mathrm{d}\xi)$. □

6.7 Affine finite-rank operator-valued processes

In this section we construct a sequence of finite-rank operator-valued affine processes associated with $(\phi_d(\cdot, \mathbf{P}_d(u)), (\psi_d(\cdot, \mathbf{P}_d(u)))_{d\in\mathbb{N}}$, the sequence of Galerkin approximations given by Proposition 6.1. The existence of this sequence of processes is asserted in Proposition 6.2. First, in Section 6.7.1 we project the given admissible parameter set onto spaces of finite-rank operators and prove that for every rank $d \in \mathbb{N}$, the projected sets can be identified with *matrix-valued admissible parameter sets* as in [42, Definition 2.3]. In Section 6.7.2 we derive from this the existence of an associated sequence of affine processes with values in positive semi-definite matrices. Subsequently, in Section 6.7.3 we transform this sequence back into the space of self-adjoint Hilbert-Schmidt operators and prove that this transformed sequence satisfies the asserted properties in Proposition 6.2.

6.7.1 Finite-rank admissible parameters

Assume that we are in the setting of Section 6.3. In particular, let (b, B, m, μ) be an admissible parameter set as in Definition 2.3 and let $(\mathcal{H}_d, \mathbf{P}_d)_{d\in\mathbb{N}}$ be a finite-rank projection scheme in $\mathcal{H}$ (with respect to the orthonormal basis $(\mathbf{e}_{i,j})_{i\leq j\in\mathbb{N}}$ in $\mathcal{H}$). For any two measurable spaces $(\mathcal{E}_1, \mathcal{B}_1)$ and $(\mathcal{E}_2, \mathcal{B}_2)$ and measurable function $f\colon (\mathcal{E}_1, \mathcal{B}_1) \to (\mathcal{E}_2, \mathcal{B}_2)$, we denote the push-forward of a measure $\mu_1\colon \mathcal{E}_1 \to [0,\infty]$ with respect to f by $f_*\mu_1$, i.e. $f_*\mu_1(A) = \mu(f^{-1}(A))$ for any $A \in \mathcal{B}_2$ and note that $f_*\mu_1$ is a proper measure on $(\mathcal{E}_2, \mathcal{B}_2)$.

For every $d \in \mathbb{N}$, we define the sets $E_d \coloneqq \{\xi \in \mathcal{H}^+ \colon\ 0 < \|\mathbf{P}_d(\xi)\| \leq 1,\ \|\xi\| > 1\}$ and $E_d^0 \coloneqq \{\xi \in \mathcal{H}^+ \colon \mathbf{P}_d(\xi) \neq 0\}$. Then we introduce the following notion:

Definition 6.13. For every $d \in \mathbb{N}$ we define the parameters (b_d, B_d, m_d, μ_d) and M_d as follows:

i) The measure $m_d \colon \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\}) \to [0, \infty]$ is defined as the push-forward of m with respect to $\mathbf{P}_d$, i.e.

$$m_d(\mathrm{d}\xi) \coloneqq (\mathbf{P}_{d*}m)(\mathrm{d}\xi).$$

ii) The vector $b_d \in \mathcal{H}_d$ is given by

$$b_d \coloneqq \mathbf{P}_d(b) + \int_{E_d} \mathbf{P}_d(\xi)\, m(\mathrm{d}\xi). \tag{6.31}$$

iii) The $\mathcal{H}_d^+$-valued measure $\mu_d \colon \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\}) \to \mathcal{H}_d^+$ is defined as the $\mathbf{P}_d$-projection of the push-forward of μ with respect to $\mathbf{P}_d$, i.e.

$$\mu_d(\mathrm{d}\xi) \coloneqq \mathbf{P}_d(\mathbf{P}_{d*}\mu)(\mathrm{d}\xi).$$

Moreover, we define the $\mathcal{H}_d^+$-valued measure M_d on $\mathcal{H}_d^+ \setminus \{0\}$ as follows: For every $A \in \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\})$ we set

$$M_d(A) \coloneqq \int_{\mathcal{H}_d^+ \setminus \{0\}} \mathbf{1}_A(\mathbf{P}_d(\xi)) \frac{1}{\|\xi\|^2} \mathbf{P}_d(\mu(\mathrm{d}\xi)). \tag{6.32}$$

iv) The linear operator $B_d \colon \mathcal{H}_d \to \mathcal{H}_d$ is defined by

$$B_d(u) \coloneqq \mathbf{P}_d(B(u)) + \int_{E_d} \mathbf{P}_d(\xi)\, \langle u, M(\mathrm{d}\xi) \rangle, \quad u \in \mathcal{H}_d. \tag{6.33}$$

Remark 6.14. Note that by the definition of M_d in (6.32) we have

$$\int_{E_d} \mathbf{P}_d(\xi)\, \langle u, M_d(\mathrm{d}\xi) \rangle = \int_{E_d} \frac{\mathbf{P}_d(\xi)}{\|\xi\|^2}\, \langle u, \mathbf{P}_d(\mu(\mathrm{d}\xi)) \rangle, \quad u \in \mathcal{H}_d, \tag{6.34}$$

and for all $u \in \mathcal{H}_d$ we have

$$\begin{aligned}\Big\| \int_{E_d} \frac{\mathbf{P}_d(\xi)}{\|\xi\|^2} \langle u, \mathbf{P}_d(\mu(\mathrm{d}\xi)) \rangle \Big\| &\leq \int_{\mathcal{H}^+ \cap \{\|\xi\| > 1\}} \|\xi\|^{-1} \langle u, \mu(\mathrm{d}\xi) \rangle \\ &\leq \|u\| \|\mu(\mathcal{H}^+ \cap \{\|\xi\| > 1\})\| < \infty.\end{aligned}$$

This shows that the integral in (6.33) is well defined (in a Bochner sense) and uniformly norm-bounded in $d \in \mathbb{N}$. Similarly, for all $d \in \mathbb{N}$ the integral part in (6.31) satisfies

$$\int_{E_d} \mathbf{P}_d(\xi)\, m(\mathrm{d}\xi) \leq \int_{\mathcal{H}^+ \cap \{\|\xi\|>1\}} \xi\, m(\mathrm{d}\xi) \in \mathcal{H}^+.$$

In the following two lemmas we give some observations that we will use in the next section.

Lemma 6.15. *Let $M(x, \mathrm{d}\xi)\colon \mathcal{B}(\mathcal{H}^+ \setminus \{0\}) \to [0,\infty]$ be the kernel defined in* (2.3) *and let μ_d and M_d be as in Definition* 6.13 iii)*. Moreover, for every $x \in \mathcal{H}_d^+$ we define the measure $M_d(x, \mathrm{d}\xi)\colon \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\}) \to [0,\infty]$ by*

$$M_d(x, A) := \int_{\mathcal{H}^+ \setminus \{0\}} \mathbf{1}_A(\mathbf{P}_d(\xi)) \frac{1}{\|\xi\|^2} \langle x, \mathbf{P}_d(\mu(\mathrm{d}\xi))\rangle, \quad A \in \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\}). \tag{6.35}$$

Then for every $A \in \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\})$ and $x \in \mathcal{H}_d^+$ we have $M_d(x, A) = \langle x, M_d(A)\rangle$ and

$$\langle x, \mu_d(A)\rangle = \int_{\mathcal{H}^+ \setminus \{0\}} \mathbf{1}_A(\mathbf{P}_d(\xi)) \|\xi\|^2 M(x, \mathrm{d}\xi). \tag{6.36}$$

Proof. Recall that $M(x, \mathrm{d}\xi) = \|\xi\|^{-2}\langle x, \mu(\mathrm{d}\xi)\rangle$ is a measure on $\mathcal{B}(\mathcal{H}^+ \setminus \{0\})$. Now let $x \in \mathcal{H}_d^+$ and $A \in \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\})$, then by definition we have $\langle x, \mu_d(A)\rangle = \langle x, (\mathbf{P}_{d*}\mu)(A)\rangle$ and we obtain by the variation-of-constant formula for pushforward measures that

$$\begin{aligned} \langle x, (\mathbf{P}_{d*}\mu)(A)\rangle &= \int_{\mathcal{H}^+ \setminus \{0\}} \mathbf{1}_A(\mathbf{P}_d(\xi)) \langle x, \mu(\mathrm{d}\xi)\rangle \\ &= \int_{\mathcal{H}^+ \setminus \{0\}} \mathbf{1}_{\mathbf{P}_d}(\mathbf{P}_d(\xi)) \|\xi\|^2 \langle x, M(\mathrm{d}\xi)\rangle. \end{aligned}$$

Note that $\mathbf{1}_A(\mathbf{P}_d(\xi))$ vanishes on the set $\{\xi \in \mathcal{H}^+\colon \mathbf{P}_d(\xi) = 0\}$ since $\mathbf{1}_A(\mathbf{P}_d(\xi)) = \mathbf{1}_A(0) = 0$ as $0 \in A^c$ for every $A \subseteq \mathcal{H}_d^+ \setminus \{0\}$, i.e. we technically integrate over E_d^0 only and $\mathbf{P}_d\colon (E_d^0, \mathcal{B}(E_d^0)) \to (\mathcal{H}_d^+ \setminus \{0\}, \mathcal{B}(\mathcal{H}_d^+ \setminus \{0\}))$ is measurable. ☐

In the next lemma we show that the function F_d and R_d from Section 6.3.1 can be expressed in terms of the parameters b_d, B_d, m_d and M_d from Definition 6.13. This will help us to associate the Galerkin approximations to affine processes in the subsequent section.

Lemma 6.16. *For every $u \in \mathcal{H}_d^+$ we can express $F_d(u)$ and $R_d(u)$ by means of the parameters b_d, B_d, m_d and M_d as follows:*

$$F_d(u) = \langle b_d, u\rangle - \int_{\mathcal{H}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, u\rangle} - 1 + \langle \chi(\xi), u\rangle\right) m_d(\mathrm{d}\xi), \tag{6.37a}$$

$$R_d(u) = B_d^*(u) - \int_{\mathcal{H}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, u\rangle} - 1 + \langle \chi(\xi), u\rangle\right) M_d(\mathrm{d}\xi). \tag{6.37b}$$

Proof. We only proof the identity (6.37b) as the proof of (6.37a) is similar. Let $u \in \mathcal{H}_d^+$ and note that we have $\langle \xi, u\rangle = \langle \mathbf{P}_d(\xi), u\rangle$ for every $\xi \in \mathcal{H}^+$ since $\mathbf{P}_d$ is an orthogonal projection with respect to $\langle\cdot,\cdot\rangle$. Setting $M(\mathrm{d}\xi) := \|\xi\|^{-2}\mu(\mathrm{d}\xi)$ we see from the definition of R_d, Definition 6.13 iv) and (6.34) that

$$\begin{aligned} R_d(u) &= \mathbf{P}_d(B(u)) - \int_{\mathcal{H}^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \mathbf{P}_d(\xi), u\rangle} - 1 + \langle \mathbf{P}_d(\chi(\xi)), u\rangle\right) \frac{\mathbf{P}_d(\mu(\mathrm{d}\xi))}{\|\xi\|^2} \\ &= B_d(u) - \int_{\mathcal{H}^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \mathbf{P}_d(\xi), u\rangle} - 1 + \langle \chi(\mathbf{P}_d(\xi)), u\rangle\right) \mathbf{P}_d(M(\mathrm{d}\xi)). \end{aligned} \tag{6.38}$$

Note that on the set $\{\xi \in \mathcal{H}^+ \colon \mathbf{P}_d(\xi) = 0\}$ the integrand on the right-hand side of (6.38) vanishes and hence we see that the integral coincides with

$$\int_{E_d^0} \left(\mathrm{e}^{-\langle \mathbf{P}_d(\xi), u\rangle} - 1 + \langle \chi(\mathbf{P}_d(\xi)), u\rangle\right) \mathbf{P}_d(M(\mathrm{d}\xi)). \tag{6.39}$$

From the definition of M_d in (6.32) and by the change-of-variables formula for push-forward measures, we conclude that the integral in (6.39) is equal to

$$\int_{\mathcal{H}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, u\rangle} - 1 + \langle \chi(\xi), u\rangle\right) M_d(\mathrm{d}\xi),$$

which inserted back into (6.38) proves the identity (6.37b). □

6.7.2 Identification with matrix-valued affine processes

For every $d \in \mathbb{N}$ we denote by $(\mathbb{M}_d, \langle\cdot,\cdot\rangle_d)$ the space of all real $d \times d$-matrices equipped with the trace inner-product $\langle x, y\rangle_d := \mathrm{Tr}(y^\intercal x)$ for $x, y \in \mathbb{M}_d$, where $y^\intercal \in \mathbb{M}_d$ denotes the transpose of y. The norm $\|\cdot\|_d$ induced by $\langle\cdot,\cdot\rangle_d$ is called the *Frobenius norm*, which is nothing else than the Hilbert-Schmidt norm in the case of $H = \mathbb{R}_d$. As before, we denote the subspace of $\mathbb{M}_d$ consisting of all the symmetric $d\times d$-matrices by $\mathbb{S}_d$. For $d \in \mathbb{N}$ we denote by $\{v_1, \ldots, v_d\}$ the standard basis of $\mathbb{R}_d$ and define the coordinate system $\Phi_d \colon \mathbb{R}_d \to H_d$ associated with the basis $\{e_1, \ldots, e_d\}$ of H_d by

$$\Phi_d(v_i) = e_i, \quad \text{for } i = 1, \ldots, d. \tag{6.40}$$

The coordinate system Φ_d identifies the d-dimensional subspace H_d with $\mathbb{R}_d$ and we can represent every linear operator $A \in \mathcal{L}(H_d)$ as a $d \times d$-matrix by using the mapping $i_d \colon \mathcal{L}(H_d) \to \mathbb{S}_d$ given by

$$i_d(A) := \Phi_d^{-1} \circ A \circ \Phi_d, \tag{6.41}$$

where under the usual matrix-identification we shall understand $i_d(A)$ as an element in $\mathbb{M}_d$. Note that whenever A is self-adjoint, its matrix representation $i_d(A)$ is self-adjoint as well, which can be seen by taking $x, y \in \mathbb{R}_d$ and the brief computation

$$\langle i_d(A)x, y\rangle_{\mathbb{R}^d} = \langle A \circ \Phi_d(x), \Phi_d(y)\rangle_H = \langle \Phi_d(x), A^* \circ \Phi_d(y)\rangle_H = \langle x, i_d(A)y\rangle_{\mathbb{R}^d}.$$

Under the mapping i_d in (6.41) we identify $\mathcal{H}_d \subseteq \mathcal{L}(H_d)$ with $\mathbb{S}_d$ and note that i_d is an isometry between $\mathbb{S}_d$ and $\mathcal{H}_d$, i.e. it identifies the Frobenius with the Hilbert-Schmidt norm. Moreover, we observe that for the cone of all symmetric positive semi-definite $d \times d$-matrices $\mathbb{S}_d^+$, the positivity of $\mathcal{H}_d^+$ is preserved under i_d, i.e. $i_d(\mathcal{H}_d^+) = \mathbb{S}_d^+$. In the following definition we introduce yet another transformation of the parameters b_d, B_d, m_d, μ_d and M_d from Definition 6.13. This time by identifying the spaces $\mathcal{H}_d$ and $\mathbb{S}_d$:

Definition 6.17. Let (b, B, m, μ) be an admissible parameter set as in Definition 2.3 and for $d \in \mathbb{N}$ let (b_d, B_d, m_d, μ_d) and M_d be as in Definition 6.13. For every $d \in \mathbb{N}$ we define the parameters $(\tilde{b}_d, \tilde{B}_d, \tilde{m}_d, \tilde{\mu}_d)$ and $\tilde{M}_d$ as follows:

i) The matrix $\tilde{b}_d \in \mathbb{S}_d^+$ is defined as $\tilde{b}_d := i_d(b_d)$.

ii) The linear operator $\tilde{B}_d \colon \mathbb{S}_d \to \mathbb{S}_d$ is given by $\tilde{B}_d := i_d \circ B_d \circ i_d^{-1}$.

iii) The measure $\tilde{m}_d \colon \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\}) \to [0, \infty]$ is defined as the push-forward of m_d with respect to i_d, i.e.

$$\tilde{m}_d(\mathrm{d}\xi) := (i_{d*} m_d)(\mathrm{d}\xi).$$

iv) The matrix-valued measure $\tilde{\mu}_d \colon \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\}) \to \mathbb{S}_d^+$ is defined as the composition of i_d and the push-forward of μ_d with respect to i_d, i.e.

$$\tilde{\mu}_d(\mathrm{d}\xi) = i_d((i_{d*}\mu_d)(\mathrm{d}\xi)).$$

Moreover, we define the $\mathbb{S}_d^+$-valued measure $\tilde{M}_d(\mathrm{d}\xi)$ as follows: For every $A \in \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\})$ we set

$$\tilde{M}_d(A) = \int_{\mathcal{H}^+ \setminus \{0\}} \mathbf{1}_A(i_d(\mathbf{P}_d(\xi))) \frac{1}{\|\xi\|^2} i_d(\mathbf{P}_d(\mu(\mathrm{d}\xi))),$$

and for every $x \in \mathbb{S}_d^+$ we write $\tilde{M}_d(x, \mathrm{d}\xi) := \langle x, \tilde{M}_d(\mathrm{d}\xi)\rangle$.

Now, let $\chi_d\colon \mathbb{S}_d \to \mathbb{S}_d$ be defined as $\chi_d(\xi) = \xi \mathbf{1}_{\|\xi\|_d \leq 1}(\xi)$. In the next lemma we show some crucial properties of the parameters $(\tilde{b}_d, \tilde{B}_d, \tilde{m}_d, \tilde{\mu}_d)$ and $\tilde{M}_d$ that allows us in a next step to identify a matrix-valued admissible parameter set with these parameters.

Lemma 6.18. *Let $d \in \mathbb{N}$ and $\tilde{b}_d$, $\tilde{B}_d$, $\tilde{m}_d$ and $\tilde{M}_d$ defined as in Definition 6.17. Then the following holds true:*

i) $\int_{\mathbb{S}_d^+\setminus\{0\}} \big(\|\xi\|_d \vee \|\xi\|_d^2\big)\tilde{m}_d(\mathrm{d}\xi) < \infty$.

ii) $\tilde{b}_d - \int_{\mathbb{S}_d^+\setminus\{0\}} \chi_d(\xi)\tilde{m}_d(\mathrm{d}\xi) \in \mathbb{S}_d^+$.

iii) For every $A \in \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\})$ we have $\tilde{M}_d(A) \in \mathbb{S}_d^+$ and

$$\int_{\mathbb{S}_d^+\setminus\{0\}} \langle \chi_d(\xi), u\rangle_d \tilde{M}_d(x, \mathrm{d}\xi) < \infty,$$

for all $x, u \in \mathbb{S}_d^+$ such that $\langle x, u\rangle_d = 0$.

iv) For all $x \in \mathbb{S}_d^+$ we have $\int_{\mathbb{S}_d^+\setminus\{0\}} \|\xi\|^2 \tilde{M}_d(x, \mathrm{d}\xi) < \infty$.

v) For all $x, u \in \mathbb{S}_d^+$ such that $\langle x, u\rangle_d = 0$ we have

$$\langle \tilde{B}_d(x), u\rangle_d - \int_{\mathcal{H}^+} \langle \chi_d(\xi), u\rangle_d \, \langle x, \tilde{M}_d(\mathrm{d}\xi)\rangle_d \geq 0. \tag{6.42}$$

Proof. First, note that for every $E \in \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\})$ we have

$$i_{d*}(\mathbf{P}_{d*}m)(E) = m(\mathbf{P}_d^{-1}(i_d^{-1}(E))) = m((i_d \circ \mathbf{P}_d)^{-1}(E)) = (i_d \circ \mathbf{P}_d)_* m(E),$$

and $\tilde{m}_d(\mathrm{d}\xi) = (i_d \circ \mathbf{P}_d)_* m(\mathrm{d}\xi)$ by definition and the analogous statement holds for the measure $\tilde{M}_d$. To show Lemma 6.18 i) we split up the integral as follows

$$\begin{aligned}\int_{\mathbb{S}_d^+\setminus\{0\}} \big(\|\xi\|_d \vee \|\xi\|_d^2\big)\, \tilde{m}_d(\mathrm{d}\xi) = &\int_{\{\xi\in\mathbb{S}_d^+\colon\, 0<\|\xi\|_d\leq 1\}} \|\xi\|_d \, \tilde{m}_d(\mathrm{d}\xi) \\ &+ \int_{\{\xi\in\mathbb{S}_d^+\colon\, \|\xi\|_d > 1\}} \|\xi\|_d^2 \, \tilde{m}_d(\mathrm{d}\xi).\end{aligned} \tag{6.43}$$

In the following we treat the two integrals on the right-hand side of (6.43) separately.

By the change-of-variable formula for pushforward measures and since i_d is an isometry, i.e. $\|\xi\| = \|i_d(\xi)\|_d$ for $\xi \in \mathcal{H}_d^+$, we deduce the following for the first integral in (6.43)

$$\begin{aligned}
\int_{\{\xi\in\mathbb{S}_d^+ : \, 0<\|\xi\|_d\leq 1\}} \|\xi\|_d \, \tilde{m}_d(\mathrm{d}\xi) &= \int_{\{\xi\in\mathcal{H}_d^+ : \, 0<\|i_d(\xi)\|_d\leq 1\}} \|i_d(\xi)\|_d \, m_d(\mathrm{d}\xi) \\
&= \int_{\{\xi\in\mathcal{H}_d^+ : \, 0<\|\xi\|\leq 1\}} \|\xi\| \, m_d(\mathrm{d}\xi) \\
&= \int_{\{\xi\in\mathcal{H}^+ : \, 0<\|\mathbf{P}_d(\xi)\|\leq 1\}} \|\mathbf{P}_d(\xi)\| \, m(\mathrm{d}\xi) \qquad (6.44) \\
&\leq \int_{\mathcal{H}^+\setminus\{0\}} \Big(\sum_{i=1}^{d}\sum_{j=i}^{d} \langle \xi, \mathbf{e}_{i,j}\rangle^2 \Big)^{\frac{1}{2}} m(\mathrm{d}\xi) \\
&\leq \sum_{i=1}^{d}\sum_{j=i}^{d} \int_{\mathcal{H}^+\setminus\{0\}} |\langle \xi, \mathbf{e}_{i,j}\rangle| \, m(\mathrm{d}\xi) < \infty, \qquad (6.45)
\end{aligned}$$

where the inequality in (6.45) follows from part (b) in Definition 2.3 i), which yields

$$\int_{\mathcal{H}^+\cap\{0<\|\xi\|\leq 1\}} |\langle \xi, \mathbf{e}_{i,j}\rangle| \, m(\mathrm{d}\xi) < \infty, \quad \forall\, 1 \leq i \leq j \leq d,$$

together with part (a) of Lemma 6.18 i) which yields

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

for all $1 \leq i \leq j \leq d$. Similarly, for the second integral on the right-hand side of (6.43) we see that

$$\begin{aligned}
\int_{\{\xi\in\mathbb{S}_d^+ : \, \|\xi\|_d>1\}} \|\xi\|_d^2 \, \tilde{m}_d(\mathrm{d}\xi) &\leq \int_{\mathcal{H}^+\setminus\{0\}} \|\mathbf{P}_d(\xi)\|^2 \, m(\mathrm{d}\xi) \\
&\leq \int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^2 m(\mathrm{d}\xi) < \infty,
\end{aligned}$$

which follows again from part (b) in Definition 2.3 i). Next, we prove the assertion in Lemma 6.18 ii). Note first, that by definition we have

$$\tilde{b}_d = i_d(b_d) = i_d(\mathbf{P}_d(b)) + \int_{E_d} i_d(\mathbf{P}_d(\xi)) \, m(\mathrm{d}\xi).$$

Moreover, from Definition 2.3 ii) it follows that $b \in \mathcal{H}^+$ and since $\mathbf{P}_d(\mathcal{H}^+) = \mathcal{H}_d^+$ and $i_d(\mathcal{H}_d^+) = \mathbb{S}_d^+$ we see that $\tilde{b}_d \in \mathbb{S}_d^+$. Moreover, since

$$\mathbf{1}_{E_d} - \mathbf{1}_{\mathcal{H}^+ \cap \{0 < \|\mathbf{P}_d(\xi)\| \leq 1\}} = -\mathbf{1}_{\mathcal{H}^+ \cap \{0 < \|\xi\| \leq 1\}}$$

and

$$\begin{aligned}\mathbf{P}_d(I_m) &= \sum_{i=1}^{d}\sum_{j=i}^{d} \langle I_m, \mathbf{e}_{i,j}\rangle \mathbf{e}_{i,j} = \sum_{i=1}^{d}\sum_{j=i}^{d} \Big(\int_{\mathcal{H}^+ \setminus \{0\}} \langle \chi(\xi), \mathbf{e}_{i,j}\rangle \, m(\mathrm{d}\xi) \Big) \mathbf{e}_{i,j} \\ &= \int_{\mathcal{H}^+ \setminus \{0\}} \mathbf{P}_d(\chi(\xi)) \, m(\mathrm{d}\xi),\end{aligned}$$

we conclude that

$$\tilde{b}_d - \int_{\mathbb{S}_d^+ \setminus \{0\}} \chi_d(\xi) \tilde{m}_d(\mathrm{d}\xi) = i_d\big(\mathbf{P}_d\big(b - I_m\big)\big) \geq_{\mathbb{S}_d^+} 0,$$

where it follows from Definition 2.3 ii) that $b - I_m \in \mathcal{H}^+$. We continue with Lemma 6.18 iii) and show first that the measure $\tilde{M}_d$ is a sigma-finite measure on $\mathbb{S}_d^+ \setminus \{0\}$ such that for every $A \in \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\})$ we have $\tilde{M}_d(A) \in \mathbb{S}_d^+$. For this, note that by definition $\mu(E) \in \mathcal{H}^+$ for all $E \in \mathcal{B}(\mathcal{H}^+ \setminus \{0\})$. Hence, this applied to the measurable set $(i_d \circ \mathbf{P}_d)^{-1}(A) \subseteq \mathcal{H}_d^+$ gives $\tilde{M}_d(A) \in \mathbb{S}_d^+$. Note that for every $x \in \mathbb{S}_d^+$ the kernel $\tilde{M}_d(x, \mathrm{d}\xi) \colon \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\}) \to [0, \infty]$, for all $A \in \mathcal{B}(\mathbb{S}_d^+ \setminus \{0\})$, satisfies

$$\tilde{M}_d(x, A) = \int_{\mathcal{H}^+ \setminus \{0\}} \mathbf{1}_A(i_d(\mathbf{P}_d(\xi))) \frac{1}{\|\xi\|^2} \langle x, i_d(\mathbf{P}_d(\mu(\mathrm{d}\xi)))\rangle_d,$$

or equivalently $\tilde{M}_d(x, \mathrm{d}\xi) = i_d(i_{d*} M_d(x, \mathrm{d}\xi))$. Moreover, let $x, u \in \mathbb{S}_d^+$ such that $\langle x, u\rangle_d = 0$, then

$$\begin{aligned}\int_{\mathbb{S}_d^+ \setminus \{0\}} \langle \chi_d(\xi), u\rangle_d \tilde{M}_d(x, \mathrm{d}\xi) &= \int_{\mathcal{H}^+ \setminus \{0\}} \langle \chi_d(i_d(\mathbf{P}_d(\xi))), u\rangle_d \frac{\langle x, i_d(\mathbf{P}_d(\mu(\mathrm{d}\xi)))\rangle_d}{\|\xi\|^2} \\ &= \int_{\mathcal{H}^+ \setminus \{0\}} \langle \chi(\mathbf{P}_d(\xi)), i_d^{-1}(u)\rangle \frac{\langle i_d^{-1}(x), \mathbf{P}_d(\mu(\mathrm{d}\xi))\rangle}{\|\xi\|^2} \\ &\leq \int_{\mathcal{H}^+ \setminus \{0\}} \langle \mathbf{P}_d(\xi), i_d^{-1}(u)\rangle \frac{\langle i_d^{-1}(x), \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2} \\ &= \int_{\mathcal{H}^+ \setminus \{0\}} \langle \xi, i_d^{-1}(u)\rangle \frac{\langle i_d^{-1}(x), \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2} < \infty. \quad (6.46)\end{aligned}$$

Note that in the last inequality (6.46) we used Definition 2.3 iii) and since $i_d^{-1}(x)$, $i_d^{-1}(u) \in \mathcal{H}^+$ satisfy $\langle i_d^{-1}(x), i_d^{-1}(u)\rangle = \langle x, u\rangle_d = 0$.

The property in (6.50) follows from

$$\int_{\mathbb{S}_d^+\setminus\{0\}} \|\xi\|_d^2 \tilde{M}_d(x,\mathrm{d}\xi) \leq \int_{\mathcal{H}^+\setminus\{0\}} \|\mathbf{P}_d(\xi)\|^2 \frac{\langle i_d^{-1}(x), \mu(\mathrm{d}\xi)\rangle}{\|\xi\|^2}$$
$$\leq \left\langle i_d^{-1}(x), \mu(\mathcal{H}^+ \setminus \{0\})\right\rangle < \infty.$$

Finally, we show Lemma 6.18 v). For this let $x, u \in \mathbb{S}_d^+$ be such that $\langle x, u\rangle_d = 0$ and note that

$$\int_{\mathbb{S}_d^+ \cap \{0<\|\xi\|_d\leq 1\}} \langle \chi_d(\xi), u\rangle_d \, \tilde{M}_d(x,\mathrm{d}\xi) = \int_{\mathcal{H}^+\cap\{0<\|\mathbf{P}_d(\xi)\|\leq 1\}} \langle \xi, i_d^{-1}(u)\rangle M(i_d^{-1}(x),\mathrm{d}\xi),$$

as well as

$$\langle \tilde{B}_d(x), u\rangle_d = \langle B(i_d^{-1}(x)), i_d^{-1}(u)\rangle_d + \int_{E_d} \langle \mathbf{P}_d(\xi), i_d^{-1}(u)\rangle_d \, M(i_d^{-1}(x),\mathrm{d}\xi).$$

Now again, by the identity $\mathbf{1}_{E_d} - \mathbf{1}_{\mathcal{H}^+\cap\{0<\|\mathbf{P}_d(\xi)\|\leq 1\}} = -\mathbf{1}_{\mathcal{H}^+\cap\{0<\|\xi\|\leq 1\}}$ and since $\langle i_d^{-1}(x), i_d^{-1}(u)\rangle = 0$ we conclude the inequality (6.42) from

$$\langle B(i_d^{-1}(x)), i_d^{-1}(u)\rangle - \int_{\mathcal{H}^+\cap\{0<\|\xi\|\leq 1\}} \langle \xi, i_d^{-1}(u)\rangle \, M(i_d^{-1}(x),\mathrm{d}\xi) \geq 0,$$

which holds true by Definition 2.3 iv) and proves the last assertion of Lemma 6.18. □

In the following proposition we assert the existence of a unique affine process on $\mathbb{S}_d^+$ associated with an *matrix-valued admissible parameter set* built from the parameters $\tilde{b}_d$, $\tilde{B}_d$, $\tilde{m}_d$ and $\tilde{M}_d$ and paths in $D(\mathbb{R}^+, \mathbb{S}_d^+)$.

Proposition 6.19. *Let $d \in \mathbb{N}$ and $(\tilde{b}_d, \tilde{B}_d, \tilde{m}_d, \tilde{\mu}_d)$ and $\tilde{M}_d$ be as in Definition 6.17. Then there exists a unique Markov process $(\tilde{X}^d, (\tilde{\mathbb{P}}_x)_{x\in\mathbb{S}_d^+})$, where $\tilde{\mathbb{P}}_x^d$ denotes the law of $\tilde{X}^d$ given $\tilde{X}_0^d = x \in \mathbb{S}_d^+$, with paths in $D(\mathbb{R}^+, \mathbb{S}_d^+)$ and such that for every $x \in \mathbb{S}_d^+$ we have*

$$\mathbb{E}_{\tilde{P}_x^d}\left[\mathrm{e}^{-\langle \tilde{X}_t^d, u\rangle_d}\right] = \mathrm{e}^{-\tilde{\phi}_d(t,u) - \langle x, \tilde{\psi}_d(t,u)\rangle_d}, \quad t \geq 0,\, u \in \mathbb{S}_d^+, \tag{6.47}$$

for $(\tilde{\phi}_d(\cdot,u), \tilde{\psi}_d(\cdot,u))$ the unique solution of the following equations:

$$\begin{cases} \dfrac{\partial \tilde{\phi}_d(t,u)}{\partial t} = \tilde{F}_d(\tilde{\phi}_d(t,u)), & \tilde{\phi}_d(0,u) = 0, \quad (6.48a) \\ \dfrac{\partial \tilde{\psi}_d(t,u)}{\partial t} = \tilde{R}_d(\tilde{\psi}_d(t,u)), & \tilde{\psi}_d(0,u) = u, \quad (6.48b) \end{cases}$$

where the functions $\tilde{F}_d\colon \mathbb{S}_d^+ \to \mathbb{R}$ *and* $\tilde{R}_d\colon \mathbb{S}_d^+ \to \mathbb{S}_d$ *are given by*

$$\tilde{F}_d(u) \coloneqq \langle \tilde{b}_d, u\rangle_d - \int_{\mathbb{S}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, u\rangle_d} - 1 + \langle \chi_d(\xi), u\rangle_d\right) \tilde{m}_d(\mathrm{d}\xi),$$
$$\tilde{R}_d(u) \coloneqq \tilde{B}_d^*(u) - \int_{\mathbb{S}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, u\rangle_d} - 1 + \langle \chi_d(\xi), u\rangle_d\right) \tilde{M}_d(\mathrm{d}\xi).$$

Moreover, for all $x \in \mathbb{S}_d^+$ *we have* $\tilde{\mathbb{P}}_x(\{\tilde{X}_t^d \in \mathbb{S}_d^+ \colon\ t \geq 0\}) = 1$ *and* $\tilde{X}^d$ *is a square-integrable semimartingale on* $\mathbb{S}_d^+$ *whose characteristic triplet* $(\tilde{A}^d, \tilde{C}^d, \nu^{\tilde{X}^d})$*, with respect to* χ_d*, is given by*

$$\tilde{A}_t^d = \int_0^t \left(\tilde{b}_d + \tilde{B}_d(\tilde{X}_s^d)\right) \mathrm{d}s; \quad \tilde{C}_t^d = 0; \quad \nu^{\tilde{X}^d}(\mathrm{d}t, \mathrm{d}\xi) = \left(\tilde{m}_d(\mathrm{d}\xi) + \tilde{M}_d(\tilde{X}_t^d, \mathrm{d}\xi)\right) \mathrm{d}t.$$

Proof. Given the parameters $\tilde{b}_d$, $\tilde{B}_d$, $\tilde{m}_d$ and $\tilde{M}_d$ we define the following adjusted constant and linear drift parameters $\tilde{c}_d$ and $\tilde{D}_d(u)$ for $u \in \mathbb{S}_d$ as

$$\tilde{c}_d \coloneqq \tilde{b}_d - \int_{\mathbb{S}_d^+\setminus\{0\}} \chi_d(\xi) \tilde{m}_d(\mathrm{d}\xi), \quad \tilde{D}_d(u) \coloneqq \tilde{B}_d^*(u) - \int_{\mathbb{S}_d^+\setminus\{0\}} \langle \chi_d(\xi), u\rangle_d \tilde{M}_d(\mathrm{d}\xi).$$

It then follows from the properties in Lemma 6.18 ii) and (6.42) that $\tilde{c}_d \in \mathbb{S}_d^+$ and $\langle \tilde{D}_d(u), x\rangle_d \geq 0$ for all $u, x \in \mathbb{S}_d^+$ with $\langle u, x\rangle_d = 0$. Together with the other properties shown in Lemma 6.18 we conclude that the parameter set $(0, \tilde{c}_d, \tilde{D}_d, 0, 0, \tilde{m}_d, \tilde{M}_d)$ is an *admissible parameter set* for $\mathbb{S}_d^+$-valued affine process according to [114, Definition 3.1]. It thus follows from [42, Theorem 2.4] and [114, Theorem 3.2] that there exists a unique affine process $(\tilde{X}^d)_{t\geq 0}$ with values in $\mathbb{S}_d^+$ such that for every $t \geq 0$ and $u \in \mathbb{S}_d^+$ the affine transform formula (6.47) holds with $(\tilde{\phi}_d(\cdot, u), \tilde{\psi}_d(\cdot, u))$ being the unique solution to the following equations:

$$\begin{cases} \dfrac{\partial \tilde{\phi}_d(t,u)}{\partial t} = \langle \tilde{c}_d, \tilde{\psi}_d(t,u)\rangle_d - \displaystyle\int_{\mathbb{S}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, \tilde{\psi}_d(t,u)\rangle_d} - 1\right) \tilde{m}_d(\mathrm{d}\xi), & (6.49\text{a}) \\ \dfrac{\partial \tilde{\psi}_d(t,u)}{\partial t} = \tilde{D}_d(\tilde{\psi}_d(t,u)) - \displaystyle\int_{\mathbb{S}_d^+\setminus\{0\}} \left(\mathrm{e}^{-\langle \xi, \tilde{\psi}_d(t,u)\rangle_d} - 1\right) \tilde{M}_d(\mathrm{d}\xi), & (6.49\text{b}) \end{cases}$$

with initial conditions $\tilde{\psi}_d(0,u) = u$ and $\tilde{\phi}_d(0,u) = 0$. Inserting the definition of $\tilde{c}_d$ and $\tilde{D}_d$ into (6.49) proves the equivalence with equations (6.48).
The existence of a càdlàg version follows from [44] and we shall denote this version again by $(\tilde{X}^d)_{t\geq 0}$. Moreover, we denote the law of $\tilde{X}^d$ given that $\tilde{X}_0^d = x$ by $\tilde{\mathbb{P}}_x^d$. Note that the first, fourth and fifth component of $(0, \tilde{c}_d, \tilde{D}_d, 0, 0, \tilde{m}_d, \tilde{M}_d)$ are zero, which correspond to a vanishing *diffusion* component as well as the absence of constant and linear *killing terms.*

By [42, Remark 2.5], this together with the moment assumption in Lemma 6.18 i) and (6.50) implies that the $\mathbb{S}_d^+$-valued affine process $(\tilde{X}^d)_{t\geq 0}$ satisfies

$$\tilde{\mathbb{P}}_x(\{\tilde{X}_t^d \in \mathbb{S}_d^+ \colon t \geq 0\}) = 1.$$

For every $d \in \mathbb{N}$ and $x \in \mathbb{S}_d^+$ the law $\tilde{\mathbb{P}}_x^d$ is thus defined on $\mathcal{B}(D(\mathbb{R}^+, \mathbb{S}_d^+))$. Note that since the diffusion part is zero, it follows from [114] that $(\tilde{X}^d)_{t\geq 0}$ must be of finite-variation.

Moreover, it follows from [42, Theorem 2.6] that the process $\tilde{X}_d$ is a semimartingale with characteristic given by $\tilde{C}_t^d = 0$, $\nu^{\tilde{X}^d}(\mathrm{d}t, \mathrm{d}\xi) = \big(\tilde{m}_d(\mathrm{d}\xi) + \tilde{M}_d(\tilde{X}_t^d, \mathrm{d}\xi)\big)\,\mathrm{d}t$ and

$$\begin{aligned}\tilde{A}^d &= \int_0^t \Big(\tilde{c}_d + \int_{\mathbb{S}_d^+\setminus\{0\}} \chi_d(\xi)\,\tilde{m}_d(\mathrm{d}\xi) + \tilde{D}_d(\tilde{X}_s^d) + \int_{\mathbb{S}_d^+\setminus\{0\}} \chi_d(\xi)\langle \tilde{X}_s^d, \tilde{M}_d(\mathrm{d}\xi)\rangle\Big)\,\mathrm{d}s \\ &= \int_0^t \big(\tilde{b}_d + \tilde{B}_d(\tilde{X}_s^d)\big)\,\mathrm{d}s,\end{aligned}$$

which proves the asserted form of the characteristic triplet $(\tilde{A}^d, \tilde{C}^d, \nu^{\tilde{X}^d})$. Lastly, we note that the process $(\tilde{X}_t^d)_{t\geq 0}$ is of finite-variation, hence locally bounded and by Lemma 6.18 we conclude that $\int_0^t \int_{\mathbb{S}_d^+\setminus\{0\}} \|\xi\|_d^2\,\tilde{\nu}^{\tilde{X}^d}(\mathrm{d}t, \mathrm{d}\xi) < \infty$ for all $t \geq 0$, which by [80, Proposition 2.29 b)] implies that X is a square-integrable martingale, i.e. $\mathbb{E}_{\tilde{P}_x^d}\left[\|\tilde{X}_t^d\|_d^2\right] < \infty$ for all $t \geq 0$. ☐

As a corollary from [114] we can sharpen the property in Lemma 6.18 iii).

Corollary 6.20. *Let the assumption of Lemma 6.18 hold. Then for every $d \in \mathbb{N}$ we have*

$$\int_{\mathbb{S}_d^+\setminus\{0\}} (\|\xi\|_d \vee \|\xi\|_d^2)\,\tilde{M}_d(\mathrm{d}\xi) < \infty. \tag{6.50}$$

Moreover, for every $d \in \mathbb{N}$ and μ as in Definition 2.3 iii) *we have*

$$\int_{\mathcal{H}^+\setminus\{0\}} \|\mathbf{P}_d(\xi)\|\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} < \infty. \tag{6.51}$$

Proof. From Proposition 6.19 it follows that $(0, \tilde{c}_d, \tilde{D}_d, 0, 0, \tilde{m}_d, \tilde{M}_d)$ is an admissible parameter set as in [42, Definition 2.3]. It thus follows from [114, Theorem 3.12] (which proves that the state-dependent jump measure $\tilde{M}_d(x, \mathrm{d}\xi)$ is of finite-variation) implies that for all $d \in \mathbb{N}$ the matrix-valued measure $\tilde{M}_d$ satisfies $\int_{\{\xi\in\mathbb{S}_d^+ \colon 0<\|\xi\|_d\leq 1\}} \|\xi\|_d \tilde{M}_d(\mathrm{d}\xi) < \infty$, which by the definition of $\tilde{M}_d(\mathrm{d}\xi)$ implies that $\int_{\{\xi\in\mathcal{H}^+ \colon 0<\|\xi\|\leq 1\}} \|\mathbf{P}_d(\xi)\|\|\xi\|^{-2}\mu(\mathrm{d}\xi) < \infty$, which proves (6.51) since $d \in \mathbb{N}$ was arbitrary. ☐

Remark 6.21. i) Note that by (6.51) we conclude that the state-dependent jump-measure $M(x, \mathrm{d}\xi) = \|\xi\|^{-2}\langle x, \mu(\mathrm{d}\xi)\rangle$ is of finite-variation in every direction $\mathbf{e}_{i,j}$, for $i \leq j \in \mathbb{N}$, and in every direction $v \in \mathcal{H}^+$ of finite-rank. However, in contrast to the finite-dimensional case in $\mathbb{S}_d^+$, see [114], this in general does not imply that $M(x, \mathrm{d}\xi)$ is of finite-variation, i.e. $\int_{\mathcal{H}^+\setminus\{0\}} \|\xi\| M(x, \mathrm{d}\xi) < \infty$ $(\forall\, x \in \mathcal{H}^+)$. Indeed, due to the infinite dimensions of $\mathcal{H}$ there are "infinite many directions", in each of which the jumps evolve with finite-variation, but in sum, over all coordinates, the variation could be infinite, see Section 6.4.

ii) The situation described in i) is a typical, although not necessary, infinite-dimensional phenomenon. Indeed, let V be an infinite-dimensional Banach space and $D_0 \subseteq V$, then the question whether $\int_{D_0} \langle \xi, u\rangle_{V^*}\, \nu(\mathrm{d}\xi) < \infty$ for all $u \in V^*$ (the Pettis integrability on D_0), implies $\int_{D_0} \|\chi(\xi)\|_V\, \nu(\mathrm{d}\xi)$ (the Bochner integrability on D_0), where V^* denotes the Banach dual of V with dual pairing $\langle\cdot,\cdot\rangle_{V^*}$, also depends on the state space V. In case of Hilbert-Schmidt operators, $D_0 = \{\xi \in \mathcal{H}^+ : 0 < \|\xi\| \leq 1\}$ and $\nu(\mathrm{d}\xi) = \nu(x, \mathrm{d}\xi)$ this implication does not hold true. In contrast, in an analogous situation on the space of trace-class operators the above implication does hold, see [122].

6.7.3 Proof: Existence of finite-rank operator-valued affine processes

For every $d \in \mathbb{N}$, let $(\tilde{X}^d, (\tilde{\mathbb{P}}_x^d)_{x\in\mathbb{S}_d^+})$ be the $\mathbb{S}_d^+$-valued affine process given by Proposition 6.19. More specifically, let $\tilde{X}^d$ be a version with paths in $D(\mathbb{R}^+, \mathbb{S}_d^+)$ realized on $\Omega = D(\mathbb{R}^+, \mathbb{S}_d^+)$ (see [42]) and denote by $\tilde{\mathbb{P}}_x^d$ the law of X^d, defined on $\mathcal{B}(D(\mathbb{R}^+, \mathbb{S}_d^+))$, given $\tilde{X}_0^d = x \in \mathbb{S}_d^+$. Moreover, let us denote by $(\tilde{\mathcal{F}}_t^d)_{t\geq 0}$ the natural filtration of the process $\tilde{X}^d$. By identifying the cones $\mathbb{S}_d^+$ and $\mathcal{H}_d^+$ under the mapping i_d^{-1}, we define the process $X^d = (X_t^d)_{t\geq 0}$ as

$$X_t^d \coloneqq i_d^{-1}(\tilde{X}_t^d) = \Phi_d \circ \tilde{X}_t^d \circ \Phi_d^{-1}, \quad t \geq 0.$$

Note that the process $(X_t^d)_{t\geq 0}$ has paths in $D(\mathbb{R}^+, \mathcal{H}_d^+)$ and the law of X^d is given by the push-forward measure $(i_d^{-1})_*\tilde{\mathbb{P}}_x$ for $x \in \mathbb{S}_d^+$, where we understand that i_d^{-1} acts pointwise on the functions in $D(\mathbb{R}^+, \mathbb{S}_d^+)$, i.e. $i_d^{-1}(D(\mathbb{R}^+, \mathbb{S}_d^+)) = D(\mathbb{R}^+, \mathcal{H}_d^+)$. Moreover, we see that $D(\mathbb{R}^+, \mathcal{H}_d^+) \subseteq D(\mathbb{R}^+, \mathcal{H}^+)$ for all $d \in \mathbb{N}$, see [83, Remark 4.5]. For every $x \in \mathcal{H}^+$ we thus define the measure $\mathbb{P}_x^d$ on $D(\mathbb{R}^+, \mathcal{H}^+)$ as

$$\mathbb{P}_x^d(A) = (i_d^{-1})_*\tilde{\mathbb{P}}^d_{i_d(\mathbf{P}_d(x))}(A \cap D(\mathbb{R}^+, \mathcal{H}_d^+)), \quad A \in \mathcal{B}(D(\mathbb{R}^+, \mathcal{H}^+)).$$

Note that $\mathbb{P}_x^d\big(X_0^d = \mathbf{P}_d(x)\big) = \tilde{\mathbb{P}}^d_{i_d(\mathbf{P}_d(x))}\big(\tilde{X}_0^d = i_d(\mathbf{P}_d(x))\big) = 1$ and the process $(X^d, (\mathbb{P}_x^d)_{x\in\mathcal{H}^+})$ is again a Markov process realized on the space $D(\mathbb{R}^+, \mathcal{H}^+)$ with respect to its natural filtration $\mathbb{F}^d = (\mathcal{F}_t^d)_{t\geq 0}$, where we set $\mathcal{F}^d = \mathcal{F}_\infty^d$.

Note that alternatively to the above construction we could also use [93, Proposition 4.7] to embed lower-dimensional affine processes into a larger ambient space, but we prefer the direct approach here.
Moreover, for every $x \in \mathcal{H}^+$ and X^d as above with $X_0^d = \mathbf{P}_d(x)$, we denote by $\mathcal{N}_x^d$ the collection of all $\mathbb{P}_x^d$-null sets of $\mathcal{F}^d$ and set $\bar{\mathcal{F}}_t := \mathcal{F}_t \vee \mathcal{N}_x^d$ for every $t \geq 0$. We define $\bar{\mathbb{F}}^d := (\bar{\mathcal{F}}_t)_{t\geq 0}$, i.e. $\bar{\mathbb{F}}^d$ is the usual augmented filtration of X^d and note that the process X^d is still a Markov with respect to $\bar{\mathbb{F}}^d$, see [56]. In addition to that, we show in the following proposition that $(X^d, (\mathbb{P}_x^d)_{x\in\mathcal{H}})$ satisfies an affine transform formula associated with the Galerkin approximations in (6.1) and, moreover, that the process X^d is a semimartingale with respect to the stochastic basis $(\Omega, \bar{\mathcal{F}}^d, \bar{\mathbb{F}}^d, \mathbb{P}_x^d)$ as above, with $\Omega = D(\mathbb{R}^+, \mathcal{H}^+)$.

Proposition 6.22. *Let (b, B, m, μ) be an admissible parameter set and for $d \in \mathbb{N}$ let (b_d, B_d, m_d, μ_d) and M_d be as in Definition 6.13. Then for every $d \in \mathbb{N}$ the process $(X^d, (\mathbb{P}_x)_{x\in\mathcal{H}^+})$ defined as above is a Markov process on $\mathcal{H}^+$ such that for every $x \in \mathcal{H}^+$ we have*

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\mathrm{e}^{-\langle X_t^d, \mathbf{P}_d(u)\rangle}\right] = \mathrm{e}^{-\phi_d(t,\mathbf{P}_d(u)) - \langle \mathbf{P}_d(x), \psi_d(t,\mathbf{P}_d(u))\rangle}, \quad t \geq 0,\, u \in \mathcal{H}_d^+, \tag{6.52}$$

for $\big(\phi_d(\cdot, \mathbf{P}_d(u)), \psi_d(\cdot, \mathbf{P}_d(u))\big)$ the unique solution of (6.1). Moreover, for every $x \in \mathcal{H}^+$ we have

$$\mathbb{P}_x^d(\{X_t^d \in \mathcal{H}_d^+ \colon\ t \geq 0\}) = 1, \tag{6.53}$$

and $(X_t^d)_{t\geq 0}$ is a semimartingale with stochastic basis $(\Omega, \bar{\mathcal{F}}^d, \bar{\mathbb{F}}^d, \mathbb{P}_x^d)$ whose characteristic triplet (A^d, C^d, ν^{X^d}), with respect to χ, is given by:

$$A_t^d = \int_0^t b_d + B_d(X_s^d)\,\mathrm{d}s; \quad C_t^d \quad = 0; \quad \nu^{X^d}(\mathrm{d}t, \mathrm{d}\xi) = \big(m_d(\mathrm{d}\xi) + M_d(X_t^d, \mathrm{d}\xi)\big)\,\mathrm{d}t.$$

Proof. Let $d \in \mathbb{N}$, $x \in \mathcal{H}^+$ and let $(\tilde{X}_t^d)_{t\geq 0}$ be the unique affine process on $\mathbb{S}_d^+$ associated with the parameter set $(0, \tilde{c}_d, \tilde{D}_d, 0, 0, \tilde{m}_d, \tilde{M}_d)$ and such that $\tilde{X}_0^d = i_d(\mathbf{P}_d(x))$. For $u \in \mathcal{H}^+$ we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}_x^d}\left[\mathrm{e}^{-\langle \mathbf{P}_d(u), X_t^d\rangle}\right] &= \mathbb{E}_{\mathbb{P}_x^d}\left[\mathrm{e}^{-\langle \mathbf{P}_d(u), i_d^{-1}(\tilde{X}_t^d)\rangle}\right] \\
&= \mathbb{E}_{\tilde{\mathbb{P}}^d_{i_d(\mathbf{P}_d(x))}}\left[\mathrm{e}^{-\langle i_d\mathbf{P}_d(u), \tilde{X}_t^d\rangle_d}\right] \\
&= \mathrm{e}^{-\tilde{\phi}_d(t, i_d(\mathbf{P}_d(u))) - \langle i_d\mathbf{P}_d(x), \tilde{\psi}_d(t, i_d(\mathbf{P}_d(u)))\rangle_d} \\
&= \mathrm{e}^{-\tilde{\phi}_d(t, i_d(\mathbf{P}_d(u))) - \langle \mathbf{P}_d(x), i_d^{-1}\tilde{\psi}_d(t, i_d(\mathbf{P}(u)))\rangle}.
\end{aligned}$$

This proves that the process X_t^d satisfies the affine transform formula with functions $\tilde{\phi}_d(t, i_d(\mathbf{P}_d(u)))$ and $i_d^{-1}(\tilde{\psi}_d(t, i_d(\mathbf{P}_d(u)))$. Therefore, in order to prove (6.5), it is left to show that $\big(\phi_d(\cdot, \mathbf{P}_d(u)), \psi_d(\cdot, \mathbf{P}_d(u))\big)$, the unique solution of (6.1a)-(6.1b) coincides with the function $\big(\tilde{\phi}_d(\cdot, i_d(\mathbf{P}_d(u))), i_d^{-1}(\tilde{\psi}_d(\cdot, i_d(\mathbf{P}_d(u))))\big)$.

For this, let us again consider $K\colon \mathcal{H}\times\mathcal{H}\to\mathbb{R}$ given by $K(u,v):=\mathrm{e}^{-\langle u,v\rangle}-1+\langle\chi(u),v\rangle$ and for every $u\in\mathcal{H}^+$, we set $\tilde{u}_d:=i_d(\mathbf{P}_d(u))$. Then we see that for all $t\geq 0$ and $u\in\mathcal{H}^+$ the function $i_d^{-1}(\tilde{\psi}_d(\cdot,i_d(\mathbf{P}_d(u))))$ satisfies the following equation:

$$\begin{aligned}\frac{\partial i_d^{-1}(\tilde{\psi}_d(t,\tilde{u}_d))}{\partial t}&=i_d^{-1}(\tilde{R}_d(\tilde{\psi}_d(t,\tilde{u}_d)))\\&=i_d^{-1}(\tilde{B}_d^*(\tilde{\psi}_d(t,\tilde{u}_d)))-\int_{\mathbb{S}_d^+\setminus\{0\}}K\big(\xi,\tilde{\psi}_d(t,i_d(\mathbf{P}_d(u)))\big)i_d^{-1}(\tilde{M}_d(\mathrm{d}\xi))\\&=B_d^*(i_d^{-1}(\tilde{\psi}(t,\tilde{u}_d))-\int_{\mathcal{H}_d^+\setminus\{0\}}K\big(\xi,i_d^{-1}(\psi_d(t,\tilde{u}_d))\big)\,M_d(\mathrm{d}\xi)),\end{aligned}$$

and $i_d^{-1}(\tilde{\psi}_d(0,\tilde{u}_d))=\tilde{u}_d=i_d^{-1}(i_d\mathbf{P}_d(u))=\mathbf{P}_d(u)$. But, since (6.1b) is uniquely solved by $\psi_d(\cdot,\mathbf{P}_d(u))$ we conclude that $\psi_d(\cdot,\mathbf{P}_d(u))=i_d^{-1}(\tilde{\psi}_d(\cdot,i_d(\mathbf{P}_d(u))))$. Similarly, for $\tilde{\phi}_d(\cdot,\tilde{u}_d)$ we find

$$\begin{aligned}\frac{\partial\tilde{\phi}_d(t,\tilde{u}_d)}{\partial t}&=\tilde{F}_d(\tilde{\psi}_d(t,\tilde{u}_d))\\&=\langle\tilde{b}_d,\tilde{\psi}_d(t,\tilde{u}_d)\rangle-\int_{\mathbb{S}_d^+\setminus\{0\}}K\big(\xi,\tilde{\psi}_d(t,\tilde{u}_d)\big)\,\tilde{m}_d(\mathrm{d}\xi)\\&=\langle b_d,i_d^{-1}\tilde{\psi}_d(t,\tilde{u}_d)\rangle-\int_{\mathcal{H}_d^+\setminus\{0\}}K\big(\xi,i_d^{-1}\tilde{\psi}_d(t,\tilde{u}_d)\big)\,m_d(\mathrm{d}\xi),\end{aligned}$$

and $\tilde{\phi}_d(0,\tilde{u}_d)=0$. Again by the uniqueness of the solution to (6.1a) we conclude that $\phi_d(\cdot,\mathbf{P}_d(u))=\tilde{\phi}_d(\cdot,i_d(\mathbf{P}_d(u)))$, which finally proves (6.5). Moreover, the property (6.53) follows from Proposition 6.19 and

$$\mathbb{P}_x^d\big(\big\{X_t^d\in\mathcal{H}_d^+\colon\ t\geq 0\big\}\big)=\tilde{\mathbb{P}}^d_{i_d(\mathbf{P}_d(x))}\big(\big\{\tilde{X}_t^d\in\mathbb{S}_d^+\colon\ t\geq 0\big\}\big)=1.$$

The asserted form of the characteristic triplet and square-integrability follows immediately from the analogous property in the matrix-valued case and an application of the linear isometric transformation i_d^{-1}. □

With Proposition 6.22 we have already shown the first part of Theorem 6.2. In the next proposition we assert some additional properties of the process X^d. In particular, we show that X^d solves the martingale problem for $\mathcal{G}^d$, from which we conclude that the second assertion of Theorem 6.2 holds true, and we extend the operators $(\mathcal{G}^d)_{d\in\mathbb{N}}$ from Fourier-basis elements to linear and quadratic functions on $\mathcal{H}^+$.

Proposition 6.23. *For $d \in \mathbb{N}$ and $x \in \mathcal{H}_d$, let X^d denote the affine process on $\mathcal{H}_d^+$ with $X_0^d = \mathbf{P}_d(x)$ as in Proposition 6.22. Then the process $(\bar{J}_t^d)_{t\geq 0}$, given by*

$$\bar{J}_t^d := X_t^d - \mathbf{P}_d(x) - \int_0^t \Big(b_d + B_d(X_s^d) - \int_{\mathcal{H}_d^+ \cap \{\|\xi\|>1\}} \xi \left(m_d(\mathrm{d}\xi) + M_d(X_s^d, \mathrm{d}\xi) \right) \Big) \,\mathrm{d}s, \tag{6.54}$$

is a square-integrable martingale on $\mathcal{H}_d$. Moreover, we define for every $f \in$ $\mathrm{dom}(\mathcal{G}^d) := \mathrm{lin}\left\{ \mathrm{e}^{-\langle \cdot, u\rangle}, \langle \cdot, u\rangle, \langle \cdot, u\rangle^2 \colon u \in \mathcal{H}_d^+ \right\}$ the operator $\mathcal{G}^d$ as

$$\mathcal{G}^d f(x) = \langle b_d + B_d(x), f'(x)\rangle + \int_{\mathcal{H}_d^+ \setminus \{0\}} \big(f(x+\xi) - f(x) - \langle \chi(\xi), f'(x)\rangle \big) \, \nu(x, \mathrm{d}\xi), \tag{6.55}$$

where $f'(x)$ denotes the first derivative of f at $x \in \mathcal{H}^+$. Then for all $f \in \mathrm{dom}(\mathcal{G}^d)$ the process

$$\Big(f(X_t^d) - f(\mathbf{P}_d(x)) - \int_0^t \mathcal{G}^d f(X_s) \,\mathrm{d}s \Big)_{t\geq 0}, \tag{6.56}$$

is a real-valued martingale.

Proof. Note first that we can extend the operator B_d to $\mathcal{H}$ by setting $B_d(u) :=$ $B_d(\mathbf{P}_d(u))$ for $u \in \mathcal{H}^+$ and the measures m_d and μ_d to $\mathcal{B}(\mathcal{H}^+ \setminus \{0\})$ by setting $m_d(A) = m_d(A \cap (\mathcal{H}_d^+ \setminus \{0\}))$ for $A \in \mathcal{B}(\mathcal{H}^+ \setminus \{0\})$ and analogously for μ_d. We denote these extended parameters again by B_d, m_d and μ_d and note that (b_d, B_d, m_d, μ_d) satisfies the conditions in Definition 2.3. The representation (6.54) thus follows from Proposition 3.5. Moreover, we see that the operator $\mathcal{G}^d$ defined in (6.55) on $\mathrm{dom}(\mathcal{G}^d)$ coincides with the *weak generator* as in Definition 3.2 and that the processes in (6.56) are real-valued martingales for all $f \in \mathrm{dom}(\mathcal{G}^d)$ thus follows from Proposition 3.6. Note, in particular, that $\mathcal{G}^d$ applied to $\mathrm{e}^{-\langle \cdot, u\rangle}$ evaluated at $x \in \mathcal{H}^+$ can be computed as

$$\begin{aligned}
\mathcal{G}^d \,\mathrm{e}^{-\langle \cdot, u\rangle}(x) = \Big(&- \langle b_d + B_d(x), \mathbf{P}_d(u)\rangle \\
&+ \int_{\mathcal{H}_d^+ \setminus \{0\}} \big(\mathrm{e}^{-\langle \xi, \mathbf{P}_d(u)\rangle} - 1 + \langle \chi(\xi), \mathbf{P}_d(u)\rangle \big) \, \nu(x, \mathrm{d}\xi) \Big) \, \mathrm{e}^{-\langle x, \mathbf{P}_d(u)\rangle} \\
= \big(-F_d(u) &- \langle x, R_d(u)\rangle \big) \, \mathrm{e}^{-\langle x, \mathbf{P}_d(u)\rangle},
\end{aligned}$$

and we see that $\mathcal{G}^d$ defined in (6.55) coincides with $\mathcal{G}^d$ in (6.4), which explains the notation. The second assertion of Theorem 6.2 follows since by construction we consider X to be the canonical process of $\mathbb{P}_x$ on $\Omega = D(\mathbb{R}^+, \mathcal{H}^+)$, which by the arguments above solves the martingale problem for $\mathcal{G}^d$ on the stochastic basis as in Theorem 6.2. □

6.8 Proof: Weak convergence of positive finite-rank operator-valued affine processes

Let (b, B, m, μ) be an admissible parameter set and for every $d \in \mathbb{N}$ let X^d denote the associated affine finite-rank operator-valued process given by Proposition 6.22. In this section we study the tightness and weak-convergence of the sequence $(X^d)_{d\in\mathbb{N}}$ on the space $D(\mathbb{R}^+, \mathcal{H}^+)$ equipped with the Skorohod topology. More precisely, for every $x \in \mathcal{H}^+$ we consider the sequence $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ of laws of X^d, given that $X_0^d = \mathbf{P}_d(x)$, defined on the Borel-σ-algebra $\mathcal{B}(D(\mathbb{R}^+, \mathcal{H}^+))$ and study its weak convergence as d tends to infinity. For this, we shall first prove that the sequence of laws $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ is tight on $\mathcal{B}(D(\mathbb{R}^+, \mathcal{H}^+))$, which we prove in Section 6.8.1. Subsequently, in Section 6.8.2, we prove weak convergence of $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ to a unique probability measure $\mathbb{P}_x$ on $\mathcal{B}(D(\mathbb{R}^+, \mathcal{H}^+))$, the canonical process X of which turns out to be the desired affine process on $\mathcal{H}^+$ with $X_0 = x$ and we prove the remaining assertions of Theorem 6.4.

6.8.1 Tightness

To prove the tightness of the sequence $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ we use the *Aldous criterion* in [87, Theorem 2.2.2], which we shall recall in the beginning of the proof of Proposition 6.25 below. We first need the following lemma:

Lemma 6.24. *Let $x \in \mathcal{H}^+$, $T > 0$ and for every $d \in \mathbb{N}$ denote by $(\bar{J}_t^d)_{t\geq 0}$ the square-integrable martingale given by* (6.54)*. Then there exists a constant $K_T \geq 0$ such that the following inequalities hold true:*

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq t\leq T} \|X_t^d\|^2\right] \leq K_T(1 + \|x\|^2), \tag{6.57}$$

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq t\leq T} \|\bar{J}_t^d\|^2\right] \leq K_T(1 + \|x\|^2). \tag{6.58}$$

Moreover, K_T can be chosen independently of $d \in \mathbb{N}$.

Proof. Let $d \in \mathbb{N}$ and (b_d, B_d, m_d, μ_d) and M_d be as in Definition 6.13. Then define $\hat{b}_d \coloneqq b_d + \int_{\mathcal{H}_d^+\cap\{\|\xi\|>1\}} \xi\, m_d(\mathrm{d}\xi)$ and the function $\hat{B}_d \colon \mathcal{H}_d^+ \to \mathcal{H}_d$ by

$$\hat{B}_d(u) \coloneqq B_d(u) + \int_{\mathcal{H}_d^+\cap\{\|\xi\|>1\}} \xi\, \langle u, M_d(\mathrm{d}\xi)\rangle, \quad u \in \mathcal{H}_d.$$

By Proposition 6.23 we have $X_t^d = \mathbf{P}_d(x) + H_t^d + \bar{J}_t^d$ for every $t \in [0, T]$, where $(\bar{J}_t^d)_{0\leq t\leq T}$ denotes the square-integrable martingale in (6.54) on $[0, T]$.

Moreover, we write $(H_t^d)_{0\leq t\leq T}$ for the finite-variation process given by

$$H_t^d \coloneqq \int_0^t \big(\hat{b}_d + \hat{B}_d(X_s^d)\big)\,\mathrm{d}s, \quad 0\leq t\leq T. \tag{6.59}$$

Therefore, we obtain

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq t\leq T}\|X_t^d\|^2\right] \leq 3\|\mathbf{P}_d(x)\|^2 + 3\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq t\leq T}\|H_t^d\|^2\right] + 3\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq t\leq T}\|\bar{J}_t^d\|^2\right]. \tag{6.60}$$

Inserting (6.59) into the second term on the right-hand side of (6.60) yields

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq t\leq T}\|H_t^d\|^2\right] &\leq 2T^2\|\hat{b}_d\|^2 + 2\|\hat{B}_d\|^2_{\mathcal{L}(\mathcal{H}_d)}\int_0^T \mathbb{E}_{\mathbb{P}_x^d}\left[\|X_s^d\|^2\right]\mathrm{d}s \\ &\leq 2T^2\|\hat{b}\|^2 + 2\|\hat{B}\|^2_{\mathcal{L}(\mathcal{H})}\int_0^T \mathbb{E}_{\mathbb{P}_x^d}\left[\|X_s^d\|^2\right]\mathrm{d}s, \end{aligned}\tag{6.61}$$

where the latter inequality for $\hat{b} \coloneqq b + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\xi\, m(\mathrm{d}\xi)$ and linear function $\hat{B}(\cdot) \coloneqq B(\cdot) + \int_{\mathcal{H}^+\cap\{\|\xi\|\geq 1\}}\xi\,\langle\,\cdot\,, M(\mathrm{d}\xi)\rangle$ holds by Remark 6.14. For the second term in (6.60), we recall from [117, Theorem 20.6] that

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq t\leq T}\|\bar{J}_t^d\|^2\right] \leq 4\mathbb{E}_{\mathbb{P}_x^d}\left[\big\langle\bar{J}^d\big\rangle_T\right], \tag{6.62}$$

where we denote by $\big(\big\langle\bar{J}^d\big\rangle_t\big)_{0\leq t\leq T}$ the predictable quadratic covariation process of the square-integrable martingale $(\bar{J}_t^d)_{0\leq t\leq T}$. Now, let $(\mathbf{e}_{i,j})_{i\leq j\in\mathbb{N}}$ be the same orthonormal basis of $\mathcal{H}$ that we used throughout this section. For $i\leq j\in\mathbb{N}$ we set $\bar{J}_t^{(i,j),d} \coloneqq \langle\bar{J}_t^d, \mathbf{e}_{i,j}\rangle$ and denote by $\big(\big\langle\bar{J}^{(i,j),d}\big\rangle_t\big)_{t\geq 0}$ the unique real-valued increasing process such that

$$\big((\bar{J}^{(i,j),d})^2 - \big\langle\bar{J}^{(i,j),d}\big\rangle\big)_{0\leq t\leq T},$$

is a martingale. Moreover, as in [117, Section 20], we denote by $\big(\big\langle\bar{J}^d\big\rangle_t\big)_{0\leq t\leq T}$ the unique predictable and increasing process such that $\big(\|\bar{J}^d\|^2 - \big\langle\bar{J}^d\big\rangle\big)_{0\leq t\leq T}$ is a martingale. Note that $\big\langle\bar{J}^d\big\rangle_t = \sum_{i\leq j}^d\big\langle\bar{J}^{(i,j),d}\big\rangle_t$ for every $0\leq t\leq T$ and thus only the form of the processes $\big(\big\langle\bar{J}^{(i,j),d}\big\rangle_t\big)_{0\leq t\leq T}$ for $1\leq i\leq j\leq d$ is left to compute. By an application of the *Carré-du-champs formula*, see e.g. [87, Lemma 3.1.3], we see that

$$\big\langle\bar{J}^{(i,j),d}\big\rangle_t = \int_0^t \mathcal{G}^d\langle X_s^d, \mathbf{e}_{i,j}\rangle^2 - 2\langle X_s^d, \mathbf{e}_{i,j}\rangle\mathcal{G}^d\langle X_s^d, \mathbf{e}_{i,j}\rangle\,\mathrm{d}s, \quad 0\leq t\leq T. \tag{6.63}$$

Note here that the operator $\mathcal{G}^d$ is as in (6.55) extended by linearity to the set $\operatorname{lin}(\{\langle\cdot,\mathbf{P}_d(u)\rangle,\langle\cdot,\mathbf{P}_d(u)\rangle^2 \colon u\in\mathcal{H}\})$, as we did in Lemma 2.18 and Proposition 2.38. We thus obtain

$$\mathcal{G}^d\langle x,\mathbf{e}_{i,j}\rangle = \langle b_d + B_d(x),\mathbf{e}_{i,j}\rangle + \int_{\mathcal{H}^+\cap\{\|\xi\|>1\}}\langle\xi,\mathbf{e}_{i,j}\rangle\left(m_d(\mathrm{d}\xi)+\langle x,M_d(\mathrm{d}\xi)\rangle\right), \tag{6.64}$$

$$\mathcal{G}^d\langle x,\mathbf{e}_{i,j}\rangle^2 = \int_{\mathcal{H}^+\setminus\{0\}}\langle\xi,\mathbf{e}_{i,j}\rangle^2\left(m_d(\mathrm{d}\xi)+\langle x,M_d(\mathrm{d}\xi)\rangle\right) + 2\langle x,\mathbf{e}_{i,j}\rangle\mathcal{G}^d\langle x,\mathbf{e}_{i,j}\rangle. \tag{6.65}$$

Inserting (6.64) and (6.65) into (6.63) yields

$$\left\langle \bar{J}^{(i,j),d}\right\rangle_t = \int_0^t\int_{\mathcal{H}^+\setminus\{0\}}\langle\xi,\mathbf{e}_{i,j}\rangle^2\left(m_d(\mathrm{d}\xi)+\langle X_s^d,M_d(\mathrm{d}\xi)\rangle\right)\,\mathrm{d}s,\quad t\in[0,T].$$

Now, since we have

$$\sum_{i\le j}^d\Big(\int_{\mathcal{H}_d^+\setminus\{0\}}\langle\xi,\mathbf{e}_{i,j}\rangle^2\,m_d(\mathrm{d}\xi)\Big) = \int_{\mathcal{H}_d^+\setminus\{0\}}\|\xi\|^2\,m_d(\mathrm{d}\xi)\le\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\,m(\mathrm{d}\xi),$$

and moreover for every $s\in[0,t]$

$$\begin{aligned}\sum_{i\le j}^d\Big(\int_{\mathcal{H}_d^+\setminus\{0\}}\langle\xi,\mathbf{e}_{i,j}\rangle^2\langle X_s^d,M_d(\mathrm{d}\xi)\rangle\Big) &= \int_{\mathcal{H}_d^+\setminus\{0\}}\|\xi\|^2\langle X_s^d,M_d(\mathrm{d}\xi)\rangle\\ &= \int_{\mathcal{H}^+\setminus\{0\}}\frac{\|\mathbf{P}_d(\xi)\|^2}{\|\xi\|^2}\langle X_s^d,\mathbf{P}_d(\mu(\mathrm{d}\xi))\rangle\\ &\le\langle X_s^d,\mu(\mathcal{H}^+\setminus\{0\})\rangle,\end{aligned}$$

we conclude that for every $d\in\mathbb{N}$ and $0\le t\le T$ the following inequality holds

$$\left\langle\bar{J}^d\right\rangle_t = \sum_{i\le j}^d\left\langle\bar{J}^{(i,j),d}\right\rangle_t\le\int_0^t\Big(\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\,m(\mathrm{d}\xi)+\langle X_s^d,\mu(\mathcal{H}^+\setminus\{0\})\rangle\Big)\,\mathrm{d}s.$$

From this it follows that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_x^d}\left[\left\langle\bar{J}^d\right\rangle_T\right] &\le T\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\,m(\mathrm{d}\xi)+\|\mu(\mathcal{H}^+\setminus\{0\})\|\Big(\int_0^T\mathbb{E}_{\mathbb{P}_x^d}\left[\|X_s^d\|\right]\mathrm{d}s\Big)\\ &\le T\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\,m(\mathrm{d}\xi)+\|\mu(\mathcal{H}^+\setminus\{0\})\|\Big(\int_0^T\mathbb{E}_{\mathbb{P}_x^d}\left[1+\|X_s^d\|^2\right]\mathrm{d}s\Big).\end{aligned} \tag{6.66}$$

Hence, inserting (6.66) and (6.61) back into (6.60) gives

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\le t\le T}\|X_t^d\|^2\right] \le 6T^2\|\hat{b}\|^2 + \|\mathbf{P}_d(x)\|^2 \\ + 12T\Big(\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\, m(\mathrm{d}\xi) + T\|\mu(\mathcal{H}^+\setminus\{0\})\|\Big) \\ + 6\big(\|\hat{B}\|_{\mathcal{L}(\mathcal{H})} + 12\|\mu(\mathcal{H}^+\setminus\{0\})\|\big)\int_0^T \mathbb{E}_{\mathbb{P}_x^d}\left[\|X_s^d\|^2\right]\mathrm{d}s.\end{aligned}$$

Therefore, setting $K_{1,T} = 6T^2\|\hat{b}\|^2 + 12T\big(\int_{\mathcal{H}^+\setminus\{0\}}\|\xi\|^2\, m(\mathrm{d}\xi) + T\mu(\mathcal{H}^+\setminus\{0\})\big)$ and $K_2 = 6\big(\|\hat{B}\|_{\mathcal{L}(\mathcal{H})} + 12\mu(\mathcal{H}^+\setminus\{0\})\big)$ (where we note that $K_{1,T}$ and K_2 do not depend on $d\in\mathbb{N}$) and by applying Gronwall's inequality we find that

$$\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\le t\le T}\|X_t^d\|^2\right] \le \mathrm{e}^{K_2T}(K_{1,T} + \|\mathbf{P}_d(x)\|^2) \le \tilde{K}_{1,T}(1+\|x\|^2),$$

for some $\tilde{K}_{1,T}$, independent of $d\in\mathbb{N}$, which proves inequality (6.57). Inserting, this back into (6.66) yields (6.58) for a suitable $\tilde{K}_{2,T}$ and choosing $K_T = \max(\tilde{K}_{1,T}, \tilde{K}_{2,T})$ proves the assertion. □

Recall the Hilbert space $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}})$ from (6.7) and let Assumption $\mathcal{E}$ be satisfied. Then the embedding of $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}})$ into $(\mathcal{L}_2(H), \langle\cdot,\cdot\rangle)$ is compact as well, i.e. $\mathcal{V}\subset\subset\mathcal{L}_2(H)$, see [139, Proposition 2.1]. Moreover, we note that $\mathcal{H}_d\subseteq\mathcal{V}_0\cap\mathcal{H}$ for all $d\in\mathbb{N}$, see [128]. In the next proposition we prove that the sequence $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ is tight on $\mathcal{B}(D(\mathbb{R}^+,\mathcal{H}^+))$.

Proposition 6.25. *Let Assumption $\mathcal{E}$ be satisfied. Then for every $x\in\mathcal{H}^+$ the sequence $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ of laws of $(X^d)_{d\in\mathbb{N}}$ is a tight sequence of measures on $\mathcal{B}(D(\mathbb{R}^+,\mathcal{H}^+))$.*

Proof. Let $x\in\mathcal{H}^+$. As mentioned before, we use the *tightness criterion from Aldous*, see [87, Theorem 2.2.2]. For the readers convenience we recall in the following the two sufficient conditions implying the tightness of $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$:

i) For every $t\ge 0$ the sequence of laws of $(X_t^d)_{d\in\mathbb{N}}$ form a tight sequence of probability measures on $\mathcal{B}(\mathcal{H}^+)$, the Borel-$\sigma$-algebra on $\mathcal{H}^+$.

ii) For every $T>0$, $\varepsilon>0$, $\eta>0$ there exists a $\delta>0$ and $N_0\in\mathbb{N}$ such that for every sequence of stopping times $(\tau_d)_{d\in\mathbb{N}}$ with $\tau_d\le T$ for all $d\in\mathbb{N}$, we have:

$$\sup_{d\ge N_0}\sup_{0\le\theta\le\delta}\mathbb{P}_x^d(\|X_{\tau_d}^d - X_{\tau_d+\theta}^d\|\ge\eta)\le\varepsilon. \tag{6.67}$$

We begin with the first condition: Recall that for all $d \in \mathbb{N}$ the processes X^d satisfies $\mathbb{P}_x^d(\{X_t^d \in \mathcal{H}_d^+ : t \geq 0\}) = 1$. In particular, for every fixed $t \geq 0$ it holds that $\mathbb{P}_x^d(X_t^d \in \mathcal{H}_d^+) = 1$. Now, note that $\mathcal{H}_d^+ \subseteq \mathcal{V}_0 \cap \mathcal{H}^+$ for all $d \in \mathbb{N}$ and since $\mathcal{V}$ is compactly embedded in $\mathcal{L}_2(H)$ and $\mathcal{H}^+$ is a closed subset of $\mathcal{L}_2(H)$, we see that also $\mathcal{V}_0 \cap \mathcal{H}^+$ is compact in $\mathcal{H}^+$. Hence, we see that $\mathbb{P}_x^d(\{X_t^d \in \mathcal{V}_0 \cap \mathcal{H}^+\}) = 1$ for every $d \in \mathbb{N}$, which proves the tightness of the sequence of laws of $(X_t^d)_{d\in\mathbb{N}}$. Since $t \geq 0$ was arbitrary, we therefore conclude that condition i) is satisfied. We continue with the second condition. For this let $T > 0$, $\varepsilon > 0$, $\eta > 0$ and let $(\tau_d)_{d\in\mathbb{N}}$ be a sequence of stopping times such that $\tau_d \leq T$ for all $d \in \mathbb{N}$. As before in the proof of Lemma 6.24 we consider for every $t \geq 0$ the decomposition $X_t^d = \mathbf{P}_d(x) + H_t^d + \bar{J}_t^d$ into the finite variation part $(H_t^d)_{t\geq 0}$ given by (6.59) and the purely-discontinuous martingale part $(\bar{J}_t^d)_{t\geq 0}$ in (6.54). For the finite-variation part we compute

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}_x^d}\left[\|H_{\tau_d}^d - H_{\tau_d+\theta}^d\|^2\right] &\leq \mathbb{E}_{\mathbb{P}_x^d}\left[\|\int_{\tau_d}^{\tau_d+\theta} \big(\hat{b}_d + \hat{B}_d(X_s^d)\big)\,\mathrm{d}s\|^2\right] \\
&\leq \theta^2 \mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq\tau\leq\theta} \big(\|\hat{b}\| + \|\hat{B}\|_{\mathcal{L}(\mathcal{H})}\|X_{\tau_d+\tau}^d\|^2\big)\right] \\
&\leq \theta^2\big(\|\hat{b}\| + \|\hat{B}\|_{\mathcal{L}(\mathcal{H})} K_{T+\theta}(1 + \|\mathbf{P}_d(x)\|^2)\big), \qquad (6.68)
\end{aligned}$$

where in the last inequality we used (6.57) and that $\tau_d \leq T$ by assumption. Similarly, for the martingale part we first set $Q_m = \int_{\mathcal{H}^+\setminus\{0\}} \|\xi\|^2\, m(\mathrm{d}\xi)$ and find

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}_x^d}\left[\|\bar{J}_{\tau_d+\theta}^d - \bar{J}_{\tau_d}^d\|^2\right] &\leq 4\mathbb{E}_{\mathbb{P}_x^d}\left[\langle \bar{J}^d\rangle_{\tau_d+\theta} - \langle \bar{J}^d\rangle_{\tau_d}\right] \\
&\leq 4\mathbb{E}_{\mathbb{P}_x^d}\left[\int_{\tau_d}^{\tau_d+\theta} \Big(Q_m + \langle X_s^d, \mu(\mathcal{H}^+ \setminus \{0\})\rangle\Big)\,\mathrm{d}s\right] \\
&\leq 4\theta\mathbb{E}_{\mathbb{P}_x^d}\left[\sup_{0\leq\tau\leq\theta} \Big(Q_m + \langle X_{\tau_d+\tau}^d, \mu(\mathcal{H}^+ \setminus \{0\})\rangle\Big)\right] \\
&\leq 4\theta\big(Q_m + \|\mu(\mathcal{H}^+ \setminus \{0\})\| K_{T+\theta}(1 + \|\mathbf{P}_d(x)\|^2)\big). \qquad (6.69)
\end{aligned}$$

By an application of Markov's inequality we thus see that

$$\begin{aligned}
\mathbb{P}_x^d\big(\|X_{\tau_d+\theta}^d - X_{\tau_d}^d\| > \eta\big) &\leq \mathbb{P}_x^d\big(\|H_{\tau_d+\theta}^d - H_{\tau_d}^d\| + \|\bar{J}_{\tau_d+\theta}^d - \bar{J}_{\tau_d}^d\| > \eta\big) \\
&\leq \frac{2}{\eta^2}\big(\mathbb{E}_{\mathbb{P}_x^d}\left[\|H_{\tau_d+\theta}^d - H_{\tau_d}^d\|^2\right] + \mathbb{E}_{\mathbb{P}_x^d}\left[\|\bar{J}_{\tau_d+\theta}^d - \bar{J}_{\tau_d}^d\|^2\right]\big),
\end{aligned}$$

and therefore by inserting (6.68) and (6.69) we obtain

$$\mathbb{P}_x^d\big(\|X_{\tau_d+\theta}^d - X_{\tau_d}^d\| > \eta\big) \leq \theta\frac{\hat{K}_{T+\theta}}{\eta^2}(1 + \|\mathbf{P}_d(x)\|^2), \qquad (6.70)$$

for a $\hat{K}_{T+\theta}$ which is independent of $d \in \mathbb{N}$ and continuous in θ.

Moreover, since $\|\mathbf{P}_d(x)\| \le \|x\|$ for all $d \in \mathbb{N}$, we find a $\delta > 0$ small enough such that

$$\sup_{d \ge N_0} \sup_{0 \le \theta \le \delta} \mathbb{P}_x^d\big(\|X^d_{\tau_d+\theta} - X^d_{\tau_d}\| > \eta\big) \le \delta \frac{\hat{K}_{T+\delta}}{\eta^2}(1 + \|x\|^2) \le \varepsilon,$$

for arbitrary $N_0 \in \mathbb{N}$. This proves the second condition above and it therefore follows from the Aldous criterion that the sequence $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ is a tight sequence of probability measures on $\mathcal{B}(D(\mathbb{R}^+, \mathcal{H}^+))$. □

6.8.2 Weak convergence of the finite-rank operator-valued affine processes

In this section we prove weak convergence of the sequence $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ of laws of $(X^d)_{d\in\mathbb{N}}$ given $X_0^d = \mathbf{P}_d(x)$ to a unique affine process X with law $\mathbb{P}_x$. By Proposition 6.25 we already know that $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ is tight, which by the Prokhorov characterization of relative weak compactness, implies that every subsequence of $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ admits a weakly convergent subsequence. If we show that all those convergent subsequences have the same limit $\mathbb{P}_x$, we can conclude that already $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ converges weakly to $\mathbb{P}_x$, see also [56, Chapter 3]. We are thus left with proving *uniqueness*, which we approach via *martingale problems*, see [56, Chapter 4]. Recall that for every $d \in \mathbb{N}$ the process in (6.6) is a martingale on $(\Omega, \bar{\mathcal{F}}^d, \bar{\mathbb{F}}^d, \mathbb{P}_x)$, in which case we say that X^d, respectively its law $\mathbb{P}_x^d$, solves the *martingale problem* for $\mathcal{G}^d$ with initial condition $X_0^d = \mathbf{P}_d(x)$. Next, we formulate a martingale problem for the operator $\mathcal{G}$ defined on the set $\big\{\mathrm{e}^{-\langle\cdot,u\rangle} : u \in \mathcal{H}^+\big\}$ as

$$\mathcal{G}\,\mathrm{e}^{-\langle\cdot,u\rangle}(x) := \big(F(u) + \langle x, R(u)\rangle\big)\,\mathrm{e}^{-\langle x,u\rangle}, \quad x \in \mathcal{H}^+. \tag{6.71}$$

Definition 6.26. Let $\Omega = D([0,T], \mathcal{H}^+)$, $\mathbb{P}$ be a probability measure on $\mathcal{B}(\Omega)$ with canonical process $(X_t)_{t\ge0}$. Let $\mathcal{G}$ be as in (6.71) defined on $\mathcal{D}$ and let $x \in \mathcal{H}^+$. We then call $\mathbb{P}$ a solution to the martingale problem for $\mathcal{G}$ with initial condition $\mathbb{P}(X_0 = x) = 1$ if for every $f \in \mathcal{D}$ the process

$$\left(f(X_t) - f(x) - \int_0^t \mathcal{G}f(X_s)\,\mathrm{d}s\right)_{t\ge0}, \tag{6.72}$$

is a martingale on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\ge0}$ denotes the natural filtration of $(X_t)_{t\ge0}$.

That the martingale problem has at least one solution is the assertion of the following proposition.

Proposition 6.27. *Let $x \in \mathcal{H}^+$. Then every weak limit $\mathbb{P}_x$ of a convergent subsequence of $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ solves the martingale problem posed in Definition 6.26. Moreover, the canonical process of $\mathbb{P}_x$ on $D(\mathbb{R}^+, \mathcal{H}^+)$ is continuous in probability.*

Proof. From the tightness of $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ the existence of a weakly convergent subsequence follows from the Prokhorov theorem. Let $\mathbb{P}_x$ be such a weak limit of some subsequence $(\mathbb{P}_x^{d_n})_{n\in\mathbb{N}}$. We know from Theorem 6.2, that for every $d\in\mathbb{N}$ the process X^d, respectively its law $\mathbb{P}_x^d$, solves the martingale problem for $\mathcal{G}^d$ with $X_0^d=\mathbf{P}_d(x)$, in particular this holds for all $(\mathbb{P}_x^{d_n})_{n\in\mathbb{N}}$. Now, note that for every $u\in\mathcal{H}^+$ we have

$$\begin{aligned}\sup_{x\in\mathcal{H}^+}|\,\mathrm{e}^{-\langle x,\mathbf{P}_d(u)\rangle}-\mathrm{e}^{-\langle x,u\rangle}|&\leq\sup_{x\in\mathcal{H}^+}\mathrm{e}^{-\langle x,u\rangle}\langle x,u-\mathbf{P}_d(u)\rangle\\&\leq\sup_{x\in\mathcal{H}^+}\mathrm{e}^{-\langle x,u\rangle}\|x\|\|\mathbf{P}_d^{\perp}(u)\|\to0,\quad\text{as }d\to\infty,\end{aligned}$$

and we also find that

$$\begin{aligned}\sup_{x\in\mathcal{H}^+}|\mathcal{G}^d\,\mathrm{e}^{-\langle\cdot,\mathbf{P}_d(u)\rangle}(x)-\mathcal{G}\,\mathrm{e}^{-\langle\cdot,u\rangle}(x)|&=\sup_{x\in\mathcal{H}^+}\big(|\langle x,R_d(\mathbf{P}_d(u))-R(u)\rangle|\,\mathrm{e}^{-\langle x,u\rangle}\big)\\&\quad+|F_d(\mathbf{P}_d(u))-F(u))|\\&\leq\sup_{x\in\mathcal{H}^+}\mathrm{e}^{-\langle x,u\rangle}\|x\|\|R_d(\mathbf{P}_d(u))-R(u)\|\\&\quad+|F_d(\mathbf{P}_d(u))-F(u)|\to0,\quad\text{as }d\to\infty,\end{aligned}$$

where the latter limit holds true as $\sup_{x\in\mathcal{H}^+}\mathrm{e}^{-\langle x,u\rangle}\|x\|$ is bounded, F and R are continuous on $\mathcal{H}^+$, see Lemma 6.11, and $\|\mathbf{P}_d^{\perp}(u)\|=\|\mathbf{P}_d(u)-u\|\to0$ as $d\to\infty$. It thus follows from [56, Lemma 5.1] that the weak limit $\mathbb{P}_x$ of $(\mathbb{P}_x^{d_n})_{n\in\mathbb{N}}$ solves the martingale problem in Definition 6.26. The continuity in probability of the canonical process of $\mathbb{P}_x$ is a consequence of Aldous criterion and follows from [87, Theorem 3.3.1]. ☐

Next, we prove that the limit $\mathbb{P}_x$ of every convergent subsequence of $(\mathbb{P}_x^d)_{d\in\mathbb{N}}$ is unique, which in turn proves Theorem 6.4 i) and we also prove the remaining assertions of Theorem 6.4.

Proof of Theorem 6.4. For $x\in\mathcal{H}^+$, denote by $\mathbb{P}_x$ the limit of some subsequence of $(\mathbb{P}_x^{d_n})_{n\in\mathbb{N}}$ and let $X=(X_t)_{t\geq0}$ the canonical process of $\mathbb{P}_x$ on $\Omega=D(\mathbb{R}^+,\mathcal{H}^+)$. By Proposition 6.27 $(X,\mathbb{P}_x)$ is a solution to the martingale problem posed in Definition 6.26. Moreover, let $T\geq0$ arbitrary, $u\in\mathcal{H}^+$ and define the functions $f_u(t,x)\colon[0,T]\times\mathcal{H}^+\to\mathbb{R}^+$ by $f_u(t,x)=\mathrm{e}^{-\phi(T-t,u)-\langle x,\psi(T-t,u)\rangle}$, where $(\phi(\cdot,u),\psi(\cdot,u))$ is the unique solution of (2.8) on $[0,T]$. We see that $f_u\in C_b^{1,1}([0,T]\times\mathcal{H}^+)$ and it thus follows from [56, Theorem 4.7.1] that since X solves the martingale problem for $\mathcal{G}$, the process $(X_t,t)_{t\geq0}$ solves the associated time-dependent martingale problem, i.e. the process

$$\Big(f_u(t,X_t)-f_u(0,x)-\int_0^t\mathcal{G}f_u(s,X_s)+\frac{\partial}{\partial s}f_u(s,X_s)\,\mathrm{d}s\Big)_{0\leq t\leq T}\tag{6.73}$$

is a martingale for every $u\in\mathcal{H}^+$.

Moreover, we see that

$$\begin{aligned}\frac{\partial}{\partial t} f_u(t,x) &= \Big(\frac{\partial \phi}{\partial t}(T-t,u) + \langle x, \frac{\partial \psi}{\partial t}(T-t,u)\rangle\Big) f_u(t,x)\\ &= \left(F(\psi(T-t,u)) + \langle x, R(\psi(T-t,u))\rangle\right) f_u(t,x)\,, \end{aligned} \tag{6.74}$$

which inserted into (6.73) nullifies the term $\mathcal{G} f_u(s, X_s)$, compare with (6.71). Hence, we see that the process $(f_u(t, X_t) - f_u(0, x))_{t \le T}$ must be a martingale. This implies in particular, that $\mathbb{E}_{\mathbb{P}_x}[f_u(t, X_t)] = f_u(0, x)$ for all $0 \le t \le T$, i.e. for $t = T$ we obtain

$$\mathbb{E}_{\mathbb{P}_x}[f_u(T, X_T)] = \mathbb{E}_{\mathbb{P}_x}\left[\mathrm{e}^{\langle X_T, u\rangle}\right] = \mathrm{e}^{-\phi(T,u) - \langle x, \psi(T,u)\rangle}, \quad u \in \mathcal{H}^+.$$

Since $T > 0$ was arbitrary, this implies that the affine transform formula (6.8) holds true. Note, that since this holds for every $u \in \mathcal{H}^+$ and the Laplace transform is measure determining on $\mathcal{H}^+$, see Lemma 3.19, it follows that X_t is unique in law for every fixed $t \ge 0$. But, since the process X is the solution to the martingale problem in Definition 6.26, it follows from [56, Theorem 4.4.2 (a)] that the pointwise uniqueness already implies the uniqueness in distribution (i.e. uniqueness of the solution to the martingale problem on $[0, T]$). Again, since T was arbitrary, this then proves Theorem 6.4 i).
Next we show Theorem 6.4 ii). Note that in the first part, we just proved that the limit of convergent subsequences of $(\mathbb{P}^d_x)_{d\in\mathbb{N}}$ is given by $\mathbb{P}_x$ and that the associated process X satisfies the affine transform formula, i.e. the sequence $(X^d)_{d\in\mathbb{N}}$ converges weakly to X on $D(\mathbb{R}^+, \mathcal{H}^+)$. We can thus continue with the convergence rate in (6.9). First, note that by standard estimates and (6.8), (6.5) we obtain

$$\begin{aligned}\Big|\mathbb{E}_{\mathbb{P}_x}\left[\mathrm{e}^{-\langle u, X_t\rangle}\right] - \mathbb{E}_{\mathbb{P}^d_x}\left[\mathrm{e}^{-\langle u, X^d_t\rangle}\right]\Big| &\le |\phi(t,u) - \phi_d(t, \mathbf{P}_d(u))|\\ &\quad + \|x\|\|\psi(t,u) - \psi_d(t, \mathbf{P}_d(u))\|,\end{aligned}$$

and it thus follows from Corollary 6.12 that there exists a C_T independent of $d \in \mathbb{N}$, such that

$$\sup_{t\in[0,T], \|u\|_{\mathcal{V}} \le 1} \Big|\mathbb{E}_{\mathbb{P}_x}\left[\mathrm{e}^{-\langle u, X_t\rangle}\right] - \mathbb{E}_{\mathbb{P}^d_x}\left[\mathrm{e}^{-\langle u, X^d_t\rangle}\right]\Big| \le C_T \|\mathbf{P}^\perp_d\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}(1 + \|x\|), \tag{6.75}$$

Note that if the conditions $\|\mu(\mathcal{H}^+ \setminus \{0\})\|_{\mathcal{V}} < \infty$ and $B^*(\mathcal{V}_0) \subseteq \mathcal{V}_0$ do not hold then we see that the convergence rate in part i) of Remark 6.5 follows from Proposition 6.1 instead of Corollary 6.12. □

6.9 Proof: Finite-rank approximation of affine stochastic covariance models

We begin with the proof of Proposition 6.7 i). Let $x \in \mathcal{H}^+$ and denote by X the square-integrable Markovian semimartingale associated with the admissible parameters (b, B, m, μ) given by Theorem 6.4. Moreover, let us denote the stochastic basis on which X is defined by $(\Omega^1, \bar{\mathcal{F}}^1, \bar{\mathbb{F}}^1, \mathbb{P}_x)$ and let $(\Omega^2, \mathcal{F}^2, (\mathcal{F}^2_t)_{t\geq 0}, \mathbb{P}^2)$ be another filtered probability space, which satisfies the usual conditions and carries a cylindrical Brownian motion $W\colon [0,\infty) \times \Omega^2 \to H$. We set

$$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}) \coloneqq (\Omega^1 \times \Omega^2, (\bar{\mathcal{F}}^1 \otimes \mathcal{F}^2), (\bar{\mathcal{F}}^1_t \otimes \mathcal{F}^2_t)_{t\geq 0}, \mathbb{P}_x \otimes \mathbb{P}^2)\,,$$

denote the expectation with respect to $\mathbb{P}$ by $\mathbb{E}$ and consider X and W as processes on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Now, it was already shown in Lemma 3.7 and Theorem 3.14 that the model (Y, X) is well-defined and satisfies the affine transform formula (3.27). In addition, note that by assumption $(\mathcal{A}, \operatorname{dom}(\mathcal{A}))$ generates a strongly continuous semigroup on V and $D^{1/2} \in \mathcal{L}_2(H, V)$, which implies that for every initial value $Y_0 = y \in V$ the solution $(Y_t)_{t\geq 0}$ exists in V, i.e. $\mathbb{P}(\{Y_t \in V\colon\ t \geq 0\}) = 1$, see e.g. [48]. Now, let $d \in \mathbb{N}$ and consider $(Y^d_t, X^d_t)_{t\geq 0}$ with X^d as given by Proposition 6.2 where in particular $X^d_0 = \mathbf{P}_d(x)$.
It follows from Proposition 6.23 that the process X^d is a square-integrable affine Markovian semimartingale on some stochastic basis $(\Omega^d, \bar{\mathcal{F}}^d, \bar{\mathbb{F}}^d, \mathbb{P}^d_x)$ and from the representation (6.54) we see that the process $(Y^d_t, X^d_t)_{t\geq 0}$ satisfies

$$\mathrm{d}\begin{bmatrix} Y^d_t \\ X^d_t \end{bmatrix} = \left(\begin{bmatrix} 0 \\ \hat{b}_d \end{bmatrix} + \begin{bmatrix} \mathcal{A}(Y^d_t) \\ \hat{B}_d(X^d_t) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} D^{1/2}(X^d_t)^{1/2} & 0 \\ 0 & 0 \end{bmatrix} \mathrm{d} \begin{bmatrix} W_t \\ 0 \end{bmatrix} + \mathrm{d} \begin{bmatrix} 0 \\ \bar{J}^d_t \end{bmatrix}, \tag{6.76}$$

with $(Y^d_0, X^d_0) = (y, \mathbf{P}_d(x)) \in H \times \mathcal{H}^+$, $\hat{b}_d = b_d + \int_{\mathcal{H}^+_d \cap \{\|\xi\|>1\}} \xi \, m_d(\mathrm{d}\xi)$ and

$$\hat{B}(u) = B_d(\mathbf{P}_d(u)) + \int_{\mathcal{H}^+_d \cap \{\|\xi\|>1\}} \xi \, \langle u, M_d(\mathrm{d}\xi)\rangle, \quad \forall\, u \in \mathcal{H}.$$

Thus, the model is of the form in Definition 3.8 and Remark 3.10. Therefore, Proposition 6.7 i) follows from immediately from Theorem 3.14. We continue with Proposition 6.7 ii). Note that for every $d \in \mathbb{N}$ the process Y^d is the mild solution of the first component in (6.76) given by

$$Y^d_t = S(t)y + \int_0^t S(t-s) D^{1/2} (X^d_s)^{1/2} \, \mathrm{d}W_s, \quad t \geq 0,$$

where the stochastic integral with respect to the cylindrical Brownian motion is well-defined, see Lemma 3.7. From Theorem 6.4 ii) we already know that X^d converges weakly to X on $D(\mathbb{R}^+, \mathcal{H}^+)$ as d tends to infinity.

From this and [102, Theorem 4.2] we conclude the weak convergence of the stochastic integrals, i.e.

$$\Big(\int_0^t S(t-s)D^{1/2}(X_s^d)^{1/2}\,\mathrm{d}W_s\Big)_{t\geq 0} \overset{d\to\infty}{\Rightarrow} \Big(\int_0^t S(t-s)D^{1/2}(X_s)^{1/2}\,\mathrm{d}W_s\Big)_{t\geq 0},$$

which in turn implies that also the joint processes $(Y^d, X^d)_{d\in\mathbb{N}}$ converges weakly to (Y, X) as d tends to infinity.

We now come to the convergence rate in (6.22). Let $T \geq 0$, $\tilde{u} \in H$ and $u = (\mathrm{i}\tilde{u}, 0)$. Note that for every $M \geq 0$ we find a local Lipschitz constant of the map $u \mapsto \tilde{R}(h, u)$ on $\{u \in \mathcal{H}^+ : \|u\| \leq M\}$ independent of $h \in H$ and $d \in \mathbb{N}$, when we replace R by $\mathbf{P}_d \circ R \circ \mathbf{P}_d$, compare this with Lemma 6.11. Again by a similar argument as in the proof of Proposition 6.1, we conclude that for every $M \geq 0$ there exists a $\tilde{H}_M > 0$, independent of $d \in \mathbb{N}$, such that $\|\psi_2(t,u)\| \vee \|\psi_{2,d}(t,u)\| \leq \tilde{H}_M$ for all $t \in [0,T]$ and all $\tilde{u} \in H$ with $\|\tilde{u}\|_H \leq M$. Note, that in contrast to the proof of Proposition 6.1 we assume $\psi_2(0,u) = 0$, but have the additional inhomogeneous-term $-\frac{1}{2}(D^{1/2}\psi_1(t,u) \otimes D^{1/2}\psi_1(t,u))$ where $\psi_1(\cdot,u)$ solves (3.24b), i.e. it is given by $\psi_1(t,u) = \mathrm{i}S(t)\tilde{u}$. This implies that the $\tilde{H}_M$ depends on $\tilde{u}$ and is in general different to the H_M we used before.

So let T, M and $\tilde{H}_M$ as above. Then Inserting the affine-transform formulas (3.27) and (6.21) of (Y, X) and (Y^d, X^d), respectively, and using standard estimations we obtain

$$\sup_{t\in[0,T]}\Big|\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\langle Y_t,\tilde{u}\rangle_H}\Big]-\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\langle Y_t^d,\tilde{u}\rangle_H}\Big]\Big| \leq \sup_{t\in[0,T]}\big(|\Phi(t,u)-\Phi_d(t,u)| + \|x\|\|\psi_2(t,u)-\psi_{2,d}(t,u)\|\big). \tag{6.77}$$

The terms on the right-hand side of (6.77) can be estimated similarly to what we saw in the proof of Proposition 6.1. Indeed, for $u \in \mathcal{H}^+$ and $s \geq 0$ we define the function $\tilde{R}_d(s,u) \coloneqq \mathbf{P}_d(R(\mathbf{P}_d(\psi_{2,d}(s,u)))) - \frac{1}{2}(D\psi_1(s,u))^{\otimes 2}$ and we see that

$$\begin{aligned}\|\mathbf{P}_d(\psi_2(t,u)) - \psi_{2,d}(t,u)\| &\leq \int_0^t \|\mathbf{P}_d\big(\tilde{R}(\psi_1(s,u),\psi_2(s,u))\big) - \tilde{R}_d(s,u)\|\,\mathrm{d}s\\ &\leq \int_0^t \|R(\psi_2(s,u)) - R(\mathbf{P}_d(\psi_{2,d}(s,u)))\|\,\mathrm{d}s\\ &\leq \tilde{L}_M^{(1)}\int_0^t \|\psi_2(s,u)-\psi_{2,d}(s,u)\|\,\mathrm{d}s,\end{aligned}$$

where $\tilde{L}_M^{(1)}$ denotes a Lipschitz constant of R on the set $\big\{u \in \mathcal{H}^+ : \|u\| \leq \tilde{H}_M\big\}$ given by Lemma 6.11. Since $u_2 = 0$, we see that an application of the variation-of-constant formula yields

$$\psi_2(t,u) = \int_0^t \Big(\mathrm{e}^{(t-s)B^*}K_{\psi_2(s,u)}(\xi)\frac{\mu(\mathrm{d}\xi)}{\|\xi\|^2} - \frac{1}{2}\,\mathrm{e}^{(t-s)B^*}(D^{1/2}\psi_1(s,u))^{\otimes 2}\Big)\,\mathrm{d}s.$$

Thus by inserting $\psi_1(t,u) = \mathrm{i}S(t)\tilde{u}$ we see that

$$\begin{aligned}\|\mathbf{P}_d^{\perp}(\psi_2(t,u))\| \leq \tilde{H}_M^2 t \sup_{s\in[0,t]} \|\mathbf{P}_d^{\perp}\,\mathrm{e}^{sB^*}\mu(\mathcal{H}^+\setminus\{0\}))\| \\ + \frac{1}{2}t \sup_{s\in[0,t]} \|\mathbf{P}_d^{\perp}\,\mathrm{e}^{(t-s)B^*}(D^{1/2}S(s)\tilde{u})^{\otimes 2}\|. \end{aligned} \tag{6.78}$$

Note that since $\tilde{u}\in V$ and $D^{1/2}S(s)(V)\subseteq V$ for every $s\geq 0$ by assumption, we have $\mathrm{e}^{(t-s)B^*}(D^{1/2}S(s)\tilde{u})^{\otimes 2}\in\mathcal{V}_0$ for every $s\leq t$, and we therefore conclude that

$$\sup_{s\in[0,t]} \|\mathbf{P}_d^{\perp}(\mathrm{e}^{(t-s)B^*}D^{1/2}S(s)\tilde{u})^{\otimes 2})\| \leq \|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}\,\mathrm{e}^{t\|B\|_{\mathcal{L}(\mathcal{H})}}\|D^{1/2}\|\|\tilde{u}\|_V^2 M_1^2\,\mathrm{e}^{2tw},$$

where we used that $\|D^{1/2}S(s)\tilde{u}\|_V^2 \leq \|D^{1/2}\|\|\tilde{u}\|_V^2 M_1^2\,\mathrm{e}^{2tw}$ for the type $M_1\geq 0$ and growth bound $w\in\mathbb{R}$ of the semigroup $(S(t))_{t\geq 0}$ on V, i.e. $\|S(t)\|_{\mathcal{L}(V)}\leq M_1\,\mathrm{e}^{tw}$ for $t\geq 0$. Hence, we conclude as in the proof of Proposition 6.1 that for $\check{K}_t^d$ with

$$\check{K}_t^d = \big(\tilde{H}_M^2 t\,\mathrm{e}^{t\|B\|_{\mathcal{L}(\mathcal{H})}}\|\mu(\mathcal{H}^+\setminus\{0\})\|_{\mathcal{V}} + \mathrm{e}^{t\|B\|_{\mathcal{L}(\mathcal{H})}}\|D^{1/2}\|M_1^2\,\mathrm{e}^{2tw}\big)\|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})},$$

we can bound the left-hand side in (6.77) by $\tilde{C}_T\|\mathbf{P}_d^{\perp}\|_{\mathcal{L}(\mathcal{V},\mathcal{H})}(1+\|x\|)$ with

$$\tilde{C}_T \coloneqq \mathrm{e}^{\tilde{L}_M^{(1)}T}\big(1+\tilde{L}_M^{(2)}T\big)(\mathrm{e}^{t\|B\|_{\mathcal{L}(\mathcal{H})}}\max(\tilde{H}_M^2 T\|\mu(\mathcal{H}^+\setminus\{0\})\|_{\mathcal{V}}, \|D^{1/2}\|M_1^2\,\mathrm{e}^{Tw}),$$

which proves the asserted convergence rate (6.22).

6.10 Concluding remarks

In addition to the discussed relevance in applications, we also want to highlight some theoretical aspects of this work. In particular, we discuss how this chapter contributes to the general understanding of affine processes in infinite-dimensions.

- **On the connection of finite- and infinite-rank affine processes**

In our existence proof for finite-rank operator-valued affine processes in Section 6.7, we essentially used an existence result for matrix-valued affine processes from [42], see also the proof of Proposition 6.19. The sequence $(X^d)_{d\in\mathbb{N}}$ of *finite-rank operator-valued affine processes* arises from an isomorphic transformation of a corresponding sequence of matrix-valued affine processes $(\tilde{X}^d)_{d\in\mathbb{N}}$. Consequently, we observe that $(X^d)_{d\in\mathbb{N}}$ shares many properties with their matrix-valued versions, such as the existence of càdlàg paths and that only jumps of finite-variation are admissible. We then embedded the sequence $(X^d)_{d\in\mathbb{N}}$ into the space of Hilbert-Schmidt operator-valued càdlàg functions to study the limit as the rank $d\to\infty$.

We observed that the affine transform formula and the càdlàg path property is preserved in the limit, but we also saw that the limit does not have to be of finite-variation anymore. We traced this behavior back to the difference of *Pettis* and *Bochner integrability* of the truncation function χ with respect to the compensator of the affine jump-measures.

- **On the diffusion component**

The idea to prove the existence of affine processes on infinite-dimensional state-spaces using finite-dimensional approximations is, to the best of my knowledge, novel. Finite-dimensional approximations of affine diffusion in Hilbert spaces were discussed in [147]. However, the approximation was not used for proving the existence of affine diffusion's and also no explicit convergence rates for the Laplace transforms of the processes or its associated generalized Riccati equations were established. Similarly to what we have done in Chapter 2, our approach in this chapter does not admit for any diffusion components. Indeed, note that a diffusion parameter α_d in a matrix-valued admissible parameter set has to satisfy $b_d \geq_{\mathbb{S}_d^+} (d-1)\alpha_d$, see [42, Definition 2.3]. But, since we let $d \to \infty$, this can not hold unless $b_d \to \infty$ or $\alpha_d \to 0$ as $d \to \infty$. But since $b_d \to b$, we must have $\alpha_d \to 0$, which means that any diffusion part vanishes in the limit case.

- **On Galerkin approximations of Riccati equations**

We also want to highlight our contribution to the theory of Galerkin approximation of operator-valued generalized Riccati equations. Note that even though Galerkin approximations of Riccati equations on Hilbert-Schmidt operators are well-studied in the literature, see, e.g. [128], as they naturally appear in stochastic control and filtering theory, their study is often limited to equations with monotone non-linear components such as the classic quadratic functions. In the present chapter we study Galerkin type approximations of generalized operator-valued Riccati equations with non-monotone non-linear components that are given by integrals of Lévy-Khintchine type with respect to (vector-valued) measures. For such equations we provided a thorough study of finite-rank approximation and derived explicit convergence rates for the Galerkin approximation in terms of admissible model parameters.

Part II

Multivariate Stochastic Covariance Models based on MCARMA Processes

Introduction to Part II: Multivariate Stochastic Covariance Models based on MCARMA Processes

Multivariate continuous-time autoregressive moving-average (MCARMA) processes are the continuous-time versions of the classical discrete-time VARMA models and have been studied thoroughly over the last two decades [112, 133, 30]. Similarly to their univariate analogs, the CARMA processes, MCARMA processes can be interpreted as the solutions of a higher-order SDE of the form

$$\mathrm{D}^p X_t + \tilde{A}_1 \mathrm{D}^{p-1} X_t + \ldots + \tilde{A}_p X_t = \tilde{C}_0 \mathrm{D}^{q+1} L_t + \tilde{C}_1 \mathrm{D}^q L_t + \ldots + \tilde{C}_q \mathrm{D} L_t, \quad (6.79)$$

where $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}$, $(\tilde{A}_i)_{i=1,\ldots,p}$ and $(\tilde{C}_j)_{j=0,\ldots,q}$ for $q, p \in \mathbb{N}$ are two families of linear operators and $(L_t)_{\in\mathbb{R}}$ denotes a multivariate Lévy process. Naturally, equation (6.79) asks for a rigorous definition as the paths of Lévy processes are in general not differentiable. Heuristically, however, we can interpret (6.79) as the continuous-time version of a (V)ARMA difference equation, and we therefore expect that similar key features governed by the autoregressive and moving-average structure of the defining equation find their counterpart in the continuous-time setting. The most notable feature of the (M)CARMA class is its flexible *short memory* structure. In general, short memory refers to an exponentially fast decaying auto-covariance function. In fact, the MCARMA class exhibits much more nuanced auto-covariance behavior in the short-time lags and allows, e.g., for non-monotone or sub-exponentially decaying configurations (although no polynomial or linear decaying ones, which processes with *long memory* do admit). The memory effect is observed in many time series in applications and explains the popularity of modeling with (M)CARMA processes in subjects ranging from finance over meteorology to natural science and engineering, see e.g. [88, 118, 140, 25].

In many applications, where (M)CARMA models are employed, a crucial model feature is positivity, e.g. in modeling wind speed [25, 16], the velocity field in turbulence [7] or, most notably, volatility in finance [9, 140, 27, 12]. It is therefore of great importance to understand the capability of (M)CARMA processes to model phenomena with positive states. In the univariate CARMA case positive processes were studied in [142, 141, 29, 27, 20, 119]. In particular, the authors in [142] give a set of necessary and/or sufficient parameter conditions for CARMA processes of general order driven by a Lévy subordinator to be non-negative.

In this second part of the thesis our goal is to introduce a class of matrix-valued MCARMA processes for which we can establish conditions that ensure that the processes assume values in multivariate cones. The particularly interesting cases of the closed positive orthant $\mathbb{R}_d^+$ and the cone of positive semi-definite matrices $\mathbb{S}_d^+$ are included in our analysis. In fact, due to their relevance in applications these two main examples coin our terminology of *positive multivariate CARMA*.

CARMA processes on multivariate cones

Chapter 7 provides the mathematical foundations for matrix-valued CARMA processes with values in multivariate convex cones. First, we introduce the class of matrix-valued MCARMA that extends classical $\mathbb{R}_d$-valued MCARMA processes to a more general multivariate framework. In particular, this novel class includes MCARMA processes on $\mathbb{S}_d$, the space of symmetric $d \times d$-matrices. We define matrix-valued MCARMA processes as the stationary solutions to specific continuous-time linear state space models on matrix-spaces driven by a matrix-valued Lévy process. More precisely, we show that a matrix-valued MCARMA process $(X_t)_{t\in\mathbb{R}}$ can be represented as

$$X_t = \int_{-\infty}^{+\infty} g(t-s)\, \mathrm{d}L_s, \quad t \in \mathbb{R}, \tag{6.80}$$

where $(L_t)_{t\in\mathbb{R}}$ denotes the matrix-valued (two-sided) Lévy process and the kernel $t \mapsto g(t)$ takes values in the space of linear operators on matrices. This approach, to define the MCARMA class, is similar to the classical case in [112] and, as a matter of fact, we show that matrix-valued MCARMA processes can be viewed as classical MCARMA under vectorization.

In the second part of Chapter 7 we are concerned with MCARMA processes with values in convex cones. More precisely, we establish necessary and sufficient parameter conditions for matrix-valued MCARMA processes to assume values in a convex cone, whenever the driving Lévy process is increasing with respect to that cone. From the representation (6.80) it can be seen, that this is the case whenever $t \mapsto g(t)$ maps into a set of operators leaving the respective cone invariant. We present particularly hands-on conditions for the kernel g in cases of the cones $\mathbb{R}_d^+$ and $\mathbb{S}_d^+$. Note here, that the vectorization of a positive semi-definite matrix does in general not yield a non-ngeative vector, so that specific conditions on the kernel g in the two cases $\mathbb{S}_d^+$ and $\mathbb{R}_d^+$ have to be studied case-by-case.

Stochastic covariance models based on positive semi-definite MCARMA

In Chapter 8 we propose to model the instantaneous covariance process in multivariate stochastic covariance models by (higher-order) MCARMA processes with values in the cone of positive semi-definite matrices. Indeed, this class extends the multivariate Barndorff–Nielsen-Shepard model (first-order MCARMA based model) to higher-order MCARMA models, that admit a more nuanced auto-covariance structure and have the potential to introduce certain short-memory effects to the model class. We demonstrate the capability of higher-order MCARMA based models by an exemplary analysis of stochastic covariance models based on positive semi-definite well-balanced Ornstein-Uhlenbeck processes.

CHAPTER 7

Positive Multivariate Continuous-Time Autoregressive Moving-Average Processes

Abstract of the chapter In this chapter, we study multivariate continuous-time autoregressive moving-average (MCARMA) processes with values in cones. More precisely, we introduce matrix-valued MCARMA processes defined as stationary solutions of specific continuous-time linear state-space models on matrix-spaces driven by Lévy processes. Moreover, we establish necessary and sufficient conditions for processes from this class to stay in a multivariate convex cone. We derive specific hands-on conditions in the following two cases: First, for classical MCARMA on $\mathbb{R}_d$ with values in the positive orthant $\mathbb{R}_d^+$. Second, for MCARMA processes on real square matrices taking values in the cone of positive semi-definite matrices. To illustrate our positivity criteria, we give several concrete parameter specifications ensuring the positivity of the respective MCARMA process. Positive semi-definite MCARMA processes are relevant for applications in multivariate stochastic covariance modeling and have the potential to model a variety of short memory features observed in realized (cross)-covariance processes.

This chapter is based on [17, Section 1, 2 and 3]:
Benth, F., and Karbach, S.
Positive multivariate continuous-time autoregressive moving-average processes, 2022, DOI: arXiv.2206.08782.

7.1 Introduction

The starting point of our analysis will be the formulation of linear continuous-time state space models on the space of real $n \times m$-matrices. It is well known that $\mathbb{R}_d$-valued MCARMA processes are characterized by certain configurations of continuous-time state space models, see e.g. [112, 30, 133]. Linear continuous-time state space models are essentially given by an Ornstein-Uhlenbeck (OU) type process on the Cartesian product of the particular state space, in our case the space of all $n \times m$-matrices, and a linear output operator mapping the values of the higher-dimensional Ornstein-Uhlenbeck type process into the state space again. We show that for certain specifications of such models the vectorization of stationary output processes are equivalent to MCARMA processes, as they were introduced in [112]. In this sense, the matrix-valued state space models give rise to the novel class of matrix-valued MCARMA processes. The details are given in Section 7.2 below. Since in some applications one is interested in non-stable systems, we explicitly include the class of non-stable state space models in our analysis. Non-stable models often correspond to so called non-causal MCARMA processes, i.e. MCARMA processes that are not adapted to the natural filtration. A careful distinction between the stable and non-stable case is justified, since the stability conditions may interact with the imposed positivity constraints.

Once we set up the class of matrix-valued MCARMA processes, we study their positivity. In particular, we are interested in necessary and sufficient parameter conditions such that a matrix-valued MCARMA process driven by a multivariate cone-valued Lévy process takes values solely in this cone. As noted above, in the univariate case the positivity of CARMA processes is well-studied and the relevance for applications is widely recognized. In the multivariate setting, however, positivity of MCARMA processes has not yet been studied in a systematical way. Partial results exist in the recent work [119], where the authors derive conditions ensuring the positivity of (univariate) CARMA processes. The authors made the claim that some parts of their results could be extended to the multivariate CARMA case, however, only little information about this extension are provided. Our Theorem 7.24 below supports this claim to some extent.

Note that also in a discrete-time setting studying autoregressive matrix-valued models is an active area of research, see e.g. [35]. Many articles on MCARMA processes deal with their connection to the discrete-time setting, e.g. studying high-frequency sampling for MCARMA in [59, 98] or parameter estimation of the driving noise from discrete observations in [133]. Our findings on the equivalence of the vectorized matrix-valued MCARMA and the classical MCARMA suggests that analogous connections to the discrete-time setting continue to hold in the matrix case and the good accessibility of this class is maintained accordingly.

Our main interest in studying positive semi-definite MCARMA processes comes from stochastic covariance modeling which will be our main concern in Chapter 8.

7.1.1 Layout of the chapter

This chapter is structured as follows: In Section 7.2 we introduce a class of continuous-time state space models on matrices and show the equivalence with the classical MCARMA class under vectorization. In Section 7.3 we study the cone invariance of matrix- and vector-valued MCARMA processes, with a particular focus on the closed positive orthant in $\mathbb{R}_d$ and the cone of symmetric positive semi-definite matrices. Lastly, in Section 7.5 we present an auxiliary result on a submultiplicativity property of the Hadamard product.

7.1.2 Notation and preliminaries

By $\mathbb{N}$ we denote the set of all integers and we set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For $d \in \mathbb{N}$, we denote by $\mathbb{R}_d$ the d-dimensional Euclidean space equipped with the standard inner-product $(\cdot,\cdot)_d$. The closed positive orthant in $\mathbb{R}_d$ will be denoted by $\mathbb{R}_d^+$ and the standard basis of $\mathbb{R}_d$ is denoted by $\{e_1, e_2, \dots, e_d\}$.

Matrices Let $n, m \in \mathbb{N}$ and let $\mathbb{K}$ denote either a field or a ring. Then we denote by $\mathbb{M}_{n,m}(\mathbb{K})$ the set of all $n \times m$ matrices over $\mathbb{K}$. If $n = m$ we write $\mathbb{M}_n(\mathbb{K})$ and if $\mathbb{K} = \mathbb{R}$ we simplify to $\mathbb{M}_{n,m}$. If $m = 1$ we have $\mathbb{M}_{n,1} = \mathbb{R}_n$ and use the latter notation. The p-times Cartesian product of $\mathbb{M}_{n,m}$ will be denoted by $(\mathbb{M}_{n,m})^p$, which is just equivalent to $\mathbb{M}_{pn,m}$, but we use the former notation as it is more suggestive. In the case where $\mathbb{K} = \mathbb{R}[\lambda]$ is the polynomial ring over $\mathbb{R}$ the set $\mathbb{M}_n(\mathbb{R}[\lambda])$ denotes the space of all matrix polynomials with coefficients in $\mathbb{M}_n$. We refer to [71] for a comprehensive analysis of matrix polynomials. We denote the transpose of a matrix $A \in \mathbb{M}_{n,m}$ by $A^\intercal$, which is an element in $\mathbb{M}_{m,n}$, and write $\mathbb{S}_d$ for the subspace in $\mathbb{M}_n$ consisting of all symmetric $n \times n$ matrices, i.e. all $A \in \mathbb{M}_n$ such that $A^\intercal = A$. As before, the set of all symmetric positive semi-definite $n \times n$-matrices will be denoted by $\mathbb{S}_n^+$, i.e. $\mathbb{S}_n^+ = \{A \in \mathbb{S}_n \colon (Ax, x)_n \geq 0, \forall x \in \mathbb{R}_n\}$. If necessary, we can express real $n \times m$-matrices in a component-wise notation by $A = (a_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m}$ for $a_{i,j} \in \mathbb{R}$. For all $n, m \in \mathbb{N}$ we can identify $\mathbb{M}_{n,m}$ with $\mathbb{R}_{nm}$ through the vectorization operator $\operatorname{vec} \colon \mathbb{M}_{n,m} \to \mathbb{R}_{nm}$ which transforms a matrix into a vector by stacking the columns below each other. Similarly, we denote by $\operatorname{vech} \colon \mathbb{S}_n \to \mathbb{R}_{n(n+1)/2}$ the operator that stacks only the lower triangular part of a symmetric matrix below another. On $\mathbb{M}_{n,m}$ we consider the inner-product $\langle \cdot,\cdot \rangle_{nm}$ given by $\langle A, B \rangle_{nm} = (\operatorname{vec}(A), \operatorname{vec}(B))_{nm}$ and denote the induced norm by $\|\cdot\|_{nm}$. Let $n_1, n_2, m_1, m_2 \in \mathbb{N}$, then for $A \in \mathbb{M}_{n_1,m_1}$ and $B \in \mathbb{M}_{n_2,m_2}$ we denote the Kronecker product of A and B by $A \otimes B \in \mathbb{M}_{n_2 n_1, m_2 m_1}$. We denote the Hadamard product of two matrices $A, B \in \mathbb{M}_{n,m}$ by $A \odot B$ and let $\mathbf{1}_{n,m} = (1_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m}$ stand for the matrix in $\mathbb{M}_{n,m}$ which is equal to one in every component and $\mathbb{0}_{n,m}$ denotes the $n \times m$-zero matrix. If $n = m$ we write $\mathbf{1}_n := \mathbf{1}_{n,n}$, $\mathbb{0}_n := \mathbb{0}_{n,n}$ and denote the identity matrix by $\mathbb{I}_n$.

Linear operators on matrices We denote by $\mathcal{L}(\mathbb{M}_{n_1,m_1}, \mathbb{M}_{n_2,m_2})$ the algebra of all linear operators from $\mathbb{M}_{n_1,m_1}$ to $\mathbb{M}_{n_2,m_2}$. If $n_1 = n_2 = n$ and $m_1 = m_2 = m$ we write $\mathcal{L}(\mathbb{M}_{n,m})$ and if $m = 1$, it is well known that $\mathcal{L}(\mathbb{M}_{n,1}) = \mathcal{L}(\mathbb{R}_n) \simeq \mathbb{M}_n$ and we use the latter notation. If $n, m \in \mathbb{N}$ are greater than one, we will denote elements of $\mathcal{L}(\mathbb{M}_{n,m})$ by bold face letters, e.g. $\mathbf{A} \in \mathcal{L}(\mathbb{M}_{n,m})$ versus $A \in \mathbb{M}_{n,m}$ and we reserve the calligraphic letters, e.g. $\mathcal{A}$, for linear operators mapping from or to $(\mathbb{M}_{n,m})^p$ for some integer $p > 1$. Since in the sequel we will always make sure that there is no confusion regarding the matrix space that we are operating in, we denote the identity operator in $\mathcal{L}(\mathbb{M}_{n,m})$ simply by $\mathbf{I}$ and will only index $\mathbf{I}$, when we speak of the identity in $\mathcal{L}((\mathbb{M}_{n,m})^p)$, in which case we write $\mathbf{I}_p$. For every $\mathbf{A} \in \mathcal{L}(\mathbb{M}_{n,m})$ we denote its spectrum by $\sigma(\mathbf{A})$, which, as we work in finite-dimensions, is just the set of eigenvalues of $\mathbf{A}$. Moreover, we denote the spectral bound of $\mathbf{A}$ by $\tau(\mathbf{A})$, i.e. $\tau(\mathbf{A}) := \max\{\Re(\lambda) \colon \lambda \in \sigma(\mathbf{A})\}$.

Matrix-valued Lévy processes Assume that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space satisfying the usual conditions and let $(L_t)_{t\geq 0}$ be an $\mathbb{R}_{nm}$-valued Lévy process defined on this probability basis, see [132] for a comprehensive analysis of multivariate Lévy processes. Since we can always identify $\mathbb{M}_{n,m}$ with $\mathbb{R}_{nm}$ the class of multivariate Lévy processes easily extends to matrix-valued Lévy processes. We recall that the Lévy characteristic exponent at $z \in \mathbb{M}_{n,m}$ is given by

$$\psi_L(z) = \mathrm{i}\langle \gamma_L, z\rangle_{nm} - \frac{1}{2}\langle Q_1 z, z\rangle_{nm} + \int_{\mathbb{M}_{n,m}} \big(\mathrm{e}^{\mathrm{i}\langle \xi, z\rangle_{nm}} - 1 - \mathrm{i}\langle \chi(\xi), z\rangle_{nm}\big)\nu_L(\mathrm{d}\xi), \tag{7.1}$$

where $\gamma_L \in \mathbb{M}_{n,m}$ denotes the *drift* of L, Q_1 is the *covariance operator* of the continuous part of the Lévy process, $\chi(\xi) := \xi \mathbf{1}_{\|\xi\|_{nm}\leq 1}(\xi)$ and $\nu_L \colon \mathcal{B}(\mathbb{M}_{n,m}) \to \mathbb{R}^+$ is the *Lévy measure*. We call a Lévy process $(L_t)_{t\geq 0}$ *integrable*, if $\mathbb{E}\left[\|L_t\|_{nm}\right] < \infty$ for all $t \geq 0$ and *square-integrable* whenever $\mathbb{E}\left[\|L_t\|_{nm}^2\right] < \infty$ for all $t \geq 0$. For a square-integrable Lévy process L with characteristic exponent (7.1) the mean of L_1 is denoted by μ_L and we have $\mu_L = \big(\gamma_L + \int_{\mathbb{M}_{n,m}\cap\{\xi \colon \|\xi\|_{n,m}>1\}} \xi \nu_L(\mathrm{d}\xi)\big)$. Moreover, we denote by $\mathcal{Q} \in \mathcal{L}(\mathbb{M}_{n,m})$ the covariance operator of L_1, which is given by $\mathcal{Q} = Q_1 + \int_{\mathbb{M}_{n,m}} \xi \otimes \xi \nu_L(\mathrm{d}\xi)$. For any Lévy process $(L_t^1)_{\geq 0}$ defined on the positive real line $\mathbb{R}^+$, we can choose a second, independent and identically distributed, Lévy process $(L_t^2)_{t\geq 0}$ to define a *two-sided Lévy process* $(L_t)_{t\in\mathbb{R}}$ by

$$L_t := \mathbf{1}_{\mathbb{R}^+}(t) L_t^1 - \mathbf{1}_{-\mathbb{R}^+\setminus\{0\}}(t) L_{-t-}^2,$$

where $L_{t-} = \lim_{s\nearrow t} L_s$ for all $t \geq 0$. Throughout this chapter we use the conventional and intuitive notation of stochastic integration with respect to matrix-valued integrators analogous to e.g. [10, 12].

7.2 Matrix-valued linear state-space models

Throughout this section we fix $m, n \in \mathbb{N}$ and let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ denote a filtered probability space satisfying the usual conditions. Moreover, we assume that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is rich enough to carry a $\mathbb{M}_{n,m}$-valued two-sided Lévy process $L = (L_t)_{t\in\mathbb{R}}$. We begin this section by introducing a very general class of linear *continuous-time state space models* defined on real $n \times m$-matrices:

Definition 7.1. Let $p \in \mathbb{N}$ and let the tuple $(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)$ consist of a *state transition operator* $\mathcal{A} \in \mathcal{L}\big((\mathbb{M}_{n,m})^p\big)$, an *input operator* $\mathcal{B} \in \mathcal{L}\big(\mathbb{M}_{n,m}, (\mathbb{M}_{n,m})^p\big)$, an *output operator* $\mathcal{C} \in \mathcal{L}\big((\mathbb{M}_{n,m})^p, \mathbb{M}_{n,m}\big)$ and a $\mathbb{M}_{n,m}$-valued two-sided Lévy process $L = (L_t)_{t\in\mathbb{R}}$. A *continuous-time linear state space model* on $\mathbb{M}_{n,m}$, associated with the parameter set $(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)$, consists of a *state-space equation* given by

$$\mathrm{d}Z_t = \mathcal{A}Z_t \,\mathrm{d}t + \mathcal{B}\,\mathrm{d}L_t, \quad t \in \mathbb{R}, \tag{7.2}$$

and an *observation equation* given by

$$X_t = \mathcal{C}Z_t, \quad t \in \mathbb{R}. \tag{7.3}$$

We call the $(\mathbb{M}_{n,m})^p$-valued process $(Z_t)_{t\in\mathbb{R}}$ the *state process* and the $\mathbb{M}_{n,m}$-valued process $(X_t)_{t\in\mathbb{R}}$ the *output process* of the continuous-time linear state space model (associated with $(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)$).

Continuous-time linear state space models have been rigorously studied in the deterministic and stochastic control literature over many decades, see e.g. [148, 4, 58]. We note that if $m = 1$ (that means for the state space $\mathbb{R}_n = \mathbb{M}_{n,1}$) our definition of a linear continuous-time state space model coincides with the one in [133, Definition 3.1].

We note that the state process $(Z_t)_{t\in\mathbb{R}}$ of a linear continuous-time state space model is simply a Lévy driven Ornstein-Uhlenbeck type process on the space $(\mathbb{M}_{n,m})^p$ and a solution to (7.2) is given by the variation-of-constant formula:

$$Z_t = \mathrm{e}^{(t-s)\mathcal{A}} Z_s + \int_s^t \mathrm{e}^{(t-u)\mathcal{A}} \mathcal{B}\,\mathrm{d}L_u, \quad s < t \in \mathbb{R}. \tag{7.4}$$

In general, a solution $(Z_t)_{t\in\mathbb{R}}$ to (7.2) is not unique. If for some $s \in \mathbb{R}$ we are given a $\mathcal{F}_s$-measurable random variable Z_s, then $(Z_t)_{t\geq s}$, given by (7.4), is the unique solution to (7.2) on $[s, \infty)$ adapted to the filtration $(\mathcal{F}_t)_{t\geq s}$. If the spectral bound of the transition operator $\mathcal{A}$ is strictly negative, i.e. $\tau(\mathcal{A}) < 0$, then it follows from [131, 36] that there exists a unique stationary solution to (7.2) if and only if $\mathbb{E}\left[\log(\|L_1\|_{nm})\right] < \infty$. In this case the unique stationary solution $(Z_t)_{t\in\mathbb{R}}$ is adapted and given by

$$Z_t = \int_{-\infty}^t \mathrm{e}^{(t-s)\mathcal{A}} \mathcal{B}\,\mathrm{d}L_s, \quad t \in \mathbb{R}. \tag{7.5}$$

Remark 7.2 (Uniqueness, stationarity and adaptedness)**.** Note that in case of $\tau(\mathcal{A}) \geq 0$, there could still, under certain conditions, exist a (unique) stationary solution to (7.2) on $\mathbb{R}$, see also Proposition 7.9 below. However, it might happen that the stationary solution is not adapted to the natural filtration $\mathbb{F}$, since Z_t possibly depends on the generated sigma-algebra $\sigma(L_s\colon s > t)$. A concrete example of a stationary output process with $\tau(\mathcal{A}) \geq 0$ is given in Section 8.3. If from a modeling perspective adaptedness to the natural filtration is required, then we shall either assume $\tau(\mathcal{A}) < 0$, for which the existence of a unique $\mathbb{F}$-adapted solution is known by the reasoning above. Or in case of $\tau(\mathcal{A}) \geq 0$, there may exist many solutions, but only one for every fixed $\mathcal{F}_s$-measurable initial condition, which also happens to be $\mathbb{F}$-adapted, but possibly non-stationary.

In the sequel we often distinguish between the two cases in Remark 7.2 which we call the *stable* (where $\tau(A) < 0$) and *non-stable* (where $\tau(A) \geq 0$) case (following the usual nomenclature, it is actually the *exponentially stable* and *non-exponentially stable* case). We find it appropriate to give a precise definition to prevent confusion:

Definition 7.3. We call a continuous-time state space model associated with the parameter set $(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)$ *stable*, whenever $\tau(\mathcal{A}) < 0$. If, moreover, L has finite log-moments, i.e. $\mathbb{E}\left[\log(\|L_1\|_{n,m})\right] < \infty$, then whenever we refer to the state process $(Z_t)_{t\in\mathbb{R}}$ we mean the unique stationary solution given by (7.5). In this case we call $(Z_t)_{t\in\mathbb{R}}$ the *stable state process* and $(X_t)_{t\in\mathbb{R}}$ the *stable output process*. In case of $\tau(\mathcal{A}) \geq 0$, we call the state space model *non-stable* and refer to $(Z_t)_{t\in\mathbb{R}}$ in (7.4) simply as a state process. If uniqueness is required, we may fix an initial value Z_s at $s \in \mathbb{R}$ for some $\mathcal{F}_s$-measurable measurable Z_s.

Given a state process $(Z_t)_{t\in\mathbb{R}}$ we now shift our focus to the output process $(X_t)_{t\in\mathbb{R}}$ defined in (7.3). In the next proposition we summarize some well known and easy to check properties of $(X_t)_{t\in\mathbb{R}}$.

Proposition 7.4. *Let $(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)$ be as in Definition 7.1 and let ψ_L be the characteristic exponent of L given by* (7.1)*. Then the process $(X_t)_{t\in\mathbb{R}}$ in* (7.3) *satisfies*

$$X_t = \mathcal{C}\,\mathrm{e}^{(t-s)\mathcal{A}} Z_s + \int_s^t \mathcal{C}\,\mathrm{e}^{(t-u)\mathcal{A}} \mathcal{B}\,\mathrm{d}L_u, \quad s < t \in \mathbb{R}, \tag{7.6}$$

and for every $x \in \mathbb{M}_{n,m}$ and $s \leq t$ we have

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle X_t, x\rangle_{mn}} \,|\, \mathcal{F}_s\right] = \exp\Big(\mathrm{i}\langle \mathcal{C}\,\mathrm{e}^{(t-s)\mathcal{A}} Z_s, x\rangle_{mn} + \int_s^t \psi_L\big(\mathcal{B}^*\,\mathrm{e}^{(t-u)\mathcal{A}^*}\mathcal{C}^* x\big)\,\mathrm{d}u\Big). \tag{7.7}$$

If L is integrable, then the conditional mean of X_t is finite and given by

$$\mathbb{E}\left[X_t \middle| \mathcal{F}_s\right] = \mathcal{C}\,\mathrm{e}^{(t-s)\mathcal{A}} Z_s + \int_0^{t-s} \mathcal{C}\,\mathrm{e}^{u\mathcal{A}} \mathcal{B} \mu_L \,\mathrm{d}u, \quad s < t \in \mathbb{R}. \tag{7.8}$$

If moreover L has finite log-moments and $(X_t)_{t\in\mathbb{R}}$ is a stable output process, then it is stationary and given by

$$X_t = \int_{-\infty}^{t} \mathcal{C}\,\mathrm{e}^{(t-s)\mathcal{A}} \mathcal{B}\,\mathrm{d}L_s, \quad t \in \mathbb{R}, \tag{7.9}$$

and if in addition L is integrable, then the mean of $(X_t)_{t\in\mathbb{R}}$ is given by

$$\mathbb{E}\left[X_t\right] = -\mathcal{C}\mathcal{A}^{-1}\mathcal{B}\mu_L, \quad t \in \mathbb{R}. \tag{7.10}$$

From (7.9) we see that the dynamics of a stable output process $(X_t)_{t\in\mathbb{R}}$ are only governed by the parameters $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ and the Lévy process L. More specifically, the dynamics depend solely on the action of the kernel function $g\colon \mathbb{R}^+ \to \mathcal{L}(\mathbb{M}_{n,m})$ given by $g(t) := \mathcal{C}\,\mathrm{e}^{\mathcal{A}t}\mathcal{B}$ applied to the increments of $(L_t)_{t\in\mathbb{R}}$. In contrast, we see that in the non-stable case, the dynamics of $(X_t)_{t\geq s}$ also depend on the initial value Z_s for $s < t$ through the term $\mathcal{C}\,\mathrm{e}^{(t-s)\mathcal{A}} Z_s$. As we will see, this has some consequences for the techniques available to study positivity of output processes in Section 7.3.

7.2.1 The Controller canonical form

In this section we introduce a more particular form of linear continuous-time state space models on $\mathbb{M}_{n,m}$ which in the discrete-time control literature is often called the *controller canonical form*. This controller canonical form will prove itself useful for two reasons: First, it will allow us to interpret the output processes of certain continuous-time state space models as a linear transformation of MCARMA processes as they were introduced in the seminal work [112]. Second, this form is particularly convenient to study positivity in the next section.

Let $p \in \mathbb{N}$ and denote by $\mathbf{0}$ the null operator in $\mathcal{L}(\mathbb{M}_{n,m})$, i.e. $\mathbf{0}$ maps every $x \in \mathbb{M}_{n,m}$ to the null matrix $\mathbb{0}_{n,m}$. Moreover, let $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p \in \mathcal{L}(\mathbb{M}_{n,m})$ and define the state transition operator $\mathcal{A}_p\colon (\mathbb{M}_{n,m})^p \to (\mathbb{M}_{n,m})^p$ as

$$\mathcal{A}_p := \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \mathbf{0} & \mathbf{I} \\ \mathbf{A}_p & \mathbf{A}_{p-1} & \dots & \dots & \mathbf{A}_1 \end{bmatrix}. \tag{7.11}$$

For every $\mathbf{x} = (x_1, \dots, x_p)^\intercal \in (\mathbb{M}_{n,m})^p$ we understand the operator $\mathcal{A}_p$ as

$$\mathcal{A}_p(\mathbf{x}) = \big(x_2, \dots, x_p, \sum_{i=1}^{p} \mathbf{A}_{p-i+1}(x_i)\big)^\intercal \in (\mathbb{M}_{n,m})^p.$$

Moreover, let $q \in \mathbb{N}_0$ such that $q < p$ and let $\mathbf{C}_0, \mathbf{C}_1, \dots, \mathbf{C}_{p-1} \in \mathcal{L}(\mathbb{M}_{n,m})$ with $\mathbf{C}_i = \mathbf{0}$ for every $i \in \mathbb{N}$ such that $i \geq q+1$ and $i \leq p-1$. We then define the output operator $\mathcal{C}_q \colon (\mathbb{M}_{n,m})^p \to \mathbb{M}_{n,m}$ by

$$\mathcal{C}_q \coloneqq [\mathbf{C}_0, \mathbf{C}_1, \dots, \mathbf{C}_{p-1}], \tag{7.12}$$

where $\mathcal{C}_q$ is to be understood as follows: For $\mathbf{x} = (x_1, \dots, x_p)^\intercal \in (\mathbb{M}_{n,m})^p$ we have

$$\mathcal{C}_q(\mathbf{x}) = \sum_{i=1}^{q+1} \mathbf{C}_{i-1}(x_i).$$

Finally, we define the input operator $E_p \in \mathcal{L}\big(\mathbb{M}_{n,m}, (\mathbb{M}_{n,m})^p\big)$ by

$$E_p \coloneqq e_p \otimes \mathbf{I}, \tag{7.13}$$

which for every $x \in \mathbb{M}_{n,m}$ is defined as $E_p(x) = (\mathbb{0}_{n,m}, \dots, \mathbb{0}_{n,m}, x)^\intercal \in (\mathbb{M}_{n,m})^p$. With this specification of $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$ the state process $(Z_t)_{t\in\mathbb{R}}$ becomes

$$Z_t = \mathrm{e}^{(t-s)\mathcal{A}_p} Z_s + \int_s^t \mathrm{e}^{(t-u)\mathcal{A}_p} E_p \,\mathrm{d}L_u, \quad s < t \in \mathbb{R}, \tag{7.14}$$

and the output process

$$X_t = \mathcal{C}_q \,\mathrm{e}^{(t-s)\mathcal{A}_p} Z_s + \int_s^t \mathcal{C}_q \,\mathrm{e}^{(t-u)\mathcal{A}_p} E_p \,\mathrm{d}L_u, \quad s < t \in \mathbb{R}, \tag{7.15}$$

and the analogous formulas (7.5) and (7.9) hold in the stable case. Note that the transition matrix $\mathcal{A}_p$ can be viewed as the companion block operator matrix of the following operator polynomial $\mathbf{P}$ with coefficients in $\mathcal{L}(\mathbb{M}_{n,m})$:

$$\mathbf{P}(\lambda) \coloneqq \mathbf{I}\lambda^p - \mathbf{A}_1\lambda^{p-1} - \mathbf{A}_2\lambda^{p-2} - \ldots - \mathbf{A}_p, \quad \lambda \in \mathbb{C}. \tag{7.16}$$

In the same spirit we introduce the operator polynomial $\mathbf{Q}$ which is associated with the output operator $\mathcal{C}_q$ and given by

$$\mathbf{Q}(\lambda) \coloneqq \mathbf{C}_0 + \mathbf{C}_1\lambda + \mathbf{C}_2\lambda^2 + \ldots + \mathbf{C}_q\lambda^q, \quad \lambda \in \mathbb{C}. \tag{7.17}$$

Now, recall vec: $\mathbb{M}_{n,m} \to \mathbb{R}_{nm}$ being the linear isometric isomorphism that maps every matrix $A \in \mathbb{M}_{n,m}$ to the vector of its columns by stacking the columns.

If $A \in \mathcal{L}(\mathbb{M}_{n,m})$ we denote by A^{vec} the *matrix representation* of A given by $A^{\text{vec}} := \text{vec} \circ A \circ \text{vec}^{-1}$. Note that $A^{\text{vec}} \in \mathcal{L}(\mathbb{R}_{nm}) \simeq \mathbb{M}_{nm}$ and by identification we consider A^{vec} as a $nm \times nm$-matrix. Moreover, we denote by $K^{(n,m)} \in \mathbb{M}_{nm}$ the commutation matrix, which is the unique matrix in $\mathbb{M}_{nm}$ such that for every $A \in \mathbb{M}_{n,m}$ we have $K^{(n,m)} \text{vec}(A) = \text{vec}(A^\intercal)$. We denote the inverse of $K^{(n,m)}$ by $K^{-(n,m)}$, which happens to be the transpose of $K^{(n,m)}$ as well. When the vec-operator is applied to an output process $(X_t)_{t\in\mathbb{R}}$ we obtain an $\mathbb{R}_{nm}$-valued process $(\text{vec}(X_t))_{t\in\mathbb{R}}$. By linearity, we see that the process $(\text{vec}(X_t))_{t\in\mathbb{R}}$ is again an output process of some continuous-time linear state space model on $\mathbb{R}_{nm}$. In the following proposition we show more, namely that the controller canonical form in (7.11)-(7.13) transforms, under the vec-transformation, into a controller canonical-like form of an $\mathbb{R}_{nm}$-valued state space model.

Proposition 7.5. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ such that $q < p$ and $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$ be as in* (7.11)-(7.13) *and L a two-sided Lévy process on $\mathbb{M}_{n,m}$. Moreover, let the state process $(Z_t)_{t\in\mathbb{R}}$ be given by* (7.14) *and the output process $(X_t)_{t\in\mathbb{R}}$ be as in* (7.15)*. We set $L^{\text{vec}} := \text{vec}(L)$ and define the* state transition matrix $\hat{\mathcal{A}}_p^{\text{vec}} \in \mathbb{M}_{pnm}$ *by*

$$\hat{\mathcal{A}}_p^{\text{vec}} := \begin{pmatrix} \mathbb{0}_{nm} & K^{-(n,m)} & \mathbb{0}_{nm} & \cdots & \mathbb{0}_{nm} \\ \mathbb{0}_{nm} & \mathbb{0}_{nm} & K^{-(n,m)} & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbb{0}_{nm} \\ \mathbb{0}_{nm} & \cdots & \cdots & \mathbb{0}_{nm} & K^{-(n,m)} \\ \hat{A}_p^{\text{vec}} & \hat{A}_{p-1}^{\text{vec}} & \cdots & \cdots & \hat{A}_1^{\text{vec}} \end{pmatrix}, \tag{7.18}$$

where $\hat{A}_i^{\text{vec}} := A_i^{\text{vec}} \circ K^{-(n,m)} \in \mathbb{M}_{nm}$ for every $i = 1, \ldots, p$ and A_i^{vec} denotes the matrix representation of $\mathbf{A}_i$. Moreover, we define the output matrix $\hat{\mathcal{C}}_q^{\text{vec}} \in \mathbb{M}_{nm,pnm}$ *by*

$$\hat{\mathcal{C}}_q^{\text{vec}} := \left(\hat{C}_0^{\text{vec}}, \hat{C}_1^{\text{vec}}, \ldots, \hat{C}_{p-1}^{\text{vec}}\right), \tag{7.19}$$

where $\hat{C}_j^{\text{vec}} = C_j^{\text{vec}} \circ K^{-(n,m)} \in \mathbb{M}_{nm}$ for every $j = 0, \ldots, p-1$ and C_j^{vec} denotes the matrix representation of $\mathbf{C}_j$. Lastly, we define the input matrix $\hat{E}_p^{\text{vec}} \in \mathbb{M}_{pnm,nm}$ *by*

$$\hat{E}_p^{\text{vec}} := e_p \otimes K^{-(n,m)}. \tag{7.20}$$

Then $(\text{vec}(X_t))_{t\in\mathbb{R}}$ is the output process of a $\mathbb{R}_{nm}$-valued continuous-time state space model associated with $(\hat{\mathcal{A}}_p^{\text{vec}}, \hat{\mathcal{C}}_q^{\text{vec}}, \hat{E}_p^{\text{vec}}, L^{\text{vec}})$ and such that

$$\mathrm{d}\,\text{vec}(Z_t^\intercal) = \hat{\mathcal{A}}_p^{\text{vec}} \text{vec}(Z_t^\intercal)\,\mathrm{d}t + \hat{E}_p^{\text{vec}}\,\mathrm{d}L_t^{\text{vec}}, \quad t \in \mathbb{R}, \tag{7.21}$$

$$\text{vec}(X_t) = \hat{\mathcal{C}}_q^{\text{vec}} \text{vec}(Z_t^\intercal), \quad t \in \mathbb{R}. \tag{7.22}$$

Proof. For every $t \in \mathbb{R}$ write $Z_t = (Z_t^{(1)}, Z_t^{(2)}, \dots, Z_t^{(p)})^\intercal \in (\mathbb{M}_{n,m})^p$ where for every $i = 1, \dots, p$ we denote by $Z_t^{(i)}$ the i-th $n \times m$-block matrix component of Z_t. Given the state process $(Z_t)_{t\in\mathbb{R}}$ and output process $(X_t)_{t\in\mathbb{R}}$ we show that $(\operatorname{vec}(X_t))_{t\in\mathbb{R}}$ solves (7.22) with $(\operatorname{vec}(Z_t^\intercal))_{t\in\mathbb{R}}$ being a solution to (7.21). By definition we have $X_t = \mathcal{C}_q(Z_t) = \sum_{i=1}^{q+1} \mathbf{C}_{i-1}(Z_t^{(i)})$ and hence by linearity of the vec-operation we see that

$$\operatorname{vec}(X_t) = \operatorname{vec}(\mathcal{C}_q(Z_t)) = \sum_{i=1}^{q+1} \operatorname{vec}(\mathbf{C}_{i-1}(Z_t^{(i)})) = \sum_{i=1}^{q+1} C_{i-1}^{\mathrm{vec}} \operatorname{vec}(Z_t^{(i)}), \quad t \in \mathbb{R}. \tag{7.23}$$

As before, let $K^{(pn,m)}$ be the commutation matrix such that $K^{(pn,m)} \operatorname{vec}(F) = \operatorname{vec}(F^\intercal)$ for every $F \in (\mathbb{M}_{n,m})^p$. Then $\operatorname{vec}(Z_t^{(i)}) = K^{-(n,m)} \operatorname{vec}((Z_t^{(i)})^\intercal)$ and for every $i = 1, 2, \dots, p$ we can write

$$\operatorname{vec}((Z_t^{(i)})^\intercal) = (e_i^\intercal \otimes \mathbb{I}_{nm}) \operatorname{vec}(Z_t^\intercal).$$

Hence, for every $i = 1, 2, \dots, p$ and $t \in \mathbb{R}$ we have

$$\operatorname{vec}(Z_t^{(i)}) = K^{-(n,m)}(e_i^\intercal \otimes \mathbb{I}_{nm}) K^{(pn,m)} \operatorname{vec}(Z_t). \tag{7.24}$$

We thus continue with the term $\operatorname{vec}(Z_t)$ that appears on the right-hand side of (7.24). Note that by linearity, inserting (7.14) into $\operatorname{vec}(Z_t)$ yields

$$\operatorname{vec}(Z_t) = \operatorname{vec}(\mathrm{e}^{(t-s)\mathcal{A}_p} Z_s) + \int_s^t \operatorname{vec}(\mathrm{e}^{(t-u)\mathcal{A}_p} E_p \, \mathrm{d}L_u), \quad s < t \in \mathbb{R}. \tag{7.25}$$

We know that $t \mapsto \mathrm{e}^{t\mathcal{A}_p} v_0$ is the unique solution to the linear equation $\frac{\partial}{\partial t} v(t) = \mathcal{A}_p v(t)$ with $v(0) = v_0$. Moreover, we see that $\frac{\partial}{\partial t} \operatorname{vec}(v(t)) = \mathcal{A}_p^{\mathrm{vec}} \operatorname{vec}(v(t))$ with $\operatorname{vec}(v(0)) = \operatorname{vec}(v_0)$ is uniquely solved by $\operatorname{vec}(v(t)) = \mathrm{e}^{t\mathcal{A}_p^{\mathrm{vec}}} \operatorname{vec}(v_0)$. Hence, by linearity, we conclude that $(\mathrm{e}^{t\mathcal{A}_p})^{\mathrm{vec}} = \mathrm{e}^{t\mathcal{A}_p^{\mathrm{vec}}}$ must hold for all $t \geq 0$. Therefore, the right-hand side in (7.25) coincides with

$$\mathrm{e}^{(t-s)\mathcal{A}_p^{\mathrm{vec}}} \operatorname{vec}(Z_s) + \int_s^t \mathrm{e}^{(t-u)\mathcal{A}_p^{\mathrm{vec}}} \operatorname{vec}(E_p \, \mathrm{d}L_u), \quad s < t \in \mathbb{R}. \tag{7.26}$$

Hence, by using the $K^{(pn,m)}$-commutation matrix again, we see that for all real $s < t$ we have

$$\operatorname{vec}(Z_t) = \mathrm{e}^{(t-s)\mathcal{A}_p^{\mathrm{vec}}} K^{-(pn,m)} \operatorname{vec}(Z_s^\intercal) + \int_s^t \mathrm{e}^{(t-u)\mathcal{A}_p^{\mathrm{vec}}} K^{-(pn,m)} \operatorname{vec}((E_p \, \mathrm{d}L_u)^\intercal).$$

Hence, by (7.24)

$$\operatorname{vec}(Z_t^{(i)}) = K^{-(n,m)}(e_i^\intercal \otimes \mathbb{I}_{nm})\Big(K^{(pn,m)}\,\mathrm{e}^{(t-s)\mathcal{A}_p^{\mathrm{vec}}}K^{-(pn,m)}\operatorname{vec}(Z_s^\intercal) + \int_s^t K^{(pn,m)}\,\mathrm{e}^{(t-u)\mathcal{A}_p^{\mathrm{vec}}}K^{-(pn,m)}\operatorname{vec}((E_p\,\mathrm{d}L_u)^\intercal)\Big), \quad s<t\in\mathbb{R}. \tag{7.27}$$

Now note that

$$K^{(pn,m)}\,\mathrm{e}^{t\mathcal{A}_p^{\mathrm{vec}}}K^{-(pn,m)} = \mathrm{e}^{tK^{(pn,m)}\mathcal{A}_p^{\mathrm{vec}}K^{-(pn,m)}}, \quad \text{for all } t\geq 0.$$

Next, we show that $K^{(pn,m)}\mathcal{A}_p^{\mathrm{vec}}K^{-(pn,m)} = \hat{\mathcal{A}}_p^{\mathrm{vec}}$. Let $\mathbf{F}_p = (F_1, F_2, \ldots, F_p)^\intercal \in (\mathbb{M}_{n,m})^p$ then

$$\begin{aligned}
K^{(pn,m)}\mathcal{A}_p^{\mathrm{vec}}K^{-(pn,m)}\operatorname{vec}(\mathbf{F}_p^\intercal) &= K^{(pn,m)}\operatorname{vec}\Big(\big(F_2, F_2, \ldots, F_p, \sum_{i=1}^p \mathbf{A}_{p+1-i}(F_i)\big)^\intercal\Big)\\
&= \operatorname{vec}\Big(\big(F_2^\intercal, F_2^\intercal, \ldots, F_p^\intercal, \sum_{i=1}^p \mathbf{A}_{p+1-i}(F_i)^\intercal\big)\Big)\\
&= \big(\operatorname{vec}(F_2^\intercal), \ldots, \operatorname{vec}(F_p^\intercal), \sum_{i=1}^p \operatorname{vec}(\mathbf{A}_{p+1-i}(F_i)^\intercal)\big)^\intercal\\
&= \hat{\mathcal{A}}_p^{\mathrm{vec}}(\operatorname{vec}(\mathbf{F}_p^\intercal)),
\end{aligned}$$

where in the last equation we used that $K^{-(n,m)}\operatorname{vec}(F_i^\intercal) = \operatorname{vec}(F_i)$ and

$$\sum_{i=1}^p \operatorname{vec}(\mathbf{A}_{p+1-i}(F_i)^\intercal) = \sum_{i=1}^p A_{p+1-i}^{\mathrm{vec}}K^{-(n,m)}\operatorname{vec}(F_i) = \sum_{i=1}^p \hat{A}_{p+1-i}^{\mathrm{vec}}\operatorname{vec}(F_i).$$

This and $\operatorname{vec}((E_p\,\mathrm{d}L_u)^\intercal) = \hat{E}_p^{\mathrm{vec}}\,\mathrm{d}\operatorname{vec}(L_u^\intercal)$ inserted into (7.27) imply

$$\begin{aligned}
\operatorname{vec}(Z_t^{(i)}) &= K^{-(n,m)}(e_i^\intercal \otimes \mathbb{I}_{nm})\,\mathrm{e}^{(t-s)\hat{\mathcal{A}}_p^{\mathrm{vec}}}\operatorname{vec}(Z_s^\intercal)\\
&\quad + K^{-(n,m)}(e_i^\intercal \otimes \mathbb{I}_{nm})\int_s^t \mathrm{e}^{(t-u)\hat{\mathcal{A}}_p^{\mathrm{vec}}}\hat{E}_p^{\mathrm{vec}}\,\mathrm{d}\operatorname{vec}(L_u^\intercal),
\end{aligned}$$

which finally inserted back into (7.23) proves (7.21) and (7.22). □

Remark 7.6. i) Note that in order to obtain the correct autoregressive structure in (7.21), we have to take the transpose of the state process Z_t for $t\in\mathbb{R}$. That means we first stack the columns of the first block matrix entry $Z_t^{(1)}$ below each other, then below this real vector of length nm we append the stacked columns of $Z_t^{(2)}$ and so on until we finally obtain the vector $\operatorname{vec}(Z_t^\intercal) = \big(\operatorname{vec}(Z_t^{(1)}), \operatorname{vec}(Z_t^{(2)}), \ldots, \operatorname{vec}(Z_t^{(p)})\big)^\intercal \in \mathbb{R}_{pnm}$.

ii) In case of $\mathbb{M}_{n,1} = \mathbb{R}_n$ the vec-operator is the identity and also $\mathrm{vec}(Z_t^\intercal) = \mathrm{vec}(Z_t)$, i.e. in this case Proposition 7.5 is trivial. Note further that whenever the state process $(Z_t)_{t\in\mathbb{R}}$ takes values in $(\mathbb{S}_d)^p$ driven by a Levy process L with values in $\mathbb{S}_d$, then $K^{(d,d)} = \mathbb{I}_{d^2}$ and we see that $\hat{A}_i^{\mathrm{vec}} = A_i^{\mathrm{vec}}$, $\hat{C}_j^{\mathrm{vec}} = C_j^{\mathrm{vec}}$ and $\hat{E}_p^{\mathrm{vec}} = E_p^{\mathrm{vec}}$ for $i = 1,\dots,p$ and $j = 0,\dots,q$ and the representations in (7.18)-(7.20) become considerably simpler. Moreover, in this case we could replace vec by the vech operation.

iii) The controller canonical form for $\mathbb{R}_d$-valued MCARMA processes was already used in [30] for estimating the parameters of the driving Lévy process $(L_t)_{t\in\mathbb{R}}$. Proposition 7.5 suggests that the results of [30] straightforwardly extend to the matrix-valued case.

From the representation in (7.21) and (7.22) we can read off the following *second order moment structure* for the process $(X_t)_{t\in\mathbb{R}}$, see also [10] or Proposition 5.2.

Proposition 7.7. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ such that $q < p$ and assume that $(\hat{\mathcal{A}}_p^{\mathrm{vec}}, \hat{E}_p^{\mathrm{vec}}, \hat{\mathcal{C}}_q^{\mathrm{vec}}, L^{\mathrm{vec}})$ is as in (7.18)-(7.20) with L being square-integrable and where we denote the covariance operator of $(L_t^{\mathrm{vec}})_{t\in\mathbb{R}}$ by $\mathcal{Q}^{\mathrm{vec}}$. Then the absolute second moment of the process $(\mathrm{vec}(X_t))_{t\in\mathbb{R}}$, given by (7.22), exists and we have*

$$\mathrm{Var}\left[\mathrm{vec}(X_t)|\mathcal{F}_s\right] = \hat{\mathcal{C}}_q^{\mathrm{vec}} \Sigma_{t,s}^{\mathrm{vec}} (\hat{\mathcal{C}}_q^{\mathrm{vec}})^\intercal, \quad s < t \in \mathbb{R}, \tag{7.28}$$

where

$$\Sigma_{t,s}^{\mathrm{vec}} := \int_s^t \mathrm{e}^{u\hat{\mathcal{A}}_p^{\mathrm{vec}}} \hat{E}_p^{\mathrm{vec}} \mathcal{Q}^{\mathrm{vec}} \big(\hat{E}_p^{\mathrm{vec}}\big)^\intercal \mathrm{e}^{u(\hat{\mathcal{A}}_p^{\mathrm{vec}})^\intercal} \,\mathrm{d}u.$$

Moreover, for every $h \geq 0$ the auto-covariance of $(\mathrm{vec}(X_t))_{t\in\mathbb{R}}$ satisfies

$$\mathrm{Cov}\left[\mathrm{vec}(X_{t+h}), \mathrm{vec}(X_t)|\mathcal{F}_s\right] = \hat{\mathcal{C}}_q^{\mathrm{vec}}\, \mathrm{e}^{h\hat{\mathcal{A}}_p^{\mathrm{vec}}} \Sigma_{t,s}^{\mathrm{vec}} (\hat{\mathcal{C}}_q^{\mathrm{vec}})^\intercal, \quad s < t \in \mathbb{R},\, h \geq 0. \tag{7.29}$$

If in addition $(X_t)_{t\in\mathbb{R}}$ is stable and given by (7.9), then

$$\mathrm{Var}\left[\mathrm{vec}(X_t)\right] = \hat{\mathcal{C}}_q^{\mathrm{vec}} \Sigma_\infty^{\mathrm{vec}} (\hat{\mathcal{C}}_q^{\mathrm{vec}})^\intercal \quad \forall t \in \mathbb{R}, \tag{7.30}$$

where $\Sigma_\infty^{\mathrm{vec}} := \int_0^\infty \mathrm{e}^{u\hat{\mathcal{A}}_p^{\mathrm{vec}}} \hat{E}_p^{\mathrm{vec}} \mathcal{Q}^{\mathrm{vec}} \big(\hat{E}_p^{\mathrm{vec}}\big)^\intercal \mathrm{e}^{u(\hat{\mathcal{A}}_p^{\mathrm{vec}})^\intercal} \,\mathrm{d}u$ and the auto-covariance is

$$\mathrm{Cov}\left[\mathrm{vec}(X_{t+h}), \mathrm{vec}(X_t)\right] = \hat{\mathcal{C}}_q^{\mathrm{vec}}\, \mathrm{e}^{h\hat{\mathcal{A}}_p^{\mathrm{vec}}} \Sigma_\infty (\hat{\mathcal{C}}_q^{\mathrm{vec}})^\intercal, \quad t \in \mathbb{R},\, h \geq 0. \tag{7.31}$$

A concrete example for the auto-covariance (7.29) of an output process of a linear state-space model can be found in (8.15) below.

Proposition 7.5 tells us that the process $(\mathrm{vec}(X_t))_{t\in\mathbb{R}}$ can be interpreted as the output process of an $\mathbb{R}_{nm}$-valued continuous-time state space model in a controller canonical form entry-wise composited with the commutation matrix $K^{-(n,m)}$. It is well known that every $\mathbb{R}_{nm}$-valued MCARMA process possesses a state space representation. Conversely, the result in [133, Theorem 3.3] describes precisely those state space specifications such that the associated output process gives rise to a (causal) MCARMA process. We introduce the following notion: We call the function $H\colon \mathbb{C} \to \mathcal{L}(\mathbb{M}_{n,m})$ given by

$$H(\lambda) \coloneqq \mathcal{C}(\lambda \mathbf{I}_p - \mathcal{A})^{-1}\mathcal{B}, \quad \lambda \in \mathbb{C}, \tag{7.32}$$

the *transfer function* of the continuous-time linear state space model associated with $(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)$. The transfer function H associated with $(\hat{\mathcal{A}}_p^{\mathrm{vec}}, \hat{E}_p^{\mathrm{vec}}, \hat{\mathcal{C}}_q^{\mathrm{vec}}, L^{\mathrm{vec}})$ as in Proposition 7.5 satisfies the following:

Lemma 7.8. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ such that $q < p$ and let $(\hat{\mathcal{A}}_p^{\mathrm{vec}}, \hat{E}_p^{\mathrm{vec}}, \hat{\mathcal{C}}_q^{\mathrm{vec}}, L^{\mathrm{vec}})$ be as in Proposition 7.5. Then for every $\lambda \in \mathbb{C}$ we have*

$$\hat{\mathcal{C}}_q^{\mathrm{vec}}(\lambda \mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{E}_p^{\mathrm{vec}} = \hat{Q}(\lambda)\hat{P}(\lambda)^{-1}, \tag{7.33}$$

where $\hat{Q}, \hat{P} \in \mathbb{M}_{nm}(\mathbb{R}[\lambda])$ for $\lambda \in \mathbb{C}$ are given by

$$\hat{Q}(\lambda) \coloneqq \hat{C}_0^{\mathrm{vec}} + \hat{C}_1^{\mathrm{vec}} K^{(n,m)}\lambda + \hat{C}_2^{\mathrm{vec}}(K^{(n,m)}\lambda)^2 + \ldots + \hat{C}_q^{\mathrm{vec}}(K^{(n,m)}\lambda)^q, \tag{7.34}$$

and

$$\hat{P}(\lambda) \coloneqq (K^{(n,m)}\lambda)^p - A_1^{\mathrm{vec}}(K^{(n,m)}\lambda)^{p-1} - A_2^{\mathrm{vec}}(K^{(n,m)}\lambda)^{p-2} - \ldots - A_p^{\mathrm{vec}}. \tag{7.35}$$

Proof. Let $\lambda \in \mathbb{C}$ and $F \coloneqq (F_1, F_2, \ldots, F_p)^\intercal \in \mathcal{L}(\mathbb{R}_{nm}, \mathbb{R}_{pnm})$ with $F_i \in \mathbb{M}_{nm}$ for all $i = 1, \ldots, p$ and such that $F(x) = (F_1 x, F_2 x, \ldots, F_p x)^\intercal$ for all $x \in \mathbb{R}_{nm}$. We solve the matrix equation $(\lambda \mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})F = \hat{E}_p^{\mathrm{vec}}$, where the left-hand side equals

$$\big(\lambda F_1 - K^{-(n,m)}F_2, \ldots, \lambda F_{p-1} - K^{-(n,m)}F_p, \lambda F_p - \hat{A}_p^{\mathrm{vec}}F_1 - \ldots - \hat{A}_1^{\mathrm{vec}}F_p\big)^\intercal.$$

Setting this equal to $(\mathbb{0}_{nm}, \ldots, \mathbb{0}_{nm}, K^{-(n,m)})^\intercal$ and solving for F yields

$$F_i = (\lambda K^{(n,m)})^{-(p-i)}F_p, \quad \text{for } i = 1, \ldots, p-1, \tag{7.36}$$

and inserting this into the last equation we see that the term

$$\lambda F_p - \sum_{i=1}^p \hat{A}_{p+1-i}^{\mathrm{vec}}F_i = \Big(\lambda^p (K^{(n,m)})^{p-1} - \sum_{i=1}^p \hat{A}_i^{\mathrm{vec}}(\lambda K^{(n,m)})^{p-i}\Big)(\lambda K^{(n,m)})^{1-p}F_p,$$

must coincide with the matrix $K^{-(n,m)}$.

This, by definition of $\hat{P}(\lambda)$, is equivalent to $F_p = (\lambda K^{(n,m)})^{p-1}\hat{P}(\lambda)^{-1}$. Hence, by (7.36) we see that $F_i = (\lambda K^{(n,m)})^{i-1}\hat{P}(\lambda)^{-1}$, i.e.

$$F = (\hat{P}(\lambda)^{-1}, \lambda K^{(n,m)}\hat{P}(\lambda)^{-1}, \ldots, (\lambda K^{(n,m)})^{p-1}\hat{P}(\lambda)^{-1})^{\intercal}.$$

From this and the definition of $\hat{Q}$ we conclude that

$$\hat{\mathcal{C}}_q^{\mathrm{vec}}(\lambda \mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{E}_p^{\mathrm{vec}} = \hat{\mathcal{C}}_q^{\mathrm{vec}} F = \hat{Q}(\lambda)\hat{P}(\lambda)^{-1}, \quad \lambda \in \mathbb{C},$$

which proves (7.33). □

The following proposition is the key result for the definition of matrix-valued MCARMA processes. We relegated the proof to Section 7.4.

Proposition 7.9. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ such that $q < p$ and assume that $(\hat{\mathcal{A}}_p^{\mathrm{vec}}, \hat{E}_p^{\mathrm{vec}}, \hat{\mathcal{C}}_q^{\mathrm{vec}}, L^{\mathrm{vec}})$ is as in Proposition 7.5. Moreover, let $\hat{Q}(\lambda)$ and $\hat{P}(\lambda)$ be given by* (7.34) *and* (7.35)*, respectively. Then there exist two matrix polynomials $\tilde{Q}, \tilde{P} \in \mathbb{M}_{nm}(\mathbb{R}[\lambda])$ such that*

$$\tilde{Q}(\lambda) = \tilde{C}_0\lambda^q + \tilde{C}_1\lambda^{q-1} + \ldots + \tilde{C}_q, \quad \lambda \in \mathbb{C}, \tag{7.37}$$

with $\tilde{C}_j \in \mathbb{M}_{nm}$ for $j = 0, \ldots, q$ and

$$\tilde{P}(\lambda) = \mathbb{I}_{nm}\lambda^p - \tilde{A}_1\lambda^{p-1} - \ldots - \tilde{A}_p, \quad \lambda \in \mathbb{C}, \tag{7.38}$$

with $\tilde{A}_i \in \mathbb{M}_{nm}$ for $i = 1, \ldots, p$ satisfying

$$\tilde{P}(\lambda)^{-1}\tilde{Q}(\lambda) = \hat{Q}(\lambda)\hat{P}(\lambda)^{-1} \quad \textit{for all } \lambda \in \mathbb{C}, \tag{7.39}$$

and $\det(\tilde{P}(\lambda)) = 0$ if and only if $\det(\hat{P}(\lambda)) = 0$. If moreover, the logarithmic moment condition $\mathbb{E}[\log(\|L_1\|_{nm})] < \infty$ holds and

$$\left\{\lambda \in \mathbb{C}\colon \det(\hat{P}(\lambda)) = 0\right\} \subseteq \mathbb{R}\setminus\{0\} + \mathrm{i}\mathbb{R}, \tag{7.40}$$

then equation (7.22) *has a stationary solution, unique in law, given by*

$$\mathrm{vec}(X_t) = \int_{-\infty}^{\infty} g(t-s)\, \mathrm{d}L_s^{\mathrm{vec}}, \quad t \in \mathbb{R}, \tag{7.41}$$

where $g\colon \mathbb{R} \to \mathbb{M}_{nm}$ is given by

$$g(t) := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\xi t}\tilde{P}(\mathrm{i}\xi)^{-1}\tilde{Q}(\mathrm{i}\xi)\, \mathrm{d}\xi\,, \quad t \in \mathbb{R}. \tag{7.42}$$

Following [112, Theorem 3.22], every $\mathbb{R}_{nm}$-valued MCARMA process $(Y_t)_{t\in\mathbb{R}}$ with *moving-average polynomial matrix* $\check{Q} \in \mathbb{M}_{nm}(\mathbb{R}[\lambda])$, *autoregressive polynomial matrix* $\check{P} \in \mathbb{M}_{nm}(\mathbb{R}[\lambda])$ and input Lévy process $\check{L}$ on $\mathbb{R}_{nm}$ is given by

$$Y_t = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\xi(t-s)}\check{P}(\mathrm{i}\xi)^{-1}\check{Q}(\mathrm{i}\xi)\,\mathrm{d}\xi\,\mathrm{d}\check{L}_s, \quad t \in \mathbb{R},$$

where $\check{P}$ satisfies $\big\{\lambda \in \mathbb{C}\colon \det(\check{P}(\lambda)) = 0\big\} \subseteq \mathbb{R}\setminus\{0\} + \mathrm{i}\mathbb{R}$ and $\check{L}$ is such that $\mathbb{E}\Big[\log(\|\check{L}_1\|_{nm})\Big] < \infty$. It thus follows from Proposition 7.9, that the unique stationary process $(\mathrm{vec}(X_t))_{t\geq 0}$ in (7.41) is an $\mathbb{R}_{mn}$-valued MCARMA process with moving-average polynomial matrix $\tilde{Q}$, autoregressive polynomial matrix $\tilde{P}$ and Lévy noise L^{vec}. Moreover, by [112, Remark 3.19] it can be interpreted as the solution of the higher order stochastic differential equation (6.79) with L replaced by L^{vec}. This justifies the following definition of a matrix-valued multivariate continuous-time autoregressive moving-average process (by means of transformed $\mathbb{R}_{nm}$-valued MCARMA processes):

Definition 7.10. Let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ such that $q < p$ and $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$ be as in (7.11)-(7.13). If moreover $\mathbb{E}\left[\log(\|L_1\|_{n,m})\right] < \infty$ and (7.40) is satisfied, then we call the unique output process $(X_t)_{t\in\mathbb{R}}$ in (7.15) for which $(\mathrm{vec}(X_t))_{t\in\mathbb{R}}$ is the stationary solution to (7.22), an $\mathbb{M}_{n,m}$*-valued continuous-time autoregressive moving-average (MCARMA) process of order* (p, q). In the special case where $q = 0$ and $C_0 = \mathbf{I}$, i.e. $\mathcal{C}_0 = [\mathbf{I}, \mathbf{0}, \dots, \mathbf{0}]$, we call $(X_t)_{t\in\mathbb{R}}$ an $\mathbb{M}_{n,m}$*-valued MCAR process of order* p instead. Moreover, if in the situation above we have $\tau(\mathcal{A}_p) < 0$, or equivalently $\big\{\lambda \in \mathbb{C}\colon \det(\hat{P}(\lambda)) = 0\big\} \subset (-\infty, 0) + \mathrm{i}\mathbb{R}$, then we say that $(X_t)_{t\in\mathbb{R}}$ is a *causal* MCAR(MA) process of order (p, q) (resp. p). Otherwise, if $\tau(\mathcal{A}_p) \geq 0$ we say that $(X_t)_{t\in\mathbb{R}}$ is *non-causal*.

Remark 7.11.

i) By Definition 7.10 we see that every causal MCARMA process with values in $\mathbb{M}_{n,m}$ is the output process of a stable linear state space model and non-causal MCARMA processes correspond to non-stable linear state space models. Moreover, it can be seen from the representation (7.41) (see also (7.60) in the proof of Proposition 7.9), that non-causal MCARMA processes are not adapted to the natural filtration $\mathbb{F}$, since for $t \in \mathbb{R}$, X_t depends on $\sigma(L_s\colon s > t)$.

ii) Note that the vectorized versions $\hat{P}$ and $\hat{Q}$ of the operator polynomials $\mathbf{P}$ and $\mathbf{Q}$ in (7.16) and (7.17) can in general not be interpreted as the moving-average and autoregressive polynomials of the MCARMA process. Instead, the polynomials $\tilde{Q}$ and $\tilde{P}$ in (7.39) can be naturally interpreted as such, i.e. by means of equation (6.79). If, however, $\hat{P}$ and $\hat{Q}$ commute, then $\hat{P}(\lambda) = \tilde{P}(\lambda)$ and $\hat{Q}(\lambda) = \tilde{Q}(\lambda)$ for all $\lambda \in \mathbb{C}$.

7.3 Positivity of MCARMA processes

In this section we study *positivity* of $\mathbb{M}_{n,m}$-valued MCARMA processes of order (p, q), as they were defined in Definition 7.10 above. Recall that we use the term *positive* as a synonym for *cone valued* as our main examples of multivariate cones, $\mathbb{R}_d^+$ and $\mathbb{S}_d^+$, are both termed *positive*. This section is divided into the case of causal MCARMA in Section 7.3.2 and the non-causal MCARMA case in Section 7.3.3. A careful distinction between the two cases is justified as the positivity constraints may interact with the stability conditions. Our main results are Theorem 7.16 and Theorem 7.24 below, which establish sufficient and/or necessary conditions for the positivity of $\mathbb{M}_{n,m}$-valued (non-)causal MCARMA processes.

7.3.1 Positive operators and cone valued Lévy processes

Before we study the positivity of $\mathbb{M}_{n,m}$-valued MCARMA processes in the next two sections, we recall some additional preliminaries concerning *convex (algebra) cones*, *(quasi)-positive* operators and *increasing Lévy processes.* For $n, m \in \mathbb{N}$, we consider the inner product space $(\mathbb{M}_{n,m}, (\cdot,\cdot)_{n,m})$ and assume that $\mathbb{M}_{n,m}$ is equipped with a convex cone C, i.e. $C \subseteq \mathbb{M}_{n,m}$ is such that $C + C \subseteq C$, $\lambda C \subseteq C$ for all $\lambda \in \mathbb{R}^+$ and $C \cap (-C) = \{\mathbb{0}_{n,m}\}$. Moreover, we write " $\leq_C$ " for the partial-ordering on $\mathbb{M}_{n,m}$ induced by C, i.e. for $x, y \in \mathbb{M}_{n,m}$: $x \leq_C y$ if and only if $y - x \in C$.

Positive and quasi-positive operators

We denote by $\pi(C) \subseteq \mathcal{L}(\mathbb{M}_{n,m})$ the set of all linear operators leaving the cone C invariant, i.e.

$$\pi(C) = \{A \in \mathcal{L}(\mathbb{M}_{n,m})\colon A(u) \geq_C 0 \,\text{for all } u \geq_C 0\} .$$

We call elements in $\pi(C)$ *positive operators* on $\mathbb{M}_{n,m}$. Note that the set $\pi(C)$ is a *convex algebra cone*, that means it is a convex cone such that $B_1, B_2 \in \pi(C)$ implies $B_1B_2 \in \pi(C)$. We denote by " $\preceq$ " the partial ordering on $\mathcal{L}(\mathbb{M}_{n,m})$ induced by $\pi(C)$. Moreover, we call an element $A \in \mathcal{L}(\mathbb{M}_{n,m})$ *quasi-positive* or *quasi-monotone*, if $\exp(At) \succeq 0$ for all $t \geq 0$, where $\exp(At)$ denotes the operator exponential of At. It is well known that A is quasi-positive if and only if for all $u, v \in C$ with $\langle u, v\rangle_{n,m} = 0$ we have $\langle Au, v\rangle_{n,m} \geq 0$, see e.g. [76] and Definition 2.10. Note here that we already used this notion in Chapter 2 to ensure that the linear drift of an affine process is inward pointing at the boundary.

The following two cases are our main examples for convex cones in $\mathbb{M}_{n,m}$:

a) For $m = 1$ and $n = d$ we have $(\mathbb{R}_d, (\cdot,\cdot)_d) = (\mathbb{M}_{d,1}, (\cdot,\cdot)_{d,1})$. On $\mathbb{R}_d$ we consider the positive orthant $C = \mathbb{R}_d^+$, which is a convex cone and we denote its induced partial ordering by " $\leq_d$ ". It is well known that the cone $\pi(\mathbb{R}_d^+) \subseteq \mathbb{M}_d$ of $\mathbb{R}_d^+$-preserving linear maps (matrices) is given by the set of all positive matrices (more precisely *non-negative matrices*), i.e.

$$\pi(\mathbb{R}_d^+) = \{(a_{i,j})_{1\leq i,j\leq d} \in \mathbb{M}_d \colon\ a_{i,j} \geq 0\ \forall i,j = 1,\ldots,d\}.$$

It is also well-known that a matrix $A = (a_{i,j})_{1\leq i,j\leq d} \in \mathbb{M}_d$ is quasi-positive (sometimes called *cross-positive*) if and only if $a_{i,j} \geq 0$ for all $i,j = 1,\ldots,d$ such that $i \neq j$, i.e. all off-diagonal elements are non-negative and the diagonal ones can be arbitrary, see e.g. [76].

b) In case of $n = m = d$ for some $d \in \mathbb{N}$ we consider the space of real $d \times d$-matrices $\mathbb{M}_d$. On $\mathbb{M}_d$ we consider the convex cone of all symmetric and positive-semi definite matrices $C = \mathbb{S}_d^+$ and denote the induced partial ordering by " $\leq_{\mathbb{S}_d^+}$ ". As far as we know, there is no analogous characterization of the set $\pi(\mathbb{S}_d^+)$ known. Partial results were achieved in this direction, see e.g. [135, 108] for some related results in the theory of linear preserver problems.

Multivariate Lévy processes on cones

Let $L = (L_t)_{t\in\mathbb{R}}$ denote a two-sided Lévy process on $\mathbb{M}_{n,m}$ defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and let C be a convex cone in $\mathbb{M}_{n,m}$. We define a *positive Lévy process* as follows:

Definition 7.12. We call a two-sided Lévy process $(L_t)_{t\in\mathbb{R}}$ on $\mathbb{M}_{n,m}$ *C-positive*, if it is C-valued or, equivalently, if $L_t - L_s \in C$ for all $t, s \in \mathbb{R}$ such that $s < t$.

The characteristic exponent (7.1) of a C-positive Lévy process is given by

$$\varphi_L(z) = \mathrm{i}\langle \gamma_L, z\rangle_{n,m} + \int_C (\mathrm{e}^{\mathrm{i}\langle \xi, z\rangle_{n,m}} - 1 - \mathrm{i}\langle \chi(\xi), z\rangle_{n,m})\nu_L(\mathrm{d}\xi), \quad z \in \mathbb{M}_{n,m},$$

where the drift γ_L is positive, i.e. $\gamma_L \in C$ and the Lévy measure ν_L is concentrated on $C \setminus \{\mathbb{0}_{n,m}\}$, hence jumps, small or large, of the Lévy process are positive. Moreover, note that compared to (7.1) the diffusion part vanishes, i.e. a positive Lévy process is of pure-jump type. We refer to [8] for more information on matrix-valued positive Lévy processes.

7.3.2 Positive causal MCARMA processes

Throughout this section we fix $n, m \in \mathbb{N}$ and let C be a cone in $\mathbb{M}_{n,m}$. Moreover, we assume that $L = (L_t)_{t\in\mathbb{R}}$ is a C-valued two-sided Lévy process defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ such that $\mathbb{E}\left[\log(\|L_1\|_{n,m})\right] < \infty$. We let $\mathcal{A}_p$, E_p and $\mathcal{C}_q$ be as in (7.11)-(7.13) and assume that $\tau(\mathcal{A}_p) < 0$. This of course implies that the stable output process $(X_t)_{t\in\mathbb{R}}$, associated with $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$, is a causal $\mathbb{M}_{n,m}$-valued MCARMA process given by

$$X_t = \int_{-\infty}^{t} \mathcal{C}_q \,\mathrm{e}^{(t-s)\mathcal{A}_p} E_p \,\mathrm{d}L_s, \quad t \in \mathbb{R}. \tag{7.43}$$

We begin with a small lemma on the connection between quasi-positive operators and their spectral bound:

Lemma 7.13. *Let V be a linear space and $C \subseteq V$ a convex cone. Then for every quasi-positive $\mathbf{A} \in \mathcal{L}(V)$ with $\tau(\mathbf{A}) < 0$ we have $-\mathbf{A}^{-1} \succeq 0$.*

Proof. It follows from the spectral mapping theorem that for every $\mathbf{A} \in \mathcal{L}(V)$ with $\tau(\mathbf{A}) < 0$ we have $\tau\big(\exp(\mathbf{A})\big) < 1$. It thus follows that $\mathrm{e}^{\mathbf{A}t} \to 0$ as $t \to \infty$ and therefore we see that

$$-\mathbf{A}^{-1} = \int_0^\infty \mathrm{e}^{\mathbf{A}s} \,\mathrm{d}s.$$

Hence, whenever $\mathbf{A}$ is quasi-positive, i.e. $\exp(\mathbf{A}s) \succeq 0$ for every $s \geq 0$, we have $\int_0^\infty \mathrm{e}^{\mathbf{A}s}\,\mathrm{d}s \succeq 0$ and consequently $-\mathbf{A}^{-1} \succeq 0$. □

From the representation (7.43) we see that $X_t \in C$ for every $t \in \mathbb{R}$, whenever the Lévy process $(L_t)_{t\in\mathbb{R}}$ is C-valued and $g(s) = \mathcal{C}_q \,\mathrm{e}^{s\mathcal{A}_p} E_p \in \pi(C)$ holds true for every $s \geq 0$. In the following lemma we prove a particular form of the Laplace transform of the kernel function g. The main part of the proof, the computation of the transfer function, is similar to the matrix representation case in Lemma 7.8.

Lemma 7.14. *The Laplace transform $\varphi\colon \mathbb{R}^+ \to \mathcal{L}(\mathbb{M}_{n,m})$ of the kernel $g(s) = \mathcal{C}_q \,\mathrm{e}^{s\mathcal{A}_p} E_p$ exists and is given by*

$$\varphi(\lambda) = \mathbf{Q}(\lambda)\mathbf{P}(\lambda)^{-1}, \quad \lambda \geq 0. \tag{7.44}$$

Proof. Since $\tau(\mathcal{A}_p) < 0$ we see that for $\lambda \geq 0$ the resolvent $R(\lambda, \mathcal{A}_p) = (\lambda\mathbf{I} - \mathcal{A}_p)^{-1}$ is given by the Laplace transform of the matrix semigroup $\mathrm{e}^{t\mathcal{A}_p}$, i.e.

$$R(\lambda, \mathcal{A}_p) = \int_0^\infty \mathrm{e}^{-\lambda s}\,\mathrm{e}^{s\mathcal{A}_p}\,\mathrm{d}s.$$

We thus compute

$$\varphi(\lambda) = \int_0^\infty \mathrm{e}^{-\lambda s} g(s)\, \mathrm{d}s = \mathcal{C}_q \left(\int_0^\infty \mathrm{e}^{-\lambda s}\, \mathrm{e}^{s\mathcal{A}_p}\, \mathrm{d}s \right) E_p = \mathcal{C}_q R(\lambda, \mathcal{A}_p) E_p. \quad (7.45)$$

For $\lambda \in \mathbb{C}$ we compute the term $R(\lambda, \mathcal{A}_p)E_p$ as follows: Let

$$\mathbf{F}_p := [F_1, F_2, \dots, F_p]^\intercal \in \mathcal{L}(\mathbb{M}_{n,m}, (\mathbb{M}_{n,m})^p)$$

with $F_i \in \mathcal{L}(\mathbb{M}_{n,m})$ and such that $\mathbf{F}_p$ applied to x satisfies

$$\mathbf{F}_p(x) = (F_1 x, F_2 x, \dots, F_p x)^\intercal.$$

Moreover, set $\mathbf{A} := [\mathbf{A}_p, \mathbf{A}_{p-1}, \dots, \mathbf{A}_1]$ and consider $(\lambda \mathbf{I} - \mathcal{A}_p)\mathbf{F}_p = E_p$, which is equivalent to

$$[\lambda F_1 - F_2, \lambda F_2 - F_3, \dots, \lambda F_{p-1} - F_p, \lambda F_p - \mathbf{A}\mathbf{F}_p] = [0, \dots, \mathbf{I}], .$$

Solving for $\mathbf{F}_p$ yields $F_1 = \lambda^{-1} F_2$, $F_2 = \lambda^{-1} F_3, \dots, F_{p-1} = \lambda^{-1} F_p$ and thus for every $i = 1, \dots, p-1$ we have $F_i = \lambda^{-(p-i)} F_p$. Moreover, the last equation reads as

$$\left(\lambda^p \mathbf{I} - \mathbf{A}_p - \mathbf{A}_{p-1}\lambda - \ldots - \mathbf{A}_1 \lambda^{p-1}\right) \lambda^{-(p-1)} F_p = \mathbf{I},$$

and hence, by definition of $\mathbf{P}(\lambda)$ we see that $F_p = \lambda^{p-1}\mathbf{P}(\lambda)^{-1}$ and therefore $F_i = \lambda^{i-1}\mathbf{P}(\lambda)^{-1}$ for $i = 1, \dots, p-1$. In vector notation this means

$$\mathbf{F}_p = [\mathbf{I} \circ \mathbf{P}(\lambda)^{-1}, \lambda \mathbf{I} \circ \mathbf{P}(\lambda)^{-1}, \dots, \lambda^{p-1}\mathbf{I} \circ \mathbf{P}(\lambda)^{-1}] = (1, \lambda, \dots, \lambda^{p-1}) \otimes \mathbf{P}(\lambda)^{-1}.$$

Thus inserting $\mathbf{F}_p = R(\lambda, \mathcal{A}_p)E_p$ back into (7.45) and by definition of $\mathbf{Q}$ in (7.17) yields

$$\mathcal{C}_q \mathbf{F}_p = \mathbf{C}_0 P(\lambda)^{-1} + \lambda \mathbf{C}_1 P(\lambda)^{-1} + \ldots + \lambda^{p-1}\mathbf{C}_{p-1} P(\lambda)^{-1} = \mathbf{Q}(\lambda)\mathbf{P}(\lambda)^{-1}.$$

□

Next, we introduce the fundamental property of the Laplace transforms of the kernel function $s \mapsto g(s)$ that will ensure the positivity of the associated causal MCARMA processes. The following definition is adapted from [3, Definition 5.4].

Definition 7.15. We call a function $f\colon \mathbb{R}^+ \to \mathcal{L}(\mathbb{M}_{n,m})$ *completely monotone* with respect to $\pi(C)$, if f is infinitely often differentiable and $(-1)^n f^{(n)}(\lambda) \succeq 0$ for all $\lambda > 0$ and $n \in \mathbb{N}_0$.

The following proposition is our main result on the positivity of matrix-valued causal MCARMA processes.

Theorem 7.16. *Let $(X_t)_{t\in\mathbb{R}}$ be a $\mathbb{M}_{n,m}$-valued causal MCARMA process of order (p,q) given by (7.43). Moreover, let $\mathbf{P}(\lambda)$ and $\mathbf{Q}(\lambda)$ be the associated operator polynomials in (7.16) and (7.17), respectively. Then the following holds true:*

i) $(X_t)_{t\in\mathbb{R}}$ is C-valued if and only if the map $\lambda \mapsto \mathbf{Q}(\lambda)\mathbf{P}(\lambda)^{-1}$ is completely monotone with respect to $\pi(C)$.

ii) If $\lambda \mapsto \mathbf{Q}(\lambda)$ is completely monotone with respect to $\pi(C)$ and $\mathbf{P}(\lambda)$ can be decomposed into linear factors as follows:

$$\mathbf{P}(\lambda) = \prod_{i=1}^{p}(\lambda\mathbf{I} - \hat{\mathbf{A}}_i), \tag{7.46}$$

where for all $i = 1,\dots,p$ the operator $\hat{\mathbf{A}}_i \in \mathcal{L}(\mathbb{M}_{n,m})$ is quasi-positive and $\tau(\hat{\mathbf{A}}_i) < 0$, then $(X_t)_{t\in\mathbb{R}}$ is C-valued.

Proof. By Lemma 7.14 the Laplace transform ϕ of the kernel $g(s) = \mathcal{C}_q\,\mathrm{e}^{s\mathcal{A}_p}E_p$ is given by the operator rational function $\lambda \mapsto \mathbf{Q}(\lambda)\mathbf{P}(\lambda)^{-1}$. By a vector valued version of Bernstein's theorem, see [3, Theorem 5.5], and Lemma 7.14 it follows that $s \mapsto g(s)$ is positive. Indeed, by Bernstein's theorem we have $g(s) \in \pi(C)$ for all $s \geq 0$ if and only if its Laplace transform is completely monotone with respect to $\pi(C)$, but since its Laplace transform is given by $\lambda \mapsto \mathbf{Q}(\lambda)\mathbf{P}(\lambda)^{-1}$ Theorem 7.16 i) follows.
For the second statement Theorem 7.16 ii), we see that from Theorem 7.16 i) it follows that the MCARMA process $(X_t)_{t\geq 0}$ is C-valued if and only if $\lambda \mapsto \mathbf{Q}(\lambda)\mathbf{P}(\lambda)^{-1}$ is completely monotone with respect to $\pi(C)$. Note further that for every $i = 1,\dots,p$ the resolvent $R(\lambda,\hat{\mathbf{A}}_i)$ exists for every $\lambda > 0$ and is completely monotone with respect to $\pi(C)$. Indeed, by assumption $\hat{\mathbf{A}}_i$ is quasi-positive and $\tau(\hat{\mathbf{A}}_i) < 0$. This implies that also $-\lambda\mathbf{I}+\hat{\mathbf{A}}_i$ is quasi-positive and $\tau(-\lambda\mathbf{I}+\hat{\mathbf{A}}_i) < 0$ for every $\lambda \geq 0$. Indeed, note that the operator $-\lambda\mathbf{I}$ is always quasi-positive for $\lambda > 0$, since whenever $\langle u,v\rangle_{n,m} = 0$ we have $\langle -\lambda\mathbf{I}u, v\rangle_{n,m} = -\lambda\langle u,v\rangle_{n,m} = 0$. By an application of Lemma 7.13 it follows that $(\lambda\mathbf{I} - \hat{\mathbf{A}}_i)^{-1} \succeq 0$. Moreover, for every $n \in \mathbb{N}$ we have

$$(-1)^n \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}(\lambda\mathbf{I} - \hat{\mathbf{A}}_i)^{-1} = (\lambda\mathbf{I} - \hat{\mathbf{A}}_i)^{-(1+n)}, \tag{7.47}$$

where the right-hand side of (7.47) is again positive, since $\pi(C)$ is an algebra cone. It thus follows that for every $i = 1,\dots,p$ the linear factor $(\lambda\mathbf{I} - \hat{\mathbf{A}}_i)^{-1}$ in the decomposition (7.46) is completely monotone and by assumption $\mathbf{Q}$ is completely monotone as well. Now the assertion follows since the product of completely monotone functions is again completely monotone (use the general Leibniz's rule here) and hence Theorem 7.16 ii) follows from part i). □

Remark 7.17. i) In the univariate case an analog of Theorem 7.16 i) was shown in [142, Theorem 2]. Here we extend the result to the multivariate setting. However, it is to be noted that the result can be extended even beyond finite-dimensions. In fact, also for certain Hilbert-valued CARMA processes as introduced in [23] a similar characterization can be shown.

ii) The factorization of operator polynomials into the form (7.46) is well studied in the literature, see, e.g. [127] and in particular for matrix polynomials [71]. For instance, a sufficient criteria for $\mathbf{P}(\lambda)$ to admit a factorization of the form (7.46) is that the transition operator $\mathcal{A}_p$ is diagonalizable. If an operator polynomial is factorizable, then the operators $\hat{\mathbf{A}}_i$ can be computed by iterated operator division and the additional positivity conditions can be checked thereafter.

The strength of Theorem 7.16 ii), however, lays in the fact that it provides us with a simple method to construct positive stationary MCARMA processes by choosing suitable operators $\hat{\mathbf{A}}_i$ for $i = 1, \dots, p$. This is explained in the following example by means of a second order MCARMA process:

Example 7.18. Let $n = m = d$ for some $d \in \mathbb{N}$, $p = 2$ and $\mathbf{Q}(\lambda) = \mathbf{I}$. Moreover, let $\hat{\mathbf{A}}_1, \hat{\mathbf{A}}_2 \in \mathcal{L}(\mathbb{M}_d)$ be quasi-positive with $\tau(\hat{\mathbf{A}}_1), \tau(\hat{\mathbf{A}}_2) < 0$. For example, we could choose $\hat{\mathbf{A}}_i x := \hat{A}_i x + x\hat{A}_i^*$ for some matrix $A_i \in \mathbb{M}_d$ with $\tau(\hat{A}_i) < 0$ for $i = 1, 2$. Note that in this case for $i = 1, 2$ we have $\sigma(\hat{\mathbf{A}}_i) = \sigma(\hat{A}_i) + \sigma(\hat{A}_i)$, see [129]. If we set $\mathbf{A}_1 := \hat{\mathbf{A}}_1 + \hat{\mathbf{A}}_2$ and $\mathbf{A}_2 := \hat{\mathbf{A}}_1\hat{\mathbf{A}}_2$, we see that

$$\mathbf{P}(\lambda) = (\lambda\mathbf{I} - \hat{\mathbf{A}}_1)(\lambda\mathbf{I} - \hat{\mathbf{A}}_2) = \lambda^2\mathbf{I} - \lambda\mathbf{A}_1 - \mathbf{A}_2.$$

Hence, following Theorem 7.16 ii) and since $\tau(\mathcal{A}_2) = \tau(\mathbf{A}_2)$, this specification gives rise to a C-positive causal MCAR process of order $p = 2$, whenever $\tau(\mathbf{A}_2) = \tau(\hat{\mathbf{A}}_1\hat{\mathbf{A}}_2) < 0$.

Note that the conditions in Theorem 7.16 ii) are in general not necessary for causal MCARMA processes to be positive. Indeed, in some situations we can check the conditions of Theorem 7.16 i) directly as illustrated by the following example:

Example 7.19. Let $p = 2$, $\mathbf{C}_0 \in \pi(C)$ and $\mathbf{A} \in \mathcal{L}(\mathbb{M}_{n,m})$ be invertible and such that $-\mathbf{A}^2$ is quasi-positive. If we define

$$\mathcal{A}_2 := \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{A}^2 & \mathbf{0} \end{bmatrix}, \quad \text{and} \quad \mathcal{C}_0 := [\mathbf{C}_0, \mathbf{0}],$$

then $\tau(\mathcal{A}_2) = \tau(-\mathbf{A}^2) < 0$ and the associated operator polynomials are $\mathbf{P}(\lambda) = \lambda^2\mathbf{I} + \mathbf{A}^2$ and $\mathbf{Q}(\lambda) = \mathbf{C}_0$. We thus see that $\lambda \mapsto \mathbf{Q}(\lambda)\mathbf{P}(\lambda)^{-1}$ is completely monotone, although no factorization of the form in (7.46) is available.

In the following Proposition 7.20 we extend another sufficient positivity criteria, known in the univariate case in [142, Theorem 1 e)], to $\mathbb{R}_d$-valued MCARMA processes. In order to this, we adapt the proof of [5, Theorem 1] to $\pi(\mathbb{R}_d^+)$-valued rational functions. Unsurprisingly, this multivariate version is more involved and requires some additional and rather technical assumptions.

Proposition 7.20. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ such that $q < p$ and let $(X_t)_{t\geq 0}$ be an $\mathbb{R}_d$-valued causal MCARMA given by (7.43) with parameters $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$. Let $P(\lambda)$ and $Q(\lambda)$ denote the matrix polynomials in (7.16) and (7.17), respectively, and assume that $\{\hat{C}_i \in \mathbb{M}_d \colon i = 0, \ldots, q\}$ and $\{\hat{A}_i \in \mathbb{M}_d \colon i = 1, \ldots, p\}$ are two commutative families of positive matrices such that*

$$Q(\lambda) = \prod_{i=0}^{q} (\lambda \mathbf{I} - \hat{C}_i) \quad \text{and} \quad P(\lambda) = \prod_{i=1}^{p} (\lambda \mathbf{I} - \hat{A}_i). \tag{7.48}$$

Moreover, assume that for every $\lambda > 0$ we have $Q(\lambda)P(\lambda)^{-1} = P(\lambda)^{-1}Q(\lambda)$, the matrix logarithms $\log(P(\lambda))$ and $\log(Q(\lambda))$ exist in $\mathbb{M}_d$ and for every $i = 0, \ldots, q$ there exist $l^{(i,1)}, l^{(i,2)}, \ldots, l^{(i,q+1)} \in \pi(\mathbb{R}_d^+)$ such that for every $n \in \mathbb{N}$ we have $(l^{(i,j+1)})^n \preceq l^{(i,j+1)}$ for all $j = 0, \ldots, q$,

$$\hat{C}_i \preceq \sum_{j=0}^{q} l^{(i,j+1)} \odot \hat{A}_{j+1} \quad \text{and} \quad \sum_{i=0}^{q} l^{(i,j+1)} = \mathbf{1}_d. \tag{7.49}$$

If $p > q+1$, we assume in addition that $\tau(\hat{A}_i) < 0$ for all $i = q+2, \ldots, p$. Then $(X_t)_{t\geq 0}$ is an $\mathbb{R}_d^+$-valued causal MCARMA process of order (p,q) with associated moving average polynomial Q and autoregressive polynomial P.

Proof. Following Theorem 7.16 i), it suffices to show that $\lambda \mapsto Q(\lambda)P(\lambda)^{-1}$ is completely monotone on $\mathbb{R}^+$ with respect to the cone $\pi(\mathbb{R}_d^+)$. By (7.48) we have

$$Q(\lambda)P(\lambda)^{-1} = \left(\prod_{i=0}^{q} (\lambda \mathbf{I} - \hat{C}_i) \right) \left(\prod_{i=1}^{p} (\lambda \mathbf{I} - \hat{A}_i) \right)^{-1}, \quad \forall \lambda > 0.$$

Since the multiplication of monotone functions is again completely monotone it suffices to show the complete monotonicity of $\lambda \mapsto Q(\lambda)\check{P}(\lambda)^{-1}$ for

$$\check{P}(\lambda) = \prod_{i=1}^{q} (\lambda \mathbf{I} - \hat{A}_i), \quad \lambda \in \mathbb{C},$$

since by Theorem 7.16 ii) we conclude that the maps $(\lambda \mathbf{I} - \hat{A}_i)^{-1}$ are completely monotone for all $i = q+2, \ldots, p$.

Moreover, it suffices to prove that the map $\lambda \mapsto \log\big(Q(\lambda)\check{P}(\lambda)^{-1}\big)$ is completely monotone, since for all $\lambda > 0$ we have

$$Q(\lambda)\check{P}(\lambda)^{-1} = \exp\big(\log(Q(\lambda)\check{P}(\lambda)^{-1})\big) = \sum_{n=1}^{\infty} \frac{\big(\log(Q(\lambda)\check{P}(\lambda)^{-1})\big)^n}{n!}. \qquad (7.50)$$

By assumption we know that the matrix logarithms of $Q(\lambda)$ and $\check{P}(\lambda)^{-1}$ exist for every $\lambda > 0$ and moreover the matrices $\hat{C}_0, \ldots, \hat{C}_q$, the matrices $\hat{A}_1, \ldots, \hat{A}_{q+1}$ and the matrix rational function $Q(\lambda)$ and $\check{P}(\lambda)^{-1}$ mutually commute and hence

$$\begin{aligned}\log\big(Q(\lambda)\check{P}(\lambda)^{-1}\big) &= \log\big(Q(\lambda)\big) - \log\big(\check{P}(\lambda)\big) \\ &= \sum_{j=0}^{q} \log\big(\lambda\mathbf{I} - \hat{C}_j\big) - \sum_{i=1}^{q+1} \log\big(\lambda\mathbf{I} - \hat{A}_i\big) \\ &= \int_0^\infty \mathrm{e}^{-\lambda s} s^{-1} \sum_{i=0}^{q} (\mathrm{e}^{s\hat{A}_{i+1}} - \mathrm{e}^{s\hat{C}_i})\,\mathrm{d}s. \qquad (7.51)\end{aligned}$$

From (7.51) we see that is suffices to show that $\sum_{i=0}^{q}(\mathrm{e}^{s\hat{A}_{i+1}} - \mathrm{e}^{s\hat{C}_i}) \succeq 0$ holds for all $s \geq 0$. For this note that the matrix exponential is monotone with respect to $\pi(\mathbb{R}_d^+)$, i.e. for $G_1, G_2 \in \mathbb{M}_d$ such that $G_1 \preceq G_2$ we have $\mathrm{e}^{G_1} \preceq \mathrm{e}^{G_2}$. This can be seen from the definition of the matrix exponential and since monomials are monotone with respect to $\pi(\mathbb{R}_d^+)$. Note further that the matrices $l^{(i,j+1)}$ are in $\pi(\mathbb{R}_d^+)$, i.e. entry-wise non-negative, and the same holds true for $\hat{C}_i$ for $i = 1, \ldots, q+1$ and $j = 0, \ldots, q$. Thus by Lemma 7.28 we have $(l^{(i,j+1)} \odot \hat{C}_i)^n \preceq (l^{(i,j+1)})^n \odot (\hat{C}_i)^n$ for every $n \in \mathbb{N}$ and hence for $i = 1, \ldots, q+1$ and $j = 0, \ldots, q$ we see that

$$\mathrm{e}^{l^{(i,j+1)} \odot \hat{A}_i} = \sum_{n \in \mathbb{N}} \frac{(l^{(i,j+1)} \odot \hat{A}_i)^n}{n!} \preceq \sum_{n \in \mathbb{N}} \frac{(l^{(i,j+1)})^n \odot \hat{A}_i^n}{n!} \preceq l^{(i,j+1)} \odot \sum_{n \in \mathbb{N}} \frac{\hat{A}_i^n}{n!}. \qquad (7.52)$$

This together with assumption (7.49) and the convexity of the matrix exponential imply that for all $s \geq 0$ we have

$$\begin{aligned}\sum_{i=0}^{q} \mathrm{e}^{s\hat{C}_i} \preceq \sum_{i=0}^{q} \mathrm{e}^{s\sum_{j=0}^{q} l^{(i,j+1)} \odot \hat{A}_{j+1}} &\preceq \sum_{i=0}^{q}\sum_{j=0}^{q} l^{(i,j+1)} \odot \mathrm{e}^{s\hat{A}_{j+1}} \\ &= \sum_{j=0}^{q} \Big(\sum_{i=0}^{q} l^{(i,j+1)}\Big) \odot \mathrm{e}^{s\hat{A}_{j+1}} \\ &= \sum_{j=0}^{q} \mathrm{e}^{s\hat{A}_{j+1}}.\end{aligned}$$

Hence, from (7.51) it follows that $\log\big(Q(\lambda)\check{P}(\lambda)^{-1}\big)$ is given by the Laplace transform of the $\pi(\mathbb{R}_d^+)$-valued kernel $s \mapsto s^{-1}\sum_{i=0}^{q}(\mathrm{e}^{s\hat{A}_{i+1}} - \mathrm{e}^{s\hat{C}_i})$, which by Bernstein's theorem yields the complete monotonicity of $\log\big(Q(\lambda)\check{P}(\lambda)^{-1}\big)$ and hence following the reasoning above, we conclude by Theorem 7.16 i) that $(X_t)_{t\in\mathbb{R}}$ is $\mathbb{R}_d^+$-valued. That Q and P are the moving-average, respectively, autoregressive polynomials of $(X_t)_{t\in\mathbb{R}}$ then follows from the commutativity of $Q(\lambda)$ and $P(\lambda)$ and Remark 7.11. □

Remark 7.21. i) Following [78, Theorem 6.4.15 c)] the (real) logarithm of the matrix $P(\lambda)$ exists if and only if $P(\lambda)$ is non-singular and has an even number of Jordan blocks of each size for every negative eigenvalue.

ii) Note that the technical assumption of Proposition 7.20 is best understood when departing from the condition

$$\sum_{i=0}^{q} \hat{A}_{i+1} \succeq \sum_{i=0}^{q} \hat{C}_i. \tag{7.53}$$

Indeed, note that if (7.53) holds for all $1 \leq k, n \leq d$, then we see that $\sum_{i=0}^{q}(\hat{A}_{i+1})_{k,n} \geq \sum_{i=0}^{q}(\hat{C}_i)_{k,n}$. Following the Hardy-Littlewood rearrangement inequality and Hall's marriage theorem, see also [5, Equation 3] and references therein, we see that for all $i = 0, \ldots, q$ and $k, n = 1, \ldots, d$ there exist $(l_{k,n}^{(i,j+1)})_{j=0,\ldots,q}$ with $0 \leq l_{k,n}^{(i,j+1)} \leq 1$ such that for every $1 \leq k, n \leq d$ we have

$$(\hat{C}_i)_{k,n} \leq \sum_{j=0}^{q} l_{k,n}^{(i,j+1)}(\hat{A}_{j+1})_{k,n}$$

and $\sum_{i=0}^{q} l_{k,n}^{(i,j)} = 1$. Hence, setting $l^{(i,j)} = (l_{k,n}^{(i,j)})_{1\leq k,n\leq d}$, we see that the conditions in Proposition 7.20 are met if we assume (7.53) together with $(l^{(i,j)})^n \preceq l^{(i,j)}$ for every $i = 0, \ldots, q$ and $j = 1, \ldots, q+1$ any $n \in \mathbb{N}$. Condition (7.53) is the analogous multivariate condition compared to the univariate case in [142, Theorem 1 e)]. In our case, the additional assumption (7.49) is needed in (7.52). Otherwise, the matrix exponential is not convex with respect to the Hadamard product. Note further that it follows from (7.51), that the stronger, but more accessible, condition

$$\sum_{i=0}^{q}(\mathrm{e}^{s\hat{A}_{i+1}} - \mathrm{e}^{s\hat{C}_i}) \succeq 0 \quad \forall\, s \geq 0$$

is sufficient and could replace the condition in (7.49).

7.3.3 Positive non-stable output processes

In this section we study positive non-stable output and non-causal MCARMA processes. As before, we assume that C is a cone in $\mathbb{M}_{n,m}$ for $n, m \in \mathbb{N}$ and let $L = (L_t)_{t\in\mathbb{R}}$ be a C-valued two-sided Lévy process defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ such that $\mathbb{E}[\log(\|L_1\|_{n,m})] < \infty$. Moreover, we let $\mathcal{A}_p$, E_p and $\mathcal{C}_q$ be as in (7.11)-(7.13) and assume throughout this section that $\tau(\mathcal{A}_p) \geq 0$. Analogous to the stable case in the previous section, we are interested in sufficient and necessary conditions for the positivity of non-stable output processes. Recall that every non-stable output process $(X_t)_{t\in\mathbb{R}}$, associated with $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$, has the representation

$$X_t = \mathcal{C}_q \, \mathrm{e}^{(t-s)\mathcal{A}_p} Z_s + \int_s^t \mathcal{C}_q \, \mathrm{e}^{(t-u)\mathcal{A}_p} E_p \, \mathrm{d}L_s, \quad s < t. \tag{7.54}$$

If, in addition to the above, condition (7.40) is satisfied, then there exists a unique stationary solution to (7.54) which is the associated non-causal MCARMA process. However, in contrast to the causal case, a non-causal MCARMA process does not admit representation (7.43). Indeed, following the proof of Proposition 7.9, we see that the stationary representation of $(\mathrm{vec}(X_t))_{t\in\mathbb{R}}$ in (7.41), and hence also the one of $(X_t)_{t\in\mathbb{R}}$, is considerably more complicated for studying positivity due to the kernel g_2 in decomposition (7.63), which in the non-causal case does not vanish. We therefore assess the positivity of non-stable output processes $(X_t)_{t\in\mathbb{R}}$ given by (7.54) as follows: We look for conditions on the model parameters $\mathbf{A}_1, \dots, \mathbf{A}_p$ and $\mathbf{C}_0, \dots, \mathbf{C}_q$ such that for every $s < t$ the process $(X_t)_{t\geq s}$ is C-valued whenever $Z_s \in C^p$ and $(L_t)_{t\in\mathbb{R}}$ is C-valued.

Remark 7.22. i) In the deterministic control literature a similar type of positivity is often referred to as *internal positivity*, see, e.g. [58]. We call a state space model *internal positive*, if for every non-negative initial value Z_s and every C-valued input process $(L_t)_{t\in\mathbb{R}}$ the output process $(X_t)_{t\geq s}$ assumes only values in C.

ii) Again, we may assume that Z_s is $\mathcal{F}_s$-measurable, such that the process $(X_t)_{t\geq s}$ is adapted to the natural filtration. In this case, however, $(X_t)_{t\geq s}$ is not necessarily an MCARMA process anymore. Moreover, even if a unique stationary solution (MCARMA process) exists, conditions such that $(X_t)_{t\geq s}$ is positive for all positive initial values $Z_s \in C^p$ are in general neither sufficient nor necessary for the positivity of the MCARMA process. Indeed, the stationary distribution of $(Z_t)_{t\in\mathbb{R}}$, if it exists, might not be supported on C^p, see Section 8.2.1 for an example.

Let $J \subseteq \{1, \ldots, p\}$ and denote by $C^{J,p}$ the wedge[1] in $(\mathbb{M}_{n,m})^p$ consisting of the cone C in the in the J-th Cartesian coordinates (and $\mathbb{M}_{n,m}$ otherwise). In particular for $J = \{1, \ldots, p\}$ we have $C^{J,p} = C^p$ and $C^{\{1\},p} = C \times \mathbb{M}_{n,m} \times \cdots \times \mathbb{M}_{n,m}$. To ensure the positivity of the term $\mathcal{C}_q \, \mathrm{e}^{(t-s)\mathcal{A}_p} Z_s$ in (7.54) for every $Z_s \in C^p$, we see that for all $t > s$ the operator $\mathcal{C}_q \, \mathrm{e}^{(t-s)\mathcal{A}_p}$ must map C^p into C. Note that in the case of a MCAR process, i.e. where $\mathcal{C}_0 = [\mathbf{I}, \mathbf{0}, \ldots, \mathbf{0}]$, one would think that $\mathrm{e}^{s\mathcal{A}_p}$ has to map C^p onto $C^{\{1\},p}$ only, as the output operator $\mathcal{C}_q$ only projects down onto the first block matrix component anyway. However, we have the following:

Lemma 7.23. *Let $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$ be as above and let $(X_t)_{t \geq s}$ be the associated non-stable output process in* (7.54)*. Then the following holds true:*

i) If $\mathbf{C}_j \in \pi(C)$ for all $j = 0, \ldots, q$ and $\mathcal{A}_p$ is quasi-positive with respect to C^p, then $(X_t)_{t \geq s}$ is C-valued for all initial values $Z_s \in C^p$ and every C-valued Lévy process $(L_t)_{t \in \mathbb{R}}$.

ii) If $\mathcal{A}_p$ is quasi-positive with respect to $C^{J,p}$ then $J = \{1, 2, \ldots, p\}$.

iii) Conversely, if $(X_t)_{t \geq s}$ is C-valued for all initial values $Z_s \in C^p$ and every C-valued Lévy process $(L_t)_{t \in \mathbb{R}}$, then $\mathbf{C}_j \in \pi(C)$ for all $j = 0, \ldots, q$. If in addition $C = \mathbf{C}_j^{-1}(C)$ holds for all $j = 0, \ldots, q$, then $\mathcal{A}_p$ must be quasi-positive with respect to C^p.

Proof. We proof part i) first. Suppose $\mathcal{A}_p$ is quasi-positive with respect to C^p, then by definition $\mathrm{e}^{t\mathcal{A}_p}\mathbf{z} \in C^p$ for every $\mathbf{z} \in C^p$ and $t \geq 0$. Hence, for $Z_s \in C^p$ we have $\mathrm{e}^{(t-s)\mathcal{A}_p} Z_s \in C^p$ and by definition of a C-positive Lévy process, we have $E_p(L_s - L_{s'}) \in C^p$ for all $s > s'$. This implies that $\int_s^t \mathrm{e}^{(t-s)\mathcal{A}_p} E_p \, \mathrm{d}L_s \in C^p$ for all $t > s$. If moreover $\mathbf{C}_j \in \pi(C)$ for all $j = 0, \ldots, q$ then $\mathcal{C}_q(\mathrm{e}^{(t-s)\mathcal{A}_p} Z_s) \in C$ and $\mathcal{C}_q(\int_s^t \mathrm{e}^{(t-s)\mathcal{A}_p} E_p \, \mathrm{d}L_s) \in C$, which according to (7.54) yields $X_t \in C$ for all $t \geq s$. Next, we show that the particular form of $\mathcal{A}_p$ implies that it can only be quasi-positive on $(\mathbb{M}_{n,m})^p$ with respect to the wedge $C^{J,p}$ if $J = \{1, 2, \ldots, p\}$. Recall that quasi-positivity of $\mathcal{A}_p$ with respect to $C^{J,p}$ can be equivalently characterized by the property that $\langle\langle \mathcal{A}_p \mathbf{x}, \mathbf{u} \rangle\rangle_p \geq 0$, whenever $\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle_p = 0$ with $\mathbf{x}, \mathbf{u} \in C^{J,p}$ where we denote the inner-product on $(\mathbb{M}_{n,m})^p$ by $\langle\langle \cdot, \cdot \rangle\rangle_p$. Therefore, let $\mathbf{x}$, $\mathbf{u} \in C^{J,p}$ and note that

$$\begin{aligned} \langle\langle \mathcal{A}_p \mathbf{x}, \mathbf{u} \rangle\rangle_p &= \langle\langle (x_2, x_3, \ldots, x_p, \sum_{i=1}^p \mathbf{A}_i(x_i))^\intercal, \mathbf{u} \rangle\rangle_p \\ &= \sum_{j=1}^{p-1} \langle x_{j+1}, u_j \rangle_{n,m} + \sum_{i=1}^p \langle \mathbf{A}_{p+1-i}(x_i), u_p \rangle_{n,m}. \end{aligned} \tag{7.55}$$

[1] Recall that a wedge $W \subseteq \mathbb{M}_{n,m}$ is a convex cone without the property $W \cap (-W) = \{\mathbb{0}_{n,m}\}$.

We set $J^c := \{1, 2, \ldots, p\} \setminus J$ and suppose that J^c is not empty. If $p \in J$, let $\mathbf{x}$, $\mathbf{u} \in C^{J,p}$ be such that $u_j = 0$ for every $j \in J$ and $\langle x_i, u_i \rangle_{n,m} = 0$ for $i \in J^c$, then clearly $\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle_p = 0$, but from (7.55) we see that

$$\langle\langle \mathcal{A}_p \mathbf{x}, \mathbf{u} \rangle\rangle_p = \sum_{j \in J^c} \langle x_{j+1}, u_j \rangle_{n,m},$$

which can be negative since we only assumed that $u_j \in \mathbb{M}_{n,m}$ and x_{j+1} can be chosen arbitrary as long as $\langle x_{j+1}, u_{j+1} \rangle_{n,m} = 0$ for $j + 1 \in J^c$. The case $p \in J^c$ then follows by a similar argument and we see that J^c must be empty or otherwise $\mathcal{A}_p$ can not be quasi-positive with respect to $C^{J,p}$. This implies that we must have $J = \{1, 2, \ldots, p\}$.

Lastly, for for the necessary direction in part iii). We suppose that $X_t \in C$ for all $t \geq s$, $Z_s \in C^p$ and every C-valued Lévy process $(L_t)_{t \in \mathbb{R}}$. In particular, at $t = s$ we have

$$X_s = \mathcal{C}_q Z_s = \sum_{i=1}^{q+1} \mathbf{C}_{i-1}(Z_s^{(i)}) \in C$$

for all $Z_s = (Z_s^{(1)}, Z_s^{(2)}, \ldots, Z_s^{(p)}) \in C^p$. Therefore, if we let $z \in C$ be arbitrary and set $z_s^j := e_j \otimes z$ for $j = 1, \ldots, q$, then $z_s^j \in C^p$ and by assumption we must have $X_s = \mathbf{C}_{j-1}(z) \in C$ for all $j = 1, \ldots, q + 1$. Since $z \in C$ was arbitrary we conclude that $\mathbf{C}_{j-1} \in \pi(C)$ for all $j = 1, \ldots, q + 1$. Next, assume that even $C = \mathbf{C}_j^{-1}(C)$ holds for all $j = 0, \ldots, q$ and show that in this case $\mathcal{A}_p$ must be quasi-positive with respect to the cone C^p. For this let $Z_s \in C^p$ be fixed, but arbitrary. Since the constant zero Lévy process is C-valued as well, it follows by assumption that $(X_t)_{t \in \mathbb{R}}$ given by $X_t = \mathcal{C}_q(\mathrm{e}^{(t-s)\mathcal{A}_p} Z_s)$ is C-valued for all $s < t \in \mathbb{R}$ and every $Z_s \in C^p$. Since by the first part $\mathbf{C}_j \in \pi(C)$ and by assumption even $C = \mathbf{C}_j^{-1}(C)$ for all $j = 0, \ldots, q$, we see that for every $s < t$ there exist a $J \subseteq \{1, \ldots, p\}$ such that $\{1, \ldots, q\} \subseteq J$ and $\mathrm{e}^{(t-s)\mathcal{A}_p} Z_s \in C^{J,p}$. But since this holds for every $s < t \in \mathbb{R}$, we find in every neighborhood infinitely many time points $(t_j)_{j \in \mathbb{N}}$ such that $\mathrm{e}^{(t_j - s)\mathcal{A}_p} Z_s \in C^{J,p}$ holds for the same J. From this we conclude that already $\mathrm{e}^{t\mathcal{A}_p}(C^{J,p}) \subseteq C^{J,p}$ must hold for all $t \geq 0$ and some set $J \subseteq \{1, \ldots, p\}$ (indeed, this follows first for all rational time points and extends to irrationals by continuity of the exponential). This, however, by definition means that $\mathcal{A}_p$ is quasi-positive with respect to $C^{J,p}$ and thus we conclude from part ii) that already $J = \{1, 2, \ldots, p\}$ holds, i.e. that $\mathcal{A}_p$ must be quasi-positive with respect to C^p. □

The following theorem is our main result on the internal-positivity of non-stable output processes on $\mathbb{M}_{n,m}$. Under an extra condition on the stationary distribution of the non-stable output process, it also provides a sufficient condition for the positivity of non-causal MCARMA processes.

Theorem 7.24. *Let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ with $q < p$. For $s < t \in \mathbb{R}$, let $(X_t)_{t \geq s}$ be the output process in (7.54), associated with $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$ such that $\tau(\mathcal{A}_p) \geq 0$. Then the following holds true:*

i) If $\mathbf{A}_1 \in \mathcal{L}(\mathbb{M}_{n,m})$ is quasi-positive and $\mathbf{A}_i, \mathbf{C}_j \in \pi(C)$ for $i = 2, \dots, p$ and $j = 0, \dots, q$, then $(X_t)_{t \geq s}$ is C-valued for every initial value $Z_s \in C^p$ and C-valued Lévy process $(L_t)_{t \in \mathbb{R}}$.

ii) If $C = \mathbf{C}_j^{-1}(C)$ for all $j = 0, \dots, q$. Then $(X_t)_{t \geq s}$ is C-valued for every initial value $Z_s \in C^p$ and every C-valued Lévy process $(L_t)_{t \in \mathbb{R}}$ if and only if $\mathbf{A}_1 \in \mathcal{L}(\mathbb{M}_{n,m})$ is quasi-positive and $\mathbf{A}_i \in \pi(C)$ for $i = 2, \dots, p$.

iii) If $\mathbf{C}_j \in \pi(C)$ for all $j = 0, \dots, q$ and there exists a stationary distribution of $(Z_t)_{t \in \mathbb{R}}$ supported on C^p, then the associated non-causal MCARMA process $(X_t)_{t \in \mathbb{R}}$ is C-valued.

Proof. We begin with the proof of part i). By Lemma 7.23 i) it is enough to show that $\mathbf{A}_i \in \pi(C)$ and $\mathbf{A}_1$ is quasi-positive with respect to C implies that $\mathcal{A}_p$ is quasi-positive with respect to C^p. Indeed, let $\mathbf{x} = (x_1, x_2, \dots, x_p)^\intercal \in C^p$ and $\mathbf{u} = (u_1, u_2, \dots, u_p)^\intercal \in C^p$ such that $\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle = 0$ we show that $\langle\langle \mathcal{A}_p \mathbf{x}, \mathbf{u} \rangle\rangle \geq 0$. As before we have

$$\langle\langle \mathcal{A}_p \mathbf{x}, \mathbf{u} \rangle\rangle = \sum_{j=1}^{p-1} \langle x_{j+1}, u_j \rangle_{n,m} + \sum_{i=1}^{p} \langle \mathbf{A}_{p-i+1}(x_i), u_i \rangle_{n,m}, \qquad (7.56)$$

and since $x_{j+1}, u_j \in C$ for all $j = 1, \dots, p-1$, we see that $\langle x_{j+1}, u_j \rangle_{n,m} \geq 0$ and hence the first sum in (7.56) is non-negative. Moreover, we see that $\langle \mathbf{A}_{p-i+1}(x_i), u_i \rangle \geq 0$ for $i = 1, \dots, p-1$ by assumption that $\mathbf{A}_i(C) \subseteq C$ for $i = 2, \dots, p$ and thus the remaining term of the second sum is $\langle \mathbf{A}_1(x_p), u_p \rangle_{n,m}$. By assumption we have $\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle = \sum_{j=1}^{p} \langle x_j, u_j \rangle_{n,m} = 0$, and in particular $\langle x_p, u_p \rangle_{n,m} = 0$, which by the quasi-positivity of $\mathbf{A}_1$ implies $\langle \mathbf{A}_1(x_p), u_p \rangle_{n,m} \geq 0$. Hence, $\langle\langle \mathcal{A}_p \mathbf{x}, \mathbf{u} \rangle\rangle_p \geq 0$ whenever $\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle_p = 0$, which proves the quasi-positivity of $\mathcal{A}_p$ with respect to C^p.
The second assertion is a consequence of Lemma 7.23 iii) and it is only left to prove that the quasi-positivity of $\mathcal{A}_p$ with respect to C^p implies that $\mathbf{A}_1$ is quasi-positive and $\mathbf{A}_i \in \pi(C)$ for $i = 2, \dots, p$. For this, suppose that $\mathcal{A}_p$ is quasi-positive with respect to C^p, then for every $\mathbf{x}$, $\mathbf{u} \in C^p$ with $\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle_p = 0$ we find that the term in (7.56) is non-negative. If we let $u_p, x_1 \in C$ be such that $\langle x_1, u_p \rangle_{n,m} = 0$ and if we set $\mathbf{u} = (0, \dots, 0, u_p)^\intercal$ and $\mathbf{x} = (x_1, 0, \dots, 0)^\intercal$, then we observe that $\langle\langle \mathbf{x}, \mathbf{u} \rangle\rangle_p = 0$ and (7.56) reduces to $\langle \mathbf{A}_1(x_p), u_p \rangle_{n,m} \geq 0$. Since x_1 and u_p satisfy $\langle x_1, u_p \rangle_{n,m} = 0$ but are otherwise arbitrary, it follows that $\mathbf{A}_1$ is quasi-positive. Moreover, we see that $\mathbf{A}_i \in \pi(C)$ for $i = 2, \dots, p$ follows by a similar argument. The third assertion follows immediately from part i) and formula (7.54), since by assumption Z_t is supported on C^p for all $t \in \mathbb{R}$. □

Remark 7.25. Note that the positivity condition in Theorem 7.24 i) only applies in the non-stable case, i.e. for $\tau(\mathcal{A}_p) \geq 0$. Indeed, suppose that $\mathcal{A}_p$ is quasi-positive with respect to C^p and $\tau(\mathcal{A}_p) < 0$, then by an application of Lemma 7.13 to $\mathcal{A}_p$ and the cone C^p implies that the inverse $\mathcal{A}_p^{-1}$ exists and must map positive vectors into negatives, i.e. $\mathcal{A}_p^{-1}(C^p) \subseteq -C^p$. However, the inverse of $\mathcal{A}_p$ is explicitly known as

$$\mathcal{A}_p^{-1} = \begin{bmatrix} \mathbf{A}_p^{-1}\mathbf{A}_{p-1} & \mathbf{A}_p^{-1}\mathbf{A}_{p-2} & \dots & \mathbf{A}_p^{-1} \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{7.57}$$

and obviously never satisfies $\mathcal{A}_p^{-1}(C^p) \subseteq -C^p$. Conversely, note that if $\mathbf{A}_i$ for $i = 1, \dots, p$ satisfies the assumptions in Theorem 7.24 i), then $\mathcal{A}_p$ can not satisfy $\tau(\mathcal{A}_p) < 0$. This is so since $\sigma(\mathcal{A}_p) = \sigma(\mathbf{A}_p)$ and by the Perron-Frobenius theorem, see also [76, Proposition 1], we know that there exists at least one leading eigenvalue of $\mathbf{A}_p$ which is non-negative and hence $\tau(\mathcal{A}_p) \geq 0$.

In the following two corollaries we concretize the positivity criteria in Theorem 7.24 for the case of $\mathbb{R}_d$-valued and $\mathbb{S}_d$-valued non-stable output processes.

Corollary 7.26. *Assume that $m = 1$ and $n = d$ for some $d \in \mathbb{N}$ and let $p \in \mathbb{N}$, $q \in \mathbb{N}_0$ such that $q < p$. Moreover, let $(L_t)_{t\in\mathbb{R}}$ be an $\mathbb{R}_d^+$-valued Lévy process and for $s < t \in \mathbb{R}$ we denote by $(X_t)_{t\geq s}$ the non-stable output process associated with $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$ such that $\tau(\mathcal{A}_p) \geq 0$ and $(A_i)_{i=1,\dots,p}$ and $(C_j)_{j=0,\dots,q}$ satisfy the following conditions:*

i) For every $i = 2, \dots, p$ we have $A_i = (a_{k,n}^{(i)})_{1\leq k,n\leq d}$ such that $a_{k,n}^{(i)} \geq 0$ for all $1 \leq k, n \leq d$;

ii) $A_1 = (a_{k,n}^{(1)})_{1\leq k,n\leq d}$ is such that $a_{k,n}^{(1)} \geq 0$ for all $1 \leq k, n \leq d$ with $k \neq n$;

iii) For every $j = 0, \dots, q$ we have $C_j = (c_{k,n}^{(j)})_{1\leq k,n\leq d}$ such that $c_{k,n}^{(j)} \geq 0$ for all $1 \leq k, n \leq d$.

Then $(X_t)_{t\geq s}$ is $\mathbb{R}_d^+$-valued, whenever $Z_s \in \mathbb{R}_{pd}^+$.

We can visualize the condition of Corollary 7.26 on $\mathcal{A}_p \in \mathbb{M}_{pd}$ as follows:

$$\mathcal{A}_p = \begin{pmatrix} \mathbb{0}_d & \mathbb{I}_d & \mathbb{0}_d & \cdots & \mathbb{0}_d \\ \mathbb{0}_d & \mathbb{0}_d & \mathbb{I}_d & & \vdots \\ \vdots & & & \ddots & \vdots \\ \vdots & & \ddots & & \mathbb{0}_d \\ \mathbb{0}_d & \cdots & \cdots & \mathbb{0}_d & \mathbb{I}_d \\ \underbrace{\begin{matrix} a^{(p)}_{1,1} \dots a^{(p)}_{1,d} \\ \vdots \quad \ddots \quad \vdots \\ a^{(p)}_{d,1} \dots a^{(p)}_{d,d} \end{matrix}}_{a^{(p)}_{k,n} \geq 0,\ \forall k,n} & \underbrace{\begin{matrix} a^{(p-1)}_{1,1} \dots a^{(p-1)}_{1,d} \\ \vdots \quad \ddots \quad \vdots \\ a^{(p-1)}_{d,1} \dots a^{(p-1)}_{d,d} \end{matrix}}_{a^{(p-1)}_{k,n} \geq 0,\ \forall k,n} & \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix} & \underbrace{\begin{matrix} a^{(1)}_{1,1} \dots a^{(1)}_{1,d} \\ \vdots \quad \ddots \quad \vdots \\ a^{(1)}_{d,1} \dots a^{(1)}_{d,d} \end{matrix}}_{a^{(1)}_{k,n} \geq 0,\ \forall k \neq n} \end{pmatrix}$$

We obtain an analogous result for $\mathbb{S}_d^+$-valued non-stable output processes:

Corollary 7.27. *Assume that $m = n = d$ for some $d \in \mathbb{N}$ and let $p \in \mathbb{N}$, $q \in \mathbb{N}_0$ such that $q < p$. For $s < t \in \mathbb{R}$ we denote by $(X_t)_{t \geq s}$ the output process associated with $(\mathcal{A}_p, E_p, \mathcal{C}_q, L)$ such that $\tau(\mathcal{A}_p) \geq 0$ and assume that $(L_t)_{t \in \mathbb{R}}$ is an $\mathbb{S}_d^+$-valued Lévy process. If $(\mathbf{A}_i)_{i=1,\dots,p}$ and $(\mathbf{C}_j)_{j=0,\dots,q}$ satisfy the following conditions:*

i) For every $i = 2, \dots, p$ there exists $a_i \in \mathbb{M}_d$ such that $\mathbf{A}_i(x) = a_i x a_i^*$ for every $x \in \mathbb{S}_d$;

ii) There exists an $a_1 \in \mathbb{M}_d$ such that $\mathbf{A}_1(x) = a_1 x + x a_1^*$ for every $x \in \mathbb{S}_d$;

iii) For every $j = 0, \dots, q$ there exists $c_j \in \mathbb{M}_d$ such that $\mathbf{C}_j(x) = c_j x c_j^*$ for every $x \in \mathbb{S}_d$.

Then $(X_t)_{t \geq 0}$ is $\mathbb{S}_d^+$-valued, whenever $Z_s \in (\mathbb{S}_d^+)^p$. Moreover, the state-space representation of $(\mathrm{vec}(X_t))_{t \geq 0}$ in (7.21)-(7.22) holds with operator $\mathcal{A}_p^{\mathrm{vec}}$ given by

$$\mathcal{A}_p^{\mathrm{vec}} := \begin{pmatrix} \mathbb{0}_{d^2} & \mathbb{I}_{d^2} & \mathbb{0}_{d^2} & \cdots & \mathbb{0}_{d^2} \\ \mathbb{0}_{d^2} & \mathbb{0}_{d^2} & \mathbb{I}_{d^2} & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ \vdots & & \ddots & & \mathbb{0}_{d^2} \\ \mathbb{0}_{d^2} & \cdots & \cdots & \mathbb{0}_{d^2} & \mathbb{I}_{d^2} \\ a_p \otimes a_p & a_{p-1} \otimes a_{p-1} & \cdots & \cdots & \mathbb{I}_d \otimes a_1 + a_1 \otimes \mathbb{I}_d \end{pmatrix},$$

output operator $\mathcal{C}_q^{\mathrm{vec}}$ given by $\mathcal{C}_q^{\mathrm{vec}} = [c_0 \otimes c_0, c_1 \otimes c_1, \dots, c_q \otimes c_q, \mathbb{0}_{d^2}, \dots, \mathbb{0}_{d^2}]$ and the input operator E_p^{vec} given by $E_p^{\mathrm{vec}} = e_p \otimes \mathbb{I}_{d^2}$.

Proof. This follows from Theorem 7.24 and the fact that maps of the form $x \mapsto axa^*$ for $a \in \mathbb{M}_d$ are in $\pi(\mathbb{S}_d^+)$ and maps of the form $x \mapsto ax + xa^*$ are quasi-positive. Moreover, note that

$$A_1^{\mathrm{vec}}(x) = \big(\mathbb{I}_d \otimes a_1 + a_1 \otimes \mathbb{I}_d\big)\,\mathrm{vec}(x)$$

and $A_i^{\mathrm{vec}}(x) = a_i \otimes a_i\,\mathrm{vec}(x)$ for $i = 2, \ldots, p$ and the analogous assertions hold for the operators C_j for $j = 0, \ldots, q$. □

7.4 Proof: Existence of stationary solutions

The existence of $\tilde{Q}, \tilde{P} \in \mathbb{M}_{nm}(\mathbb{R}[\lambda])$ such that (7.39) holds with $\det(\tilde{P}(\lambda)) = 0$ if and only if $\det(\hat{P}(\lambda)) = 0$ follows immediately from [89, Lemma 6.3-8] and Lemma 7.8. From (7.33) it then follows that $(\tilde{P}, \tilde{Q})$ is a left-matrix fraction description of the transfer function $H(\lambda) = \hat{\mathcal{C}}_q^{\mathrm{vec}}(\lambda\mathbf{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{E}_p^{\mathrm{vec}}$, i.e.

$$\hat{\mathcal{C}}_q^{\mathrm{vec}}(\lambda\mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{E}_p^{\mathrm{vec}} = \tilde{P}(\lambda)^{-1}\tilde{Q}(\lambda), \quad \forall\lambda \in \mathbb{C}. \tag{7.58}$$

We define the two kernels $g_1, g_2 \colon \mathbb{R} \to \mathcal{L}(\mathbb{R}_{nm}, \mathbb{R}_{nmp})$ by

$$g_i(t) \coloneqq \frac{1}{2\pi\,\mathrm{i}} \int_{\rho_i} \mathrm{e}^{\lambda t}(\lambda\mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{E}_p^{\mathrm{vec}}\,\mathrm{d}\lambda, \quad t \in \mathbb{R} \text{ and } i = 1, 2, \tag{7.59}$$

where we integrate anti-clockwise over simple closed curves ρ_1 and ρ_2 in the open left- and right-half plane of the complex field encircling the zeroes of the map $\lambda \mapsto (\lambda\mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{E}_p^{\mathrm{vec}}$. More specifically, let ρ_1 be the rectangle in the left-half plane with width M and height $2R$ (with M and R large enough such that all eigenvalues in the left-half plain are encircled) and such that the line segment

$$\rho_{\mathrm{i}R} \coloneqq \{\lambda \in \mathbb{C} \colon \Re(\lambda) = 0 \text{ and } |\Im(\lambda)| \leq R\},$$

forms an edge of ρ_1. We then define the curve ρ_2 as the reflection of ρ_1 over the imaginary axis. Moreover, for $i = 1, 2$ we denote by $\hat{\rho}_i$ the curve ρ_i without the line segment on the imaginary axis.
The complex integrals in (7.59) are well defined, since by assumption (7.40) there is no singularity of $\lambda \mapsto (\lambda\mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{E}_p^{\mathrm{vec}} = (\mathbb{I}_{nm}, \lambda\mathbb{I}_{nm}, \ldots, \lambda^p\mathbb{I}_{nm})^\intercal\tilde{P}(\lambda)^{-1}$ on the imaginary axis. For every $t \in \mathbb{R}$ we set

$$Z_t \coloneqq \int_{-\infty}^{t} g_1(t-u)\,\mathrm{d}L_u - \int_t^{\infty} g_2(t-u)\,\mathrm{d}L_u. \tag{7.60}$$

Next, we show that the integrals over the kernels g_1 and g_2 are well defined as the limit of integrals

$$\lim_{T\to\infty}\int_{-T}^{t} g_1(t-u)\,\mathrm{d}L_u \quad \text{and} \quad \lim_{T\to\infty}\int_{t}^{T} g_2(t-u)\,\mathrm{d}L_u.$$

Moreover, we show that $(Z_t)_{t\in\mathbb{R}}$ is the unique stationary solution of (7.21). First, note that for $i=1,2$ the kernel g_i satisfies the equation $\frac{\mathrm{d}}{\mathrm{d}t}g_i(t)=\hat{\mathcal{A}}_p^{\mathrm{vec}}g_i(t)$, which can be seen by similar arguments as in Lemma 7.8. Indeed, note that we have

$$\hat{\mathcal{A}}_p^{\mathrm{vec}}g_i(t)=\frac{1}{2\pi\,\mathrm{i}}\int_{\rho_i}\mathrm{e}^{\lambda t}(\lambda\mathbb{I}_{pnm}-\hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{\mathcal{A}}_p^{\mathrm{vec}}\hat{E}_p^{\mathrm{vec}}\,\mathrm{d}\lambda,$$

and the term $(\lambda\mathbb{I}_{pnm}-\hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}\hat{\mathcal{A}}_p^{\mathrm{vec}}\hat{E}_p^{\mathrm{vec}}$ can be computed by solving the following linear matrix equation

$$(\lambda\mathbb{I}_{pnm}-\hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1}F=\hat{\mathcal{A}}_p^{\mathrm{vec}}\hat{E}_p^{\mathrm{vec}},$$

for $F:=(F_1,F_2,\dots,F_p)^{\intercal}\in\mathcal{L}(\mathbb{R}_{nm},\mathbb{R}_{pnm})$ with $F_i\in\mathbb{M}_{nm}$ for all $i=1,\dots,p$. Since $\hat{\mathcal{A}}_p^{\mathrm{vec}}\hat{E}_p^{\mathrm{vec}}=(\mathbb{0}_{nm},\dots,\mathbb{0}_{nm},K^{-(n,m)},\hat{A}_1^{\mathrm{vec}})$, we can argue similarly to the proof of Lemma 7.8 and obtain

$$F=(\lambda\mathbb{I}_{nm},\dots,\lambda^{p+1}(K^{(n,m)})^p\mathbb{I}_{nm})^{\intercal}\hat{P}(\lambda)^{-1}-(\mathbb{0}_{nm},\dots,\mathbb{0}_{nm},K^{(n,m)})^{\intercal}.$$

Note that by integrating over the closed curves ρ_i for $i=1,2$, the integral over the latter term vanishes and hence for all $t\in\mathbb{R}$ we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}g_i(t)&=\frac{1}{2\pi\,\mathrm{i}}\int_{\rho_i}\mathrm{e}^{\lambda t}(\lambda\mathbb{I}_{nm},\lambda^2K^{(n,m)},\dots,\lambda^{p+1}(K^{(n,m)})^p\mathbb{I}_{nm})^{\intercal}\hat{P}(\lambda)^{-1}\,\mathrm{d}\lambda\\&=\hat{\mathcal{A}}_p^{\mathrm{vec}}g_i(t),\quad i=1,2.\end{aligned}$$

Since the homogeneous linear equation $\frac{\mathrm{d}}{\mathrm{d}t}v(t)=\hat{\mathcal{A}}_p^{\mathrm{vec}}v(t)$ is uniquely solved by $\mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}}v_0$, we see that $g_i(t)=\mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}}g_i(0)$ for $i=1,2$ and $t\in\mathbb{R}$. This has the following consequences: First, it follows that there exist $K>0$ and $\delta>0$ such that for all $u\le 0$ we have $\|g_1(-u)\|_{\mathcal{L}(\mathbb{M}_{n,m},(\mathbb{M}_{n,m})^p)}\le K\,\mathrm{e}^{-\delta|u|}$ and for all $u\ge 0$ we have $\|g_2(-u)\|_{\mathcal{L}(\mathbb{M}_{n,m},(\mathbb{M}_{n,m})^p)}\le K\,\mathrm{e}^{-\delta|u|}$. This together with $\mathbb{E}\,[\log(\|L_1\|_{nm})]<\infty$ implies the existence of the integrals $\int_{-\infty}^{t}g_1(t-u)\,\mathrm{d}L_u$ and $\int_t^{\infty}g_2(t-u)\,\mathrm{d}L_u$, respectively, as limits of integrals over the intervals $(-T,t]$, resp. $[t,T)$, for $T\to\infty$, see also [36]. Next, we show that $(Z_t)_{t\in\mathbb{R}}$ is a solution of (7.21).

Another consequence from the representation $g_i(t) = \mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}} g_i(0)$ for $i = 1, 2$ and $t \in \mathbb{R}$ is, that for every $s < t \in \mathbb{R}$ the following equality holds true:

$$\begin{aligned}\mathrm{e}^{(t-s)\hat{\mathcal{A}}_p^{\mathrm{vec}}} Z_s &= \mathrm{e}^{(t-s)\hat{\mathcal{A}}_p^{\mathrm{vec}}} \Big(\int_{-\infty}^{s} g_1(s-u)\,\mathrm{d}L_u - \int_s^\infty g_2(s-u)\,\mathrm{d}L_u \Big) \\ &= \mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}} \Big(\int_{-\infty}^{s} g_1(-u)\,\mathrm{d}L_u - \int_s^\infty g_2(-u)\,\mathrm{d}L_u \Big). \qquad (7.61)\end{aligned}$$

Now, by setting $\rho = \rho_1 + \rho_2$, the spectral representation of the matrix exponential $\mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}}$, see e.g. [103, Theorem 17.5], yields

$$\mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}} \hat{E}_p^{\mathrm{vec}} = \frac{1}{2\pi\,\mathrm{i}} \int_\rho \mathrm{e}^{\lambda t} (\lambda \mathbb{I}_{pnm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1} \hat{E}_p^{\mathrm{vec}}\,\mathrm{d}\lambda, \quad \forall t \in \mathbb{R},$$

and hence, for every $t \in \mathbb{R}$ we have

$$\int_s^t \mathrm{e}^{(t-u)\hat{\mathcal{A}}_p^{\mathrm{vec}}} \hat{E}_p^{\mathrm{vec}}\,\mathrm{d}L_u = \mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}} \Big(\int_s^t g_1(-u)\,\mathrm{d}L_u + \int_s^t g_2(-u)\,\mathrm{d}L_u \Big). \qquad (7.62)$$

By summing up (7.61) and (7.62) we obtain

$$\mathrm{e}^{t\hat{\mathcal{A}}_p^{\mathrm{vec}}} \Big(\int_{-\infty}^{t} g_1(-u)\,\mathrm{d}L_u - \int_t^\infty g_2(-u)\,\mathrm{d}L_u \Big) = Z_t,$$

which proves that $(Z_t)_{t\in\mathbb{R}}$ is a solution of (7.21). Moreover, it is easy to see that $(Z_t)_{t\in\mathbb{R}}$ is also stationary and unique in law. Now, let us set $Y_t := \hat{\mathcal{C}}_q^{\mathrm{vec}}(Z_t)$ and $h(t) := g_1(t)\mathbf{1}_{[0,\infty)}(t) - g_2(t)\mathbf{1}_{(-\infty,0)}(t)$ for all $t \in \mathbb{R}$, then

$$Y_t = \int_{-\infty}^{\infty} \hat{\mathcal{C}}_q^{\mathrm{vec}} h(t-u)\,\mathrm{d}L_u = \int_{-\infty}^{\infty} \Big(\frac{1}{2\pi\,\mathrm{i}} \int_\rho \mathrm{e}^{\lambda(t-u)} \hat{\mathcal{C}}_q^{\mathrm{vec}} (\lambda \mathbb{I}_{nm} - \hat{\mathcal{A}}_p^{\mathrm{vec}})^{-1} \hat{E}_p^{\mathrm{vec}}\,\mathrm{d}\lambda \Big)\,\mathrm{d}L_u,$$

which by (7.58) equals

$$Y_t = \int_{-\infty}^{\infty} \Big(\frac{1}{2\pi\,\mathrm{i}} \int_\rho \mathrm{e}^{\lambda(t-u)} \tilde{P}^{-1}(\lambda)\tilde{Q}(\lambda)\,\mathrm{d}\lambda \Big)\,\mathrm{d}L_u, \quad t \in \mathbb{R}, \qquad (7.63)$$

and if

$$\int_{-\infty}^{\infty} \Big(\frac{1}{2\pi\,\mathrm{i}} \int_\rho \mathrm{e}^{\lambda(t-u)} \tilde{P}^{-1}(\lambda)\tilde{Q}(\lambda)\,\mathrm{d}\lambda \Big) = \int_{-\infty}^{\infty} g(t-u)\,\mathrm{d}L_u, \quad t \in \mathbb{R}, \qquad (7.64)$$

then by uniqueness we conclude that $\mathrm{vec}(X_t) = Y_t$ for $t \in \mathbb{R}$ and that the representation (7.41) holds true. It is therefore left to prove that the identity in (7.64) holds true for every $t \in \mathbb{R}$.

In order to prove this, we set $K(z,t) \coloneqq \mathrm{e}^{-zt}\tilde{P}(z)^{-1}\tilde{Q}(z)$ and use the integration paths from above, i.e. $\rho_i = \hat{\rho}_i \pm \rho_{\mathrm{i}R}$. We then compute the left hand-side in (7.64) as follows:

$$
\begin{aligned}
h(t) &= g_1(t)\mathbf{1}_{[0,\infty)}(t) - g_2(t)\mathbf{1}_{(-\infty,0)}(t) \\
&= \frac{1}{2\pi\,\mathrm{i}}\Big(\int_{\rho_1} K(z,t)\,\mathrm{d}z\mathbf{1}_{[0,\infty)}(t) - \int_{\rho_2} K(z,t)\,\mathrm{d}z\mathbf{1}_{(-\infty,0)}(t)\Big) \\
&= \frac{1}{2\pi\,\mathrm{i}}\Big(\int_{\hat{\rho}_1} K(z,t)\,\mathrm{d}z\mathbf{1}_{[0,\infty)}(t) - \int_{\hat{\rho}_2} K(z,t)\,\mathrm{d}z\mathbf{1}_{(-\infty,0)}(t)\Big) \\
&\quad + \frac{1}{2\pi\,\mathrm{i}}\Big(\int_{-\mathrm{i}R}^{\mathrm{i}R} K(z,t)\,\mathrm{d}z\mathbf{1}_{[0,\infty)}(t) - \int_{\mathrm{i}R}^{-\mathrm{i}R} K(z,t)\,\mathrm{d}z\mathbf{1}_{(-\infty,0)}(t)\Big) \\
&= \frac{1}{2\pi\,\mathrm{i}}\Big(\int_{\hat{\rho}_1} K(z,t)\,\mathrm{d}z\mathbf{1}_{[0,\infty)}(t) - \int_{\hat{\rho}_2} K(z,t)\,\mathrm{d}z\mathbf{1}_{(-\infty,0)}(t)\Big) \\
&\quad + \frac{1}{2\pi}\Big(\int_{-R}^{R} K(\mathrm{i}\xi,t)\,\mathrm{d}\xi\mathbf{1}_{[0,\infty)}(t) + \int_{-R}^{R} K(\mathrm{i}\xi,t)\,\mathrm{d}\xi\mathbf{1}_{(-\infty,0)}(t)\Big) \\
&= \frac{1}{2\pi}\int_{-R}^{R} K(\mathrm{i}\xi,t)\,\mathrm{d}\xi + \frac{1}{2\pi\,\mathrm{i}}\Big(\int_{\hat{\rho}_1} K(z,t)\,\mathrm{d}z\mathbf{1}_{[0,\infty)}(t) - \int_{\hat{\rho}_2} K(z,t)\,\mathrm{d}z\mathbf{1}_{(-\infty,0)}(t)\Big).
\end{aligned}
$$

Now, by letting $R \to \infty$, we see that for every $t \in \mathbb{R}$ the first term converges to $g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} K(\mathrm{i}\xi,t)\,\mathrm{d}\xi$ and it remains to show that the latter term converges to zero. For this, we first consider the term $\int_{\hat{\rho}_1} K(z,t)\,\mathrm{d}z\mathbf{1}_{[0,\infty)}(t)$ for $t \geq 0$ and note that the integral over $\hat{\rho}_1$ can be split into three separate integrals: The first is

$$
\int_{\mathrm{i}R-M}^{-\mathrm{i}R-M} \mathrm{e}^{\lambda t}\tilde{P}(\lambda)^{-1}\tilde{Q}(\lambda)\,\mathrm{d}\lambda = \int_{-R}^{R} \mathrm{e}^{-\mathrm{i}\xi t}\,\mathrm{e}^{-tM}\tilde{P}(-\mathrm{i}\xi - M)^{-1}\tilde{Q}(-\mathrm{i}\xi - M)\,\mathrm{d}\xi,
$$

where the term e^{-tM} in the integral dictates the convergence to zero as $M \to \infty$ for arbitrary R. The two other terms are given by

$$
\int_{\mathrm{i}R}^{\mathrm{i}R-M} \mathrm{e}^{\lambda t}\tilde{P}(\lambda)^{-1}\tilde{Q}(\lambda)\,\mathrm{d}\lambda = -\int_{0}^{M} \mathrm{e}^{(\mathrm{i}R-\zeta)t}\tilde{P}(\mathrm{i}R-\zeta)^{-1}\tilde{Q}(\mathrm{i}R-\zeta)\,\mathrm{d}\zeta,
$$

and

$$
\int_{-\mathrm{i}R-M}^{-\mathrm{i}M} \mathrm{e}^{\lambda t}\tilde{P}(\lambda)^{-1}\tilde{Q}(\lambda)\,\mathrm{d}\lambda = \int_{0}^{-M} \mathrm{e}^{(-\mathrm{i}R+\zeta)t}\tilde{P}(-\mathrm{i}R+\zeta)^{-1}\tilde{Q}(-\mathrm{i}R+\zeta)\,\mathrm{d}\zeta,
$$

where in both cases the integrals exists for arbitrary large M and since $p > q$ implies $\tilde{P}(-\mathrm{i}R+\zeta)^{-1}\tilde{Q}(-\mathrm{i}R+\zeta) \to 0$ as $R \to \infty$.

We see that both integrals converge to zero as $R \to \infty$. Similarly, for the integral over $\hat{\rho}_2$ and all $t < 0$, we see that

$$\int_{-\mathrm{i}R+M}^{\mathrm{i}R+M} \mathrm{e}^{\lambda t} \tilde{P}(\lambda)^{-1} \tilde{Q}(\lambda) \,\mathrm{d}\lambda = \int_{-R}^{R} \mathrm{e}^{-\mathrm{i}\xi t} \,\mathrm{e}^{tM} \tilde{P}(-\mathrm{i}\xi - M)^{-1} \tilde{Q}(-\mathrm{i}\xi - M) \,\mathrm{d}\xi,$$

which, as before, converges to zero as $M \to \infty$ for every $R \geq 0$ since $t < 0$ and the matrix exponential is the dominating term. For the remaining parts, we have

$$\int_{\mathrm{i}R+M}^{\mathrm{i}R} \mathrm{e}^{\lambda t} \tilde{P}(\lambda)^{-1} \tilde{Q}(\lambda) \,\mathrm{d}\lambda = -\int_{0}^{-M} \mathrm{e}^{(\mathrm{i}R-\zeta)t} \tilde{P}(\mathrm{i}R - \zeta)^{-1} \tilde{Q}(\mathrm{i}R - \zeta) \,\mathrm{d}\zeta,$$

and

$$\int_{-\mathrm{i}R}^{-\mathrm{i}R+M} \mathrm{e}^{\lambda t} \tilde{P}(\lambda)^{-1} \tilde{Q}(\lambda) \,\mathrm{d}\lambda = -\int_{0}^{M} \mathrm{e}^{-(\mathrm{i}R-\zeta)t} \tilde{P}(\zeta - \mathrm{i}R)^{-1} \tilde{Q}(\zeta - \mathrm{i}R) \,\mathrm{d}\zeta,$$

and we see that for every $M > 0$ and $t < 0$, since $p > q$ the integrals converge to zero, whenever $R \to \infty$. Thus the only remaining term of $h(t)$ when expanding the integration domain is $g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(\mathrm{i}\xi, t) \,\mathrm{d}\xi$ which proves (7.64) and we conclude the assertions of Proposition 7.9.

7.5 Auxiliary result

In this section we only prove a submultiplicativity property of the Hadamard product of two non-negative matrices. This auxiliary results was used in the proof of Proposition 7.20.

Lemma 7.28. *Let $d \in \mathbb{N}$ and $A, B \in \mathbb{M}_d$ be non-negative matrices, i.e. $A, B \in \pi(\mathbb{R}_d^+)$. Then for all $n \in \mathbb{N}$ we have $(A \odot B)^n \preceq A^n \odot B^n$.*

Proof. We write $A = (a_{i,j})_{1 \leq i,j \leq d}$ and $B = (b_{i,j})_{1 \leq i,j \leq d}$ and prove the statement by induction over $n \in \mathbb{N}$. For $n = 1$ the statement is trivial. In case of $n = 2$, it follows from the non-negativity of $(a_{i,j})_{1 \leq i,j \leq d}$ and $(b_{i,j})_{1 \leq i,j \leq d}$ that for every $i, j \in \{1, \ldots, d\}$ we have

$$((A \odot B)^2)_{i,j} = \sum_{k=1}^{d} a_{i,k} b_{i,k} a_{k,j} b_{k,j} \leq \Big(\sum_{k=1}^{d} a_{i,k} a_{j,k}\Big)\Big(\sum_{k=1}^{d} b_{i,k} b_{j,k}\Big) = ((A^2 \odot B^2))_{i,j},$$

which proves the statement for $n = 2$. Now, suppose that $(A \odot B)^n \preceq A^n \odot B^n$ for some $n \in \mathbb{N}$. We show that also $(A \odot B)^{n+1} \preceq A^{n+1} \odot B^{n+1}$ holds.

Indeed, we have $((A \odot B)^{n+1})_{i,j} = ((A \odot B)(A \odot B)^n)_{i,j} \leq ((A \odot B)A^n \odot B^n)_{i,j}$, where we note that $((A \odot B)A^n \odot B^n)_{i,j} = \sum_{k=1}^{d} a_{i,k}(A^n)_{k,j} b_{i,k}(B^n)_{k,j}$ and hence

$$\begin{aligned}((A \odot B)^{n+1})_{i,j} &\leq \sum_{k=1}^{d} a_{i,k}(A^n)_{k,j} b_{i,k}(B^n)_{k,j} \\ &\leq \big(\sum_{k=1}^{d} a_{i,k}(A^n)_{k,j}\big)\big(\sum_{k=1}^{d} b_{i,k}(B^n)_{k,j}\big) = (A^{n+1} \odot B^{n+1})_{i,j},\end{aligned}$$

which yields the desired inequality $(A \odot B)^{n+1} \preceq A^{n+1} \odot B^{n+1}$. □

CHAPTER 8

Stochastic Covariance Models based on positive semi-definite MCARMA

Abstract of the chapter In this chapter we discuss the capability of positive semi-definite MCARMA processes to model the instantaneous covariance process in multivariate stochastic covariance models. This extends ideas known in the univariate setting [27] to the multivariate setting, and we demonstrate the potential of higher-order MCARMA based models to capture a great variety of short-memory effects observed in realized variance and cross-covariance processes in applications. We introduce the class of matrix-valued well-balanced Ornstein-Uhlenbeck processes as a particular example of an matrix-valued MCARMA process of order two and justify the relevance of higher-order MCARMA based stochastic covariance models by an exemplary analysis of the second order moment structure of positive semi-definite well-balanced Ornstein-Uhlenbeck based models.

This chapter is based on [17, Section 4]:
Benth, F., and Karbach, S.
Positive multivariate continuous-time autoregressive moving-average processes, 2022, DOI: arXiv.2206.08782.

8.1 Introduction

In this chapter we are concerned with multivariate stochastic covariance models based on positive semi-definite MCARMA processes as introduced and studied in Chapter 7. More precisely, we consider a model consisting of a d-dimensional (logarithmic)-price process together with an instantaneous covariance process given by an MCARMA process with values in the cone of positive semi-definite $d \times d$-matrices. Classical choices for the instantaneous covariance process in multivariate stochastic covariance models are either pure-jump based models, such as OU type models and superposition of these, so called supOU processes, see [10, 126], pure diffusion-based models such as the Wishart processes, see [31], or mixtures of both, e.g. affine jump-diffusion models, see [104, 42]. We propose to use MCARMA processes on symmetric and positive semi-definite matrices to model the instantaneous covariance process in multivariate stochastic volatility models. For one thing, the popular multivariate extension of the Barndorff–Nielsen and Shepard (BNS) stochastic volatility model in [126] is included in the much broader class of positive semi-definite MCARMA based models. In addition, the nuanced short memory structure of the MCARMA class has the potential to explain phenomena like non-monotone or sub-exponentially decaying auto-correlation functions that are often observed in realized (co)variance time-series of financial data. We demonstrate the capability of MCARMA based stochastic volatility models by means of *positive semi-definite well-balanced Ornstein-Uhlenbeck processes*. This class stems from a particular configuration of MCARMA processes on positive semi-definite matrices and we provide a detailed analysis of it in Section 8.3.

8.1.1 Layout of the chapter

This chapter is structured as follows: In Section 8.2 we begin with a brief introduction to multivariate stochastic covariance models. The theory is in large parts analogous to the thoroughly discussed infinite-dimensional case. In particular, we recall some results on the second order moment structure of the return and covariance processes in Section 8.2. Subsequently, in Section 8.2.1, we recall the multivariate Barndorff–Nielsen and Shepard (BNS) volatility model from [125] as a particular example of an MCAR based stochastic covariance model of order one. We motivate the use of higher-order MCARMA based stochastic covariance models to capture certain short-memory effects observed in realized variance and cross-covariance time-series. In Section 8.3, we demonstrate the gained flexibility in the second order structure when using higher-order MCARMA processes through an exemplary analysis of *positive semi-definite well-balanced Ornstein-Uhlenbeck processes.*

8.2 Second order moment structure of stochastic covariance models

Let $d \in \mathbb{N}$. We call an $\mathbb{R}_d \times \mathbb{S}_d^+$-valued process $(Y_t, X_t)_{t\geq 0}$ a *stochastic covariance model* on $\mathbb{R}_d$, whenever it consists of a *instantaneous covariance process* $(X_t)_{t\geq 0}$ on $\mathbb{S}_d^+$ and a d-dimensional (logarithmic) asset price process $(Y_t)_{t\geq 0}$ given by a stochastic differential equation of the form

$$\begin{cases} \mathrm{d}Y_t = (\alpha + X_t\beta)\,\mathrm{d}t + X_t^{1/2}\,\mathrm{d}W_t, \quad t > 0, \\ Y_0 = y \in \mathbb{R}_d, \end{cases} \tag{8.1}$$

where $\alpha \in \mathbb{R}_d$ is the drift, $\beta \in \mathbb{R}_d$ the risk-premium, and $(W_t)_{t\geq 0}$ denotes a $\mathbb{R}_d$-valued (standard) Brownian motion, see [12, 125]. The instantaneous covariance process $(X_t)_{t\geq 0}$ is assumed to be integrable with respect to $(W_t)_{t\geq 0}$ and $\mathbb{S}_d^+$-valued such that for all $t \geq 0$ the matrix square-root $X_t^{1/2}$ exists, and the stochastic integral in (8.1) is well-defined. Our goal is to have a preferably flexible class of instantaneous covariance process $(X_t)_{t\geq 0}$ such that the stochastic covariance model $(Y_t, X_t)_{t\geq 0}$ represents the stylized facts of financial data, while being sufficiently tractable for, e.g. simulations, statistical inference or option pricing. In contrast to the logarithmic price process $(Y_t)_{t\geq 0}$, the instantaneous covariance process $(X_t)_{t\geq 0}$ is not directly observable in markets. However, it can be measured indirectly from the *realized covariance* of the (squared) return process as follows: First, assume that $(X_t)_{t\geq 0}$ is square-integrable and stationary and for any real positive number Δ we define the discrete-time process $(Y_n^\Delta)_{n\in\mathbb{N}}$ by

$$Y_n^\Delta := Y_{n\Delta} - Y_{(n-1)\Delta} = \int_{(n-1)\Delta}^{n\Delta} (\alpha + X_t\beta)\,\mathrm{d}t + \int_{(n-1)\Delta}^{n\Delta} X_t^{1/2}\,\mathrm{d}W_t, \quad n \in \mathbb{N}.$$

The process $(Y_n^\Delta)_{n\in\mathbb{N}}$ is the sequence of logarithmic returns over the intervals $[(n-1)\Delta, n\Delta]$ and it can be seen that for every $n \in \mathbb{N}$ the random variable Y_n^Δ given X_n^Δ is normal distributed. More precisely, for every $n \in \mathbb{N}$ we have

$$Y_n^\Delta | X_n^\Delta \sim \mathcal{N}(\alpha\Delta + X_n^\Delta\beta, X_n^\Delta) \quad \text{with } X_n^\Delta := \int_{(n-1)\Delta}^{n\Delta} X_t\,\mathrm{d}t = X_{n\Delta}^+ - X_{(n-1)\Delta}^+,$$

where $X_t^+ := \int_0^t X_t\,\mathrm{d}s$ for $t \geq 0$, see also [9]. We define the auto-covariance function $\mathrm{acov}_{X_t} := \mathbb{R}^+ \to \mathcal{L}(\mathbb{M}_d)$ of the process $(X_t)_{t\geq 0}$ at $t \geq 0$ by $\mathrm{acov}_{X_t}(h) := \mathrm{Cov}[\mathrm{vec}(X_{t+h}), \mathrm{vec}(X_t)]$ where $h \geq 0$; Similarly, we define the auto-covariances $\mathrm{acov}_{X_n^\Delta}$ and $\mathrm{acov}_{Y_n^\Delta (Y_n^\Delta)^\intercal}(h)$ for the processes $Y_n^\Delta (Y_n^\Delta)^\intercal$ and $(X_n^\Delta)_{n\in\mathbb{N}}$, respectively, in which case we restrict to $n, h \in \mathbb{N}$.

Following [125], we observe that the second order structure of the instantaneous covariance process $(X_t)_{t\geq 0}$, respectively its discrete difference process $(X_n^\Delta)_{n\in\mathbb{N}}$, is inherited by the second order structure of the squared logarithmic return process $(Y_n^\Delta (Y_n^\Delta)^\intercal)_{n\in\mathbb{N}}$ such that

$$\mathrm{acov}_{Y_n^\Delta (Y_n^\Delta)^\intercal}(h) = \mathrm{acov}_{X_n^\Delta}(h), \quad \text{for all } h, n \in \mathbb{N}. \tag{8.2}$$

Moreover, we recall from [125, Theorem 3.2] that for general square-integrable stationary instantaneous covariance processes $(X_t)_{t\geq 0}$ we have

$$\mathrm{acov}_{X_n^\Delta}(h) = r^{++}(h\Delta + \Delta) - 2r^{++}(h\Delta) + r^{++}(h\Delta - \Delta), \quad \forall h, n \in \mathbb{N}, \tag{8.3}$$

where $r^{++}\colon \mathbb{R}^+ \to \mathcal{L}(\mathbb{M}_d)$ is given by

$$r^{++}(t) \coloneqq \int_0^t \int_0^s \mathrm{acov}_{X_n}(u)\,\mathrm{d}u\,\mathrm{d}s, \quad t \geq 0. \tag{8.4}$$

As we noted before, a prominent feature observed in many realized (co)variance time-series is the *memory effect*, which means that the observed auto-covariances of the (squared) returns exhibits sub-exponential decay or even non-monotone configurations, see, e.g. [27, 28] and the reference therein. To capture the memory effect in the short-time lags, we propose to model the instantaneous covariance process $(X_t)_{t\geq 0}$ in (8.1) by $\mathbb{S}_d^+$-valued higher-order MCARMA processes as they were introduced in Chapter 7. In the univariate case the idea to model the instantaneous variance process in stochastic volatility models by positive (higher-order) CARMA processes goes back to [27]. To demonstrate the potential of MCARMA based stochastic covariance models, we compare the two auto-covariances in (8.2) and (8.3) in the following two cases: First, for the multivariate BNS model in Section 8.2.1, which was studied in [125]. Secondly, for a stochastic covariance model based on a *positive semi-definite well-balanced OU* processes which we introduce in Section 8.3 below.

8.2.1 Ornstein-Uhlenbeck type processes on positive semi-definite matrices

Stochastic covariance models based on positive semi-definite OU type processes were studied extensively in [10, 126, 125] and we refer to it as the (multivariate) BNS stochastic volatility model. The class of matrix-valued OU processes is included in the class of matrix-valued MCARMA processes and form the class of MCAR processes of order one, see Definition 7.10. We show that the positivity criteria in Theorem 7.16 in case of a causal MCAR process of order one coincides with the known (sufficient) criteria for positive semi-definite OU processes in [10].

Let $p = 1$, $q = 0$ and set $\mathcal{A}_1 = -\mathbf{A}$ for some $\mathbf{A} \in \mathcal{L}(\mathbb{M}_d)$ such that (7.40) is satisfied and let $\mathcal{C}_q = \mathbf{I}$. Moreover, let L be a two-sided Lévy process on $\mathbb{M}_d$ and assume that $\mathbb{E}[\log(\|L_1\|_{d^2})] < \infty$. We denote the MCAR process associated with the state space representation $(\mathcal{A}_1, \mathbf{I}, \mathbf{I}, L)$ by $(X_t)_{t\in\mathbb{R}}$. This process is an Ornstein-Uhlenbeck type process on $\mathbb{M}_d$ and has the following representation:

$$X_t = \mathrm{e}^{-(t-s)\mathbf{A}} Z_s + \int_s^t \mathrm{e}^{-(t-u)\mathbf{A}}\,\mathrm{d}L_u, \quad s < t \in \mathbb{R}.$$

If we further assume that $\tau(-\mathbf{A}) < 0$, then $X_t = \int_{-\infty}^t \mathrm{e}^{-(t-s)\mathbf{A}}\,\mathrm{d}L_s$ is the unique stationary OU process adapted to the natural filtration of $(L_t)_{t\in\mathbb{R}}$. It follows from Theorem 7.16 ii), that whenever $(L_t)_{t\in\mathbb{R}}$ is $\mathbb{S}_d^+$-increasing, a sufficient condition for $(X_t)_{t\in\mathbb{R}}$ to be $\mathbb{S}_d^+$-valued is the quasi-positivity of the operator $-\mathbf{A}$ with respect to $\mathbb{S}_d^+$. By recalling the definition of quasi-positivity this means $\mathrm{e}^{-t\mathbf{A}}(\mathbb{S}_d^+) \subseteq \mathbb{S}_d^+$ for all $t \geq 0$ and we see that this criteria coincides with the usual sufficient condition for OU processes to be symmetric and positive semi-definite, see [10, Proposition 4.1].
In the following we recall some statistical properties of $\mathbb{S}_d^+$-valued stationary OU process that follow from Propositions 7.4 and 7.7 and [10, Proposition 4.7]: For all $t \geq 0$ the mean of X_t is given by

$$\mathbb{E}[X_t] = \mathbf{A}^{-1}\mu_L = \mathbf{A}^{-1}\Big(\gamma_L + \int_{\mathbb{S}_d^+\cap\{\|\xi\|_d>1\}} \xi\,\nu^L(\mathrm{d}\xi)\Big), \tag{8.5}$$

and the auto-covariance of $(X_t)_{t\geq 0}$ at $t \geq 0$ is given by

$$\operatorname{Cov}[\operatorname{vec}(X_t), \operatorname{vec}(X_{t+h})] = \mathrm{e}^{-hA^{\mathrm{vec}}}\mathcal{D}^{-1}Q^{\mathrm{vec}}, \quad h \geq 0, \tag{8.6}$$

where $\mathcal{Q}^{\mathrm{vec}} = \operatorname{vec}\circ\mathcal{Q}\circ\operatorname{vec}^{-1}$ and $\mathcal{D} \in \mathcal{L}(\mathbb{M}_{d^2})$ is given by

$$\mathcal{D}(X) = A^{\mathrm{vec}}X + X(A^{\mathrm{vec}})^{\intercal} \quad \text{with } A^{\mathrm{vec}} = \operatorname{vec}\circ\mathbf{A}\circ\operatorname{vec}^{-1}. \tag{8.7}$$

Moreover, we have

$$\int_0^t X_t\,\mathrm{d}t = -\mathbf{A}^{-1}(X_t - X_0 - L_t), \quad t \in \mathbb{R}, \tag{8.8}$$

and it follows from [125] that

$$r^{++}(t) = \Big((A^{\mathrm{vec}})^{-2}\big(\mathrm{e}^{-A^{\mathrm{vec}}t} - \mathbb{I}_{d^2}\big) + (A^{\mathrm{vec}})^{-1}t\Big)\mathcal{D}^{-1}\mathcal{Q}^{\mathrm{vec}}, \quad t \geq 0, \tag{8.9}$$

which by (8.2) and (8.3) for every $n \in \mathbb{N}$ yield

$$\operatorname{acov}_{Y_n^\Delta (Y_n^\Delta)^\intercal}(h) = \mathrm{e}^{-A^{\mathrm{vec}}\Delta(h-1)}(A^{\mathrm{vec}})^{-2}\big(\mathbb{I}_{d^2} - \mathrm{e}^{-A^{\mathrm{vec}}\Delta}\big)^2\mathcal{D}^{-1}\mathcal{Q}^{\mathrm{vec}}, \quad h \in \mathbb{N}. \tag{8.10}$$

8.3 Well-balanced Ornstein-Uhlenbeck processes on positive semi-definite matrices

In this section we introduce *matrix-valued well-balanced OU processes* that extend the univariate well-balanced OU processes from [136] to $\mathbb{S}_d$-valued processes. The authors in [136] proposed positive well-balanced OU processes as a model for the instantaneous variance process in stochastic volatility models. Here, we extend this idea to the multivariate setting by studying $\mathbb{S}_d^+$-valued well-balanced OU processes and show that this class is well-suited to model the instantaneous covariance process in multivariate stochastic covariance models. Moreover, we show that stochastic covariance models based on well-balanced OU processes exhibit auto-covariance functions of the squared logarithmic returns Y_n^Δ that have slower decay compared to the multivariate BNS model.
Let $\mathbf{A} \in \mathcal{L}(\mathbb{S}_d)$ with $\tau(-\mathbf{A}) < 0$ and define $\mathcal{A}_2 \in \mathcal{L}((\mathbb{M}_d)^2)$, $E_2 \in \mathcal{L}(\mathbb{M}_d, (\mathbb{M}_d)^2)$ and $\mathcal{C}_0 \in \mathcal{L}((\mathbb{M}_d)^2, \mathbb{M}_d)$ by

$$\mathcal{A}_2 := \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{A}^2 & \mathbf{0} \end{bmatrix}, \quad E_2 := \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \quad \text{and} \quad \mathcal{C}_0 := [-2\mathbf{A}, \mathbf{0}]. \tag{8.11}$$

Moreover, let $(L_t)_{t\in\mathbb{R}}$ be a square-integrable $\mathbb{S}_d^+$-increasing Lévy process with expectation $\mu_L \in \mathbb{S}_d^+$ and covariance operator $\mathcal{Q} \in \mathcal{L}(\mathbb{S}_d)$. Let $(X_t)_{t\geq 0}$ denote the output process of the state space model associated with $(\mathcal{A}_2, E_2, \mathcal{C}_0, L)$ and initial value Z_0, where Z_0 is a random element in $(\mathbb{S}_d)^2$ (not necessarily $\mathcal{F}_0$-measurable). The output process $(X_t)_{t\geq 0}$ is given by

$$X_t = \mathcal{C}_0 \,\mathrm{e}^{t\mathcal{A}_2} Z_0 + \int_0^t \mathcal{C}_0 \,\mathrm{e}^{(t-s)\mathcal{A}_2} E_2 \,\mathrm{d}L_s, \quad t \geq 0, \tag{8.12}$$

and we note that since $\sigma(\mathcal{A}_2) = \sigma(\mathbf{A}^2) \subseteq \mathbb{R}^+ \setminus \{0\} + \mathrm{i}\mathbb{R}$ the process $(X_t)_{t\geq 0}$ is non-stable, but (7.40) is satisfied and thus following Proposition 7.9, there exist a unique stationary solution to (8.12). Following our terminology from Definition 7.10, this stationary process is a non-causal MCARMA processes of order $(2, 0)$ and in the following proposition we define *positive semi-definite well-balanced OU processes* as the unique stationary and positive semi-definite solution to (8.12). Moreover, we specify the stationary distribution and present sufficient positivity conditions for $(X_t)_{t\geq 0}$.

Proposition 8.1. *Let $(\mathcal{A}_2, E_2, \mathcal{C}_0, L)$ be as above and in addition assume that $-\mathbf{A}$ is quasi-positive. Let $(X_t)_{t\geq 0}$ be as in (8.12) with $Z_0 = (Z_0^{(1)}, Z_0^{(2)})$ given by $Z_0^{(1)} = -\mathbf{A}^{-1}(\pi_1 + \pi_2)/2$ and $Z_0^{(2)} = (\pi_1 - \pi_2)/2$ where*

$$\pi_1 := \int_{-\infty}^0 \mathrm{e}^{s\mathbf{A}} \,\mathrm{d}L_s \quad \textit{and} \quad \pi_2 := \int_0^\infty \mathrm{e}^{-s\mathbf{A}} \,\mathrm{d}L_s.$$

Then $(X_t)_{t\geq 0}$ is stationary, $\mathbb{S}_d^+$-valued and can be represented as

$$X_t = \int_{-\infty}^{t} \mathrm{e}^{-(t-s)\mathbf{A}}\,\mathrm{d}L_s + \int_{t}^{\infty} \mathrm{e}^{-(s-t)\mathbf{A}}\,\mathrm{d}L_s, \quad t \geq 0. \tag{8.13}$$

Moreover, for all $t \geq 0$ we have

$$\mathbb{E}\left[X_t\right] = 2\mathbf{A}^{-1}\mu_L, \tag{8.14}$$

and the auto-covariance of $(X_t)_{t\geq 0}$ at $t \geq 0$ is given by

$$\begin{aligned}\operatorname{Cov}\left[\operatorname{vec}(X_{t+h}), \operatorname{vec}(X_t)\right] = \mathrm{e}^{-hA^{\mathrm{vec}}}\big(\mathrm{e}^{h\hat{\mathcal{D}}} - \mathbf{I}\big)\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}} \\ + 2\,\mathrm{e}^{-hA^{\mathrm{vec}}}\mathcal{D}^{-1}\mathcal{Q}^{\mathrm{vec}}, \quad h \geq 0,\end{aligned} \tag{8.15}$$

where $\hat{\mathcal{D}} \in \mathcal{L}(\mathbb{S}_d)$ is given by $\hat{\mathcal{D}}(X) = A^{\mathrm{vec}}X - X(A^{\mathrm{vec}})^\intercal$, $\mathcal{D}$ is as in (8.7) *and $A^{\mathrm{vec}} := \operatorname{vec}\circ\mathbf{A}\circ\operatorname{vec}^{-1}$. We call the process $(X_t)_{t\geq 0}$ in* (8.13) *a* positive semi-definite well-balanced OU process.

Proof. From the particular anti-diagonal form of $\mathcal{A}_2$ we see that for every $k \in \mathbb{N}$ the following holds true:

$$\mathcal{A}_2^{2k} = \begin{bmatrix}(\mathbf{A}^2)^k & \mathbf{0}\\ \mathbf{0} & (\mathbf{A}^2)^k\end{bmatrix} \quad \text{and} \quad \mathcal{A}_2^{2k+1} = \begin{bmatrix}\mathbf{0} & (\mathbf{A}^2)^k\\ (\mathbf{A}^2)^{k+1} & \mathbf{0}\end{bmatrix},$$

which gives

$$\mathrm{e}^{t\mathcal{A}_2} = \sum_{k=0}^{\infty}\frac{(t\mathcal{A}_2)^{2k}}{(2k)!} + \sum_{k=0}^{\infty}\frac{(t\mathcal{A}_2)^{2k+1}}{(2k+1)!} = \begin{bmatrix}\cosh(t\mathbf{A}) & \sinh(t\mathbf{A})\mathbf{A}^{-1}\\ \sinh(t\mathbf{A})\mathbf{A} & \cosh(t\mathbf{A})\end{bmatrix}, \quad t \geq 0,$$

where $\cosh(t\mathbf{A}) := \sum_{k=0}^{\infty}\frac{(t\mathbf{A})^{2k}}{2k!}$ and $\sinh(t\mathbf{A}) := \sum_{k=0}^{\infty}\frac{(t\mathbf{A})^{2k+1}}{(2k+1)!}$. Note further, that $\sinh(\mathbf{A})$, $\cosh(\mathbf{A})$ and $\mathbf{A}$ all commute mutually. Hence by (8.12) we obtain

$$\begin{aligned}X_t &= [-2\mathbf{A}, \mathbf{0}]\begin{pmatrix}\cosh(t\mathbf{A})Z_0^{(1)} + \sinh(t\mathbf{A})\mathbf{A}^{-1}Z_0^{(2)}\\ \sinh(t\mathbf{A})\mathbf{A}Z_0^{(1)} + \cosh(t\mathbf{A})Z_0^{(2)}\end{pmatrix}\\ &\quad + \int_0^t [-2\mathbf{A}, \mathbf{0}]\begin{pmatrix}\sinh((t-s)\mathbf{A})\mathbf{A}^{-1}\,\mathrm{d}L_s\\ \cosh((t-s)\mathbf{A})\,\mathrm{d}L_s\end{pmatrix}\\ &= -2\mathbf{A}\cosh(t\mathbf{A})Z_0^{(1)} - 2\sinh(\mathbf{A})Z_0^{(2)} - 2\int_0^t \sinh((t-s)\mathbf{A})\,\mathrm{d}L_s\\ &= -\mathbf{A}\,\mathrm{e}^{t\mathbf{A}}Z_0^{(1)} - \mathbf{A}\,\mathrm{e}^{-t\mathbf{A}}Z_0^{(1)} + X_t^{(1)} - X_t^{(2)},\end{aligned} \tag{8.16}$$

where in the last line (8.16) we used that $\cosh(t\mathbf{A}) = \frac{1}{2}(\exp(t\mathbf{A}) + \exp(-t\mathbf{A}))$ and $\sinh(t\mathbf{A}) = \frac{1}{2}(\exp(t\mathbf{A}) - \exp(-t\mathbf{A}))$, set $X_t^{(1)} := \mathrm{e}^{-t\mathbf{A}}Z_0^{(2)} + \int_0^t \mathrm{e}^{-(t-s)\mathbf{A}}\,\mathrm{d}L_s$, for $t \geq 0$, as well as $X_t^{(2)} := \mathrm{e}^{t\mathbf{A}}Z_0^{(2)} + \int_0^t \mathrm{e}^{(t-s)\mathbf{A}}\,\mathrm{d}L_s$.

Now, recall $\pi_1 = \int_{-\infty}^0 \mathrm{e}^{s\mathbf{A}}\,\mathrm{d}L_s$ and $\pi_2 = \int_0^\infty \mathrm{e}^{-s\mathbf{A}}\,\mathrm{d}L_s$, respectively, and inserting the initial state $Z_0^{(1)} = -\mathbf{A}^{-1}\frac{1}{2}(\pi_1+\pi_2)$ and $Z_0^{(2)} = \frac{1}{2}(\pi_1-\pi_2)$ into (8.16) yields

$$\begin{aligned} X_t &= \mathrm{e}^{t\mathbf{A}}\pi_2 + \int_0^t \mathrm{e}^{-(t-s)\mathbf{A}}\,\mathrm{d}L_s + \mathrm{e}^{-t\mathbf{A}}\pi_1 - \int_0^t \mathrm{e}^{(t-s)\mathbf{A}}\,\mathrm{d}L_s \\ &= \int_0^\infty \mathrm{e}^{(t-s)\mathbf{A}}\,\mathrm{d}L_s + \int_0^t \mathrm{e}^{-(t-s)\mathbf{A}}\,\mathrm{d}L_s + \int_{-\infty}^0 \mathrm{e}^{-(t-s)\mathbf{A}}\,\mathrm{d}L_s - \int_0^t \mathrm{e}^{(t-s)\mathbf{A}}\,\mathrm{d}L_s \\ &= \int_{-\infty}^t \mathrm{e}^{-\mathbf{A}(t-s)}\,\mathrm{d}L_s + \int_t^\infty \mathrm{e}^{-\mathbf{A}(s-t)}\,\mathrm{d}L_s, \end{aligned} \tag{8.17}$$

which proves representation (8.13). We set $\tilde{X}_t^{(1)} = \int_{-\infty}^t \mathrm{e}^{-\mathbf{A}(t-s)}\,\mathrm{d}L_s$ and $\tilde{X}_t^{(2)} = \int_t^\infty \mathrm{e}^{-\mathbf{A}(s-t)}\,\mathrm{d}L_s$. By assumption $-\mathbf{A}$ is quasi-positive and hence it follows from Section 8.2.1 that both $\tilde{X}_t^{(1)}$ and $\tilde{X}_t^{(2)}$ are positive semi-definite for all $t \geq 0$. Thus, we proved that $(X_t)_{t\geq 0}$ is stationary, positive semi-definite and possesses the representation (8.13). We continue with the computation of the expectation and auto-covariance. It is easy to see that for all $t \geq 0$ we have

$$\mathbb{E}\left[\tilde{X}_t^{(1)}\right] = \mathbb{E}\left[\tilde{X}_t^{(2)}\right] = \mathbf{A}^{-1}\mu_L,$$

which implies (8.14). It is left to prove that the auto-covariance of $(X_t)_{t\geq 0}$ at $t \geq 0$ satisfies (8.15). For this, we recall that for all $t, h \geq 0$ we have

$$\mathrm{Cov}\left[\mathrm{vec}(X_{t+h}), \mathrm{vec}(X_t)\right] = \mathbb{E}\left[\mathrm{vec}(X_{t+h})\,\mathrm{vec}(X_t)^\intercal\right] - \mathbb{E}\left[\mathrm{vec}(X_{t+h})\right]\mathbb{E}\left[\mathrm{vec}(X_t)^\intercal\right],$$

where according to (8.14) and linearity of the expectation we see that

$$\mathbb{E}\left[\mathrm{vec}(X_{t+h})\right] = 2(A^{\mathrm{vec}})^{-1}\mathbb{E}\left[\mathrm{vec}(L_1)\right],$$

and $\mathbb{E}\left[\mathrm{vec}(X_t)^\intercal\right] = 2\mathbb{E}\left[\mathrm{vec}(L_1)\right]^\intercal\left((A^{\mathrm{vec}})^\intercal\right)^{-1}$. Thus by (8.13), we are left with the terms

$$\mathbb{E}\left[\mathrm{vec}(\tilde{X}_{t+h}^{(i)})\,\mathrm{vec}(\tilde{X}_t^{(j)})^\intercal\right], \quad i, j = 1, 2. \tag{8.18}$$

Note that for $i = j = 1$ this term is the auto-covariance of the stationary OU type process $(\tilde{X}_t^{(1)})_{t\geq 0}$ adjusted by the following outer-square of the expectation:

$$\mathbb{E}\left[\mathrm{vec}(\tilde{X}_{t+h}^{(1)})\right]\mathbb{E}\left[\mathrm{vec}(\tilde{X}_t^{(1)})^\intercal\right] = (A^{\mathrm{vec}})^{-1}\mathbb{E}\left[\mathrm{vec}(L_1)\right]\mathbb{E}\left[\mathrm{vec}(L_1)^\intercal\right]\left((A^{\mathrm{vec}})^{-1}\right))^\intercal.$$

Note that following (8.6), the auto-covariance of the process $(\mathrm{vec}(\tilde{X}_t^{(1)}))_{t\geq 0}$ is given by

$$\mathrm{Cov}\left[\mathrm{vec}(\tilde{X}_{t+h}^{(1)}), \mathrm{vec}(\tilde{X}_t^{(1)})\right] = \mathrm{e}^{-hA^{\mathrm{vec}}}\mathcal{D}^{-1}\mathcal{Q}^{\mathrm{vec}}, \quad h \geq 0.$$

Similarly, for $i = j = 2$ a straightforward computation shows that the auto-covariance of $(\tilde{X}_t^{(2)})_{t\geq 0}$ is the same as the auto-covariance of $(\tilde{X}_t^{(1)})_{t\geq 0}$. For $i = 2$ and $j = 1$, we see that by the definition of a two-sided Lévy process we have

$$\mathrm{Cov}\left[\mathrm{vec}(\tilde{X}_{t+h}^{(2)}), \mathrm{vec}(\tilde{X}_t^{(1)})\right] = 0.$$

Hence, we are left with the last term in (8.18), that is $i = 1$ and $j = 2$. Note first that for every $h \geq 0$ we have

$$\mathrm{vec}(\tilde{X}_{t+h}^{(1)})\,\mathrm{vec}(\tilde{X}_t^{(2)})^\intercal = \int_{-\infty}^{t+h} \mathrm{e}^{-(t+h-u)A^{\mathrm{vec}}}\,\mathrm{d}\,\mathrm{vec}(L_u) \int_t^{\infty} \mathrm{d}\,\mathrm{vec}(L_s)^\intercal\,\mathrm{e}^{-(s-t)(A^{\mathrm{vec}})^\intercal}.$$

By using the independent increments property of $(L_t)_{t\in\mathbb{R}}$ and by a change of variable, we see that

$$\begin{aligned}
\mathbb{E}\left[\mathrm{vec}(\tilde{X}_{t+h}^{(1)})\,\mathrm{vec}(\tilde{X}_t^{(2)})^\intercal\right] &= \int_{-\infty}^{h}\int_0^{\infty} \mathrm{e}^{-(h-u)A^{\mathrm{vec}}}\mathbb{E}\left[\mathrm{d}\,\mathrm{vec}(L_u)\,\mathrm{d}\,\mathrm{vec}(L_s)^\intercal\right]\mathrm{e}^{-s(A^{\mathrm{vec}})^\intercal} \\
&= \mathrm{e}^{-A^{\mathrm{vec}}h}\int_0^h \mathrm{e}^{sA^{\mathrm{vec}}}\,\mathcal{Q}^{\mathrm{vec}}\,\mathrm{e}^{-s(A^{\mathrm{vec}})^\intercal}\,\mathrm{d}s \\
&\quad + (A^{\mathrm{vec}})^{-1}\mathbb{E}\left[\mathrm{vec}(L_1)\right]\mathbb{E}\left[\mathrm{vec}(L_1)^\intercal\right]\left((A^{\mathrm{vec}})^\intercal\right)^{-1} \\
&= \mathrm{e}^{-hA^{\mathrm{vec}}}\int_0^h \mathrm{e}^{s\hat{\mathcal{D}}}\,\mathcal{Q}^{\mathrm{vec}}\,\mathrm{d}s \\
&\quad + (A^{\mathrm{vec}})^{-1}\mathbb{E}\left[\mathrm{vec}(L_1)\right]\mathbb{E}\left[\mathrm{vec}(L_1)^\intercal\right]((A^{\mathrm{vec}})^{-1}))^\intercal.
\end{aligned}$$

The integral in the last line can be computed as

$$\int_0^h \mathrm{e}^{s\hat{\mathcal{D}}}\,\mathcal{Q}^{\mathrm{vec}}\,\mathrm{d}s = \hat{\mathcal{D}}^{-1}\left(\mathrm{e}^{h\hat{\mathcal{D}}} - \mathbf{I}\right)\mathcal{Q}^{\mathrm{vec}}.$$

Hence, by collecting all the terms in (8.18) and since $\mathcal{D}$, $\hat{\mathcal{D}}$ and $\exp(\mathbf{A})$ mutually commute, we obtain (8.15). □

Remark 8.2. Proposition 8.1 shows that the positivity criteria for non-stable state space models from Theorem 7.24 is indeed not necessary. As noted before, this happens, as for stationary processes (such as the well-balanced OU) it would suffice to ensure the positivity only for its stationary states and not, as in Theorem 7.24, for all positive initial values Z_0. Moreover, note that the stationary distribution does not have to be supported on the positive cone $(\mathbb{S}_d^+)^p$ for the output process $(X_t)_{t\geq 0}$ to be positive.

8.3.1 Second order structure of positive semi-definite well-balanced OU based stochastic covariance models

In this section we study the second order structure of stochastic covariance models with instantaneous covariance process given by a positive semi-definite well-balanced OU process. In particular, we compare the obtained auto-covariance of the squared (logarithmic)-return process with the corresponding auto-covariance in (8.10) in the multivariate BNS model. In the next lemma we compute the function r^{++} from (8.9) for the positive semi-definite well-balanced OU process:

Lemma 8.3. *Let $\mathbf{A} \in \mathcal{L}(\mathbb{M}_d)$ such that $-\mathbf{A}$ is quasi-positive and $\tau(-\mathbf{A}) < 0$ and denote by $(X_t)_{t\geq 0}$ the associated positive semi-definite well-balanced OU process as in Proposition 8.1. For every $t \geq 0$ we have:*

$$r^{++}(t) = \Big((A^{\mathrm{vec}})^{-2}\,\mathrm{e}^{-tA^{\mathrm{vec}}}\big(\mathcal{G}\,\mathrm{e}^{t\hat{\mathcal{D}}} - \mathbf{I}\big) - \mathcal{D}_1 t + \mathcal{D}_2\Big)\,\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}} + 2r^{++}_{OU}(t), \quad (8.19)$$

where $r^{++}_{OU}(t)$ is as in (8.9)*, $\mathcal{G} \in \mathcal{L}(\mathbb{M}_{d^2})$ is defined by $\mathcal{G}(x) := (A^{\mathrm{vec}})^2 x\big((A^{\mathrm{vec}})^\intercal\big)^{-2}$ and $\mathcal{D}_i \in \mathcal{L}(\mathbb{M}_{d^2})$ for $i = 1, 2$ is given by $\mathcal{D}_i(x) := (A^{\mathrm{vec}})^{-i}x - x\big((A^{\mathrm{vec}})^\intercal\big)^{-i}$.*

Proof. Let $T \geq 0$, then by definition of r^{++} in (8.4) we have to compute

$$r^{++}(t) = \int_0^t \int_0^s \mathrm{acov}_{X_T}(u)\,\mathrm{d}u\,\mathrm{d}s, \quad t \geq 0,$$

for $\mathrm{acov}_{X_T}(u) = \mathrm{Cov}\,[X_{T+u}, X_T]$ given by (8.15). By (8.6) we see that for every $u \geq 0$, the auto-covariance function $\mathrm{acov}_{X_T}(u)$ is the sum of two times the auto-covariance function of the classical OU process and $\mathrm{e}^{-hA^{\mathrm{vec}}}\big(\mathrm{e}^{h\hat{\mathcal{D}}} - \mathbf{I}\big)\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}}$. Thus we obtain

$$\begin{aligned}
r^{++}(t) &= \int_0^t \int_0^s \mathrm{e}^{-uA^{\mathrm{vec}}}\big(\mathrm{e}^{u\hat{\mathcal{D}}} - \mathbf{I}\big)\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}}\,\mathrm{d}u\,\mathrm{d}s + 2r^{++}_{OU}(t) \\
&= \int_0^t -\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}}\big((A^{\mathrm{vec}})^\intercal\big)^{-1}\big(\mathrm{e}^{-s(A^{\mathrm{vec}})^\intercal} - \mathbb{I}_{d^2}\big) \\
&\qquad + (A^{\mathrm{vec}})^{-1}\big(\mathrm{e}^{-sA^{\mathrm{vec}}} - \mathbb{I}_{d^2}\big)\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}}\,\mathrm{d}s + 2r^{++}_{OU}(t) \\
&= \hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}}\big((A^{\mathrm{vec}})^\intercal\big)^{-2}(\mathrm{e}^{-t(A^{\mathrm{vec}})^\intercal} - \mathbb{I}_{d^2}) - (A^{\mathrm{vec}})^{-2}(\mathrm{e}^{-tA^{\mathrm{vec}}} - \mathbb{I}_{d^2})\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}} \\
&\qquad + \hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}}\big((A^{\mathrm{vec}})^\intercal\big)^{-1}t - (A^{\mathrm{vec}})^{-1}\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}} + 2r^{++}_{OU}(t) \\
&= \big((A^{\mathrm{vec}})^{-2}\,\mathrm{e}^{-tA^{\mathrm{vec}}}\big(\mathcal{G}\,\mathrm{e}^{t\hat{\mathcal{D}}} - \mathbf{I}\big) - \mathcal{D}_1 t + \mathcal{D}_2\big)\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}} + 2r^{++}_{OU}(t).
\end{aligned}$$

□

8.3.2 Heuristics on the second order structure of positive semi-definite well-balanced OU based models

In the following we give some heuristics on the flexibility of the second order structure of positive semi-definite well-balanced OU processes in comparison with the classical OU type processes. From Lemma 8.3 and (8.3) we see that the auto-covariance function of the squared (logarithmic)-returns (8.2) in the positive semi-definite well-balanced OU based stochastic covariance model for $h \in \mathbb{N}$ has the following form:

$$\begin{aligned}\mathrm{acov}_{Y_n^\Delta (Y_n^\Delta)^\intercal}(h) &= \mathrm{e}^{-A^{\mathrm{vec}}\Delta(h-1)}(A^{\mathrm{vec}})^{-2}(\mathcal{G}\,\mathrm{e}^{\Delta(h-1)\hat{D}} - \mathbf{I})(\mathbb{I}_{d^2} - \mathrm{e}^{-A^{\mathrm{vec}}\Delta})^2\hat{\mathcal{D}}^{-1}\mathcal{Q}^{\mathrm{vec}} \\ &\quad + 2\,\mathrm{e}^{-A^{\mathrm{vec}}\Delta(h-1)}(A^{\mathrm{vec}})^{-2}\big(\mathbb{I}_{d^2} - \mathrm{e}^{-A^{\mathrm{vec}}\Delta}\big)^2\mathcal{D}^{-1}\mathcal{Q}^{\mathrm{vec}}, \qquad (8.20)\end{aligned}$$

where $n \in \mathbb{N}$ and $\Delta > 0$. Note that the term in the second line of (8.20) is simply two times the auto-covariance of the multivariate BNS model (compare with (8.10)). Thus, the interesting term is the first one in (8.20). Indeed, assume for simplicity that A^{vec} is symmetric. Then note that for every $\Delta(h-1) > 0$ we obtain the following expansion of the non-linear term in the first line of (8.20):

$$(A^{\mathrm{vec}})^{-2}(\mathcal{G}\,\mathrm{e}^{\Delta(h-1)\hat{\mathcal{D}}} - \mathbf{I})\hat{\mathcal{D}}^{-1} = \Delta(h-1) + \frac{1}{2}\hat{\mathcal{D}}(\Delta(h-1))^2 + \mathcal{O}\big((\Delta(h-1))^3\big).$$

From this, we observe the following small-time asymptotic of the auto-covariance function $\mathrm{acov}_{Y_n^\Delta (Y_n^\Delta)^\intercal}(h)$:

$$\mathrm{acov}_{Y_n^\Delta (Y_n^\Delta)^\intercal}(h) \approx \Delta(h-1)\,\mathrm{e}^{-A^{\mathrm{vec}}\Delta(h-1)}\mathcal{Q}^{\mathrm{vec}} + \mathcal{O}\big((\Delta(h-1))^2\,\mathrm{e}^{-A^{\mathrm{vec}}\Delta(h-1)}\mathcal{Q}^{\mathrm{vec}}\big),$$

for small $\Delta(h-1)$. In comparison with the auto-covariance function of the squared (logarithmic)-returns in the multivariate BNS model, the positive semi-definite well-balanced OU based model allows for a slower decay and even for non-monotone configurations in the short-time lags indicated by the term $\Delta(h–1)$ in front of $\mathrm{e}^{-A^{\mathrm{vec}}\Delta(h-1)}$, which is missing in the small-time asymptotic in the BNS model. Note that the attainable auto-covariance functions are also beyond the ones obtained from superposition of positive semi-definite OU processes, see [11]. Hence, this example demonstrate that the class of higher-order MCARMA based stochastic covariance models is indeed a flexible model class providing multiple modeling options to capture short-memory effects in observed financial and non-financial data. This class is also tractable, as their characteristic functions is known in a closed form, see Proposition 7.4.

Appendix

A.1 Integration with respect to a vector valued measure

In this section we summarize some results on vector-valued measures and integration with respect to such. The theory goes back to the work of Bartle, Dunford, and Schwartz (see, e.g., [14]) and Lewis ([107]). A good overview can be found in [120, Chapter 2]. As we work in the Hilbert-space setting (in particular, as Hilbert spaces are reflexive), the theory simplifies considerably.
Throughout this section let $(S, \mathcal{F})$ be a measurable space, let $(H, \langle \cdot, \cdot \rangle_H)$ be a real Hilbert space, and let $\mu \colon \mathcal{F} \to H$ be an H-valued measure.

Definition A.4. We say that $f \colon S \to \mathbb{R}$ is *μ-integrable* if the following two conditions are satisfied:

i) f is $\langle \mu, h \rangle$-integrable for all $h \in H$ (i.e., $f \colon S \to \mathbb{R}$ is measurable and $\int_S |f| \, \mathrm{d}|\langle \mu, h \rangle| < \infty$ for all $h \in H$), and

ii) for all $A \in \mathcal{F}$ there exists an $h_A \in H$ such that for all $h \in H$ we have $\langle h_A, h \rangle_H = \int_A f \, \mathrm{d}\langle \mu, h \rangle$.

In this case we denote h_A by $\int_A f \, \mathrm{d}\mu$. In addition, we define

$$\mathcal{L}^1(S, \mu) := \{ f \colon S \to \mathbb{R} \colon f \text{ is } \mu\text{-integrable} \}. \tag{A.1}$$

Example A.5. If f is a $\mathcal{F}$-simple function, then $f \in \mathcal{L}^1(S, \mu)$.

The following characterisation is useful (see also [107, p.163]):

Lemma A.6. *We have that $f \in \mathcal{L}^1(S,\mu)$ if and only if f is $\langle \mu, h\rangle$-integrable for all $h \in H$.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of simple functions such that $f_n \to f$ μ-a.s. and $|f_n| \leq |f|$ for all $n \in \mathbb{N}$. Let $A \in \mathcal{F}$. Note that the mapping $T\colon H \to \mathbb{R}$, $T(h) = \int_A f \,\mathrm{d}\langle \mu, h\rangle$ is linear and that

$$T(h) = \lim_{n\to\infty} \int_A f_n \,\mathrm{d}\langle \mu, h\rangle = \lim_{n\to\infty} \langle \int_A f_n \ \mathrm{d}\mu, h\rangle_H,$$

for all $h \in H$ by the dominated convergence theorem. It follows from this and the uniform boundedness principle that $\sup_{n\in\mathbb{N}} \| \int_A f_n \ \mathrm{d}\mu\|_H < \infty$, whence $T \in H^*$. The Riesz representation theorem thus ensures that there exists an $h_A \in H$ such that $\langle h_A, h\rangle_H = T(h)$ for all $h \in H$. □

Corollary A.7. *If $f \in \mathcal{L}^1(S,\mu)$ and $g\colon S \to \mathbb{R}$ is measurable and satisfies $|g| \leq f$ μ-a.s., then $g \in \mathcal{L}^1(S,\mu)$. In particular, $\mathcal{L}^1(S,\mu)$ contains all bounded measurable $\mathbb{R}$-valued functions on S.*

By [107, Corollary 1.4] we have, for any $(E_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ converging to $E \in \mathcal{F}$, that

$$\lim_{n\to\infty} \mu(E_n) = \mu(E). \tag{A.2}$$

Moreover, the dominated convergence theorem remains valid for H-valued measures:

Theorem A.8 (Theorem 2.1.7 in [120])**.** *Let $g \in L^1(S,\mu)$, let $f\colon S \to \mathbb{R}$ be μ-measurable and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of μ-measurable functions on S satisfying $|f_n(s)| \leq g(s)$ for all $s \in S$, $n \in \mathbb{N}$, and $\lim_{n\to\infty} f_n(s) = f(s)$ for all $s \in S$. Then $f, f_n \in L^1(S,\mu)$, $n \in \mathbb{N}$, and*

$$\lim_{n\to\infty} \left\| \int_S f_n \,\mathrm{d}\mu - \int_S f \,\mathrm{d}\mu \right\|_H = 0\,. \tag{A.3}$$

Finally, let $K \subset H$ be a self-dual cone and assume that $\mu\colon \mathcal{F} \to K$ is a K-valued measure. In this case we have $0 \leq_K \mu(E) \leq_K \mu(F)$ for all $E, F \in \mathcal{F}$ satisfying $E \subseteq F$, and thus also (by monotonicity of K)

$$\|\mu(E)\|_H \leq \|\mu(F)\|_H. \tag{A.4}$$

Moreover, as K is self-dual, $\langle \mu, h\rangle$ is a positive measure for all $h \in K$, whence (again by self-duality) we have

$$f \in L^1(S,\mu), f \geq 0 \Rightarrow \int_S f \ \mathrm{d}\mu \in K. \tag{A.5}$$

In particular, if $f \in L^1(S,\mu)$ is positive, and $E \in \mathcal{F}$, then

$$\int_E f \, \mathrm{d}\mu \leq_K \operatorname*{ess\,sup}_{s\in E} f(s)\mu(E). \tag{A.6}$$

This combined with the monotonicity of K implies that for every every $f \in L^1(S,\mu)$ and every $E \in \mathcal{F}$ we have (by considering f^+ and f^- separately) that

$$\left\| \int_E f \, \mathrm{d}\mu \right\|_H \leq \operatorname*{ess\,sup}_{s\in E} |f(s)| \|\mu(E)\|_H. \tag{A.7}$$

A.2 A comparison theorem

A more general version of the following comparison theorem can be found, e.g., as [50, Theorem 5.4].

Theorem A.9. *Let $(H, (\cdot,\cdot))$ be a Hilbert space, $K \subset H$ a cone, let $T > 0$, and let $F\colon [0,T] \times H \to H$. Assume that $F(t,\cdot)$ is quasi-monotone with respect to K for all $t \in [0,T]$, and that there exists a constant $L \in [0,\infty)$ such that*

$$\|F(t,x) - F(t,y)\|_H \leq L\|x-y\|_H, \quad t \in [0,T], x,y \in H. \tag{A.8}$$

Let $f,g \in C^1([0,T],H)$ satisfy $f(0) \leq_K g(0)$ and moreover assume that for all $t \in [0,T]$ we have $f'(t) - F(t,f(t)) \leq_K g'(t) - F(t,g(t))$. Then $f(t) \leq_K g(t) \in K$ for all $t \in [0,T]$.

SUMMARY
Stochastic covariance models in Hilbert spaces with jumps

In this thesis we provide the mathematical foundations for two novel classes of operator-valued stochastic processes with jumps that can be used as models for the instantaneous covariance process in stochastic covariance models in finite- and infinite-dimensional Hilbert spaces. The natural state-space for such processes is the cone of positive self-adjoint Hilbert-Schmidt operators, which is the natural infinite-dimensional version of the cone of positive semi-definite and symmetric matrices. The first class that we study, is the class of affine processes on (infinite-dimensional) positive Hilbert-Schmidt operators. The second is the class of positive semi-definite matrix-valued MCARMA processes.

Operator-valued affine processes In the first part of the thesis we introduce and study the class of affine processes on positive Hilbert-Schmidt operators, which can be viewed as the natural infinite-dimensional version of the class of affine processes on positive semi-definite matrices in [42]. In particular, we prove the existence of a broad class of positive operator-valued affine pure-jump processes admitting jumps of possibly infinite-variation governed by a jump-intensity measure that depends affine linearly on the state of the process. For proving the existence of this class, we use the theory of generalized Feller processes, which requires a thorough analysis of the associated generalized Riccati equations and the regularity of their solutions.
We propose to model the instantaneous covariance process in infinite-dimensional stochastic covariance models by an affine process on positive Hilbert-Schmidt operators and we prove that the so defined stochastic covariance model inherits the tractable affine structure of its instantaneous covariance process.

In this context, we present several examples that extend the operator-valued Barndorff–Nielsen-Shepard (BNS) type models towards a class of models admitting state-dependent jumps.

In a more specific setting, we introduce a geometric affine stochastic covariance model with jumps for forward curve dynamics in, e.g., commodity markets formulated in the Heath-Jarrow-Morton-Musiela framework and demonstrate that the virtues of affine stochastic covariance models for option-pricing via Fourier-inversion methods are maintained in this infinite-dimensional setting. In particular, we derive quasi-explicit formulas for plain vanilla call options written on forwards in terms of the solutions of the generalized Riccati equations.

In addition, we study the long-time behavior of affine processes on positive Hilbert-Schmidt operators and derive explicit rates for the convergence of the transition kernels of a subcritical affine process to its unique invariant measure in Wasserstein distance of order $p \in [1, 2]$. Moreover, we analyze the long-time behavior of the forward volatility smile in the aforementioned geometric affine stochastic covariance model and establish an intimate connection to the implied volatility of plain vanilla options on forwards modeled in the stationary covariance regime.

In addition to that, we study finite-rank approximation of affine stochastic covariance models and establish weakly convergent finite-rank approximations of affine processes on positive Hilbert-Schmidt operators and their generalized Riccati equations. This approach provides an alternative existence proof for affine processes on positive Hilbert-Schmidt operators by exploiting the connection to their matrix-valued versions and an associated Galerkin approximation of the generalized Riccati equations.

Matrix-valued MCARMA processes The second class we introduce and study in this thesis, is the class of matrix-valued MCARMA processes. MCARMA processes can be viewed as solutions to a certain class of higher-order multivariate stochastic differential equations (SDEs) and are known for their flexible covariance structure. We define the class of matrix-valued MCARMA processes as the stationary solutions of certain continuous-time linear state-space models on matrices associated to the higher-order SDEs. Moreover, we prove the equivalence of the matrix-valued versions to the classical MCARMA class under vectorization. Once we established this class, we are concerned with finding sufficient and necessary parameter conditions ensuring that an MCARMA process is cone-valued, e.g. in case of the cones $\mathbb{R}_d^+$ and $\mathbb{S}_d^+$.

Moreover, we study the potential of positive semi-definite MCARMA processes to model the instantaneous covariance process in multivariate stochastic covariance models. This extends the popular multivariate BNS model (a first order MCARMA model) towards more nuanced models based on higher-order MCARMA processes. The appeal of models based on higher-order positive semi-definite MCARMA processes lays in their flexible second order moment structure, which has the potential to model a variety of short-memory features observed in realized (cross)-covariance processes. We justify this hypothesis by an exemplary analysis of positive well-balanced Ornstein-Uhlenbeck based models.

Samenvatting
Stochastische covariantiemodellen in Hilbertruimte met sprongen

In dit proefschrift analyseren we twee nieuwe modellen waarmee de ogenblikkelijke covariantieproces in stochastische covariantiemodellen kan worden beschreven. De natuurlijke toestandsruimte voor zulke modellen is de conus van positief zelf-geadjungeerde operatoren op een (eindig- of oneindigdimensionale) Hilbertruimte. Het ene model betreft oneindigdimensionale stochastische processen die in de klasse van *affiene processen* vallen. Het andere model betreft eindigdimensionale (matrixwaardige) stochastische processen die in de klasse van *MCARMA processen* vallen.

Operatorwaardige affiene processen *Affiene processen* zijn stochastische processen waarvan de charateristieke functie wordt bepaald door de oplossing van een gewone differentiaalvergelijking. Dientengevolge is het relatief eenvoudig om de verwachtingswaarde van (een functie van) een affien proces te bepalen, hetgeen de populariteit van deze processen verklaart. In dit proefschrift worden affiene processen met waardes in de conus van positieve, zelf-geadjungeerde operatoren op een Hilbertruimte geanalyseerd: elk van de onderstaande paragrafen vat één van de hoofdstukken van dit proefschrift samen.

In 2011 hebben onderzoekers een volledige karakterisering van de affiene processen met waardes in de conus van positief semi-definiete matrices gegeven [42]. Een volledige karakterisering van alle affiene processen met waardes in de conus van positieve, zelf-geadjungeerde *operatoren* is nog niet bekend.

Wel hebben wij een 'toelaatbare parameterruimte' bepaald waarbinnen zulke operatorwaardige affiene processen bestaan. Deze toelaatbare parameterruimte staat onder andere toestandsafhankelijke sprongen toe – om specifieker te zijn: de sprongintensiteitsmaat heeft mogelijk oneindige variatie en mag affien lineair afhangen van de toestand van het proces. Voor het bewijzen van het bestaan van deze processen maken wij gebruik van de theorie over gegeneraliseerde Feller processen, die een grondige analyse van de geassocieerde gegeneraliseerde Riccati vergelijkingen en de regelmatigheid van de oplossingen vergt.

De operatorwaardige affiene processen gebruiken wij vervolgens om de ogenblikkelijke covariantie te modellen in een oneindigdimensionaal stochastisch covariantiemodel. Vervolgens tonen we aan dat het gehele model de affiene structuur erft van zijn affiene ogenblikkelijke covariantieproces. Wij beschouwen enkele concrete voorbeelden, in het bijzonder tonen we aan dat ons model een generalisatie is van het in de literatuur bekende operatorwaardige Barndorff-Nielsen-Shepard (BNS) model.

Een andere concrete toepassing betreft het modelleren van de dynamica van termijncurves in, bijvoorbeeld, grondstofmarkten. Via het zogenaamde Heath-Jarrow-Morton-Musiela raamwerk kan de dynamica van termijncurves worden geïnterpreteerd als een oneindigdimensionaal covariantiemodel. We tonen aan dat onze affiene stochastische covariantie modellen toelaten om efficiënt optiekoersen te bepalen via Fourier-inversie methoden. In het bijzonder hebben wij *plain vanilla call options on forwards* expliciet opgeschreven in termen van de oplossing van de gegeneraliseerde Riccati vergelijkingen.

We bestuderen tevens het asymptotische gedrag van de door ons ingevoerde operatorwaardige affiene processen. In het bijzonder geven we voorwaarden voor en snelheid van convergentie van de verdeling van een subkritisch affien proces naar zijn unieke invariante maat in de Wasserstein-distance van orde $p \in [1, 2]$ af. Bovendien vonden wij een interessant verband tussen het asymptotische gedrag van de volatiliteits-*smile* voor forwards gemodelleerd in het geometrische affiene stochastische covariantiemodel en de impliciete volatiliteit van plain vanilla opties op forwards gemodelleerd onder het stationaire covariantieregime.
Omdat de voornaamste motivatie voor het bestuderen van affiene processen ligt in de *berekenbaarheid* van afgeleide verwachtingswaardes, is het ook noodzakelijk om aan te tonen dat onze oneindigdimensionale modellen op een zinnige manier door eindigdimensionale processen kunnen worden benaderd. Concreet hebben we laten zien dat bepaalde eindigdimensionale benaderingen (probabilistisch) zwak convergeren naar het oneindigdimensional model, waarbij we ook een convergentiesnelheid konden bepalen. Dit leverde tevens een alternatief (en krachtiger) bewijs voor het bestaan van dergelijke oneindigdimensionale affiene processen.

Matrixwaardige MCARMA modellen Het tweede stochastische covariantiemodel dat in dit proefschrift wordt geïntroduceerd is gebaseerd op matrixwaardige MCARMA processen. Grofweg beschrijven MCARMA processen oplossingen voor multidimensionale hogere-orde lineaire stochastische differentiaalvergelijkingen (SDV); ze zijn geliefd vanwege hun flexibele covariantie structuur. Wij definiëren de klasse van matrixwaardige MCARMA processen in termen van stationaire oplossingen van bepaalde continue-tijd, lineaire toestandsruimte modellen op matrices geassocieerd met de hogere-orde SDV. Bovendien bewijzen wij dat onze matrixwaardige MCARMA processen via vectorisatie equivalent zijn aan de klassieke MCARMA modellen.

Opdat we onze matrixwaardige MCARMA modellen kunnen gebruiken voor het opstellen van een stochastisch covariantiemodel, bepalen wij noodzakelijk en voldoende voorwaarden op de parameters opdat het MCARMA process waardes aanneemt in bijvoorbeeld de positieve orthant $\mathbb{R}_d^+$ of de positieve semi-definiete matrices $\mathbb{S}_d^+$. We laten zien dat een stochastisch covariantiemodel dat gebruik maakt van onze matrixwaardige MCARMA processen een generalisatie oplevert van het populaire multidimensionale BNS model. Heuristisch gezien staat ons model het toe om een grote(re) verscheidenheid van *short-memory* kenmerken te modelleren. Zulke kenmerken worden waargenomen in gerealiseerde (cross)-covariantie processen. Deze heuristiek onderbouwen wij aan de hand van een voorbeeld met positieve semi-definiete evenwichtige Ornstein-Uhlenbeck gebaseerde modellen.

Author contributions

Chapter 2 is based on the article [38], which is a joint work with Asma Khedher and Sonja Cox, whose idea it was to study affine processes on positive Hilbert-Schmidt operators. Sven Karbach contributed with: the idea to use the theory of generalized Feller semigroups to prove existence, the implementation of the proof (in particular the approximation argument) and establishing uniform growth bounds through an analysis of the regularity of the solutions to the generalized Riccati equations. Khedher and Cox helped making the proofs rigorous and streamlining the arguments.

Chapter 3 is based on the article [39] which is a joint work with Asma Khedher and Sonja Cox, whose idea it was to use affine processes on positive Hilbert-Schmidt operators as the instantaneous covariance process of infinite-dimensional stochastic covariance models. Sven Karbach contributed with proving well-posedness of the model, showing the semimartingale property under the càdlàg paths assumption and using this together with bounded approximation of the operator $\mathcal{A}$ to establish the affine transform formula. Cox and Khedher helped making the proofs rigorous. All authors contributed to the examples equally.

Chapter 4 is based on the theoretical part of a working paper that Sven Karbach writes together with Jian He, Sonja Cox and Asma Khedher. The computations of the real- and complex exponential moments and the extension of the affine transform formula to complex inputs is the contribution of Karbach. The application to the pricing of options on forwards was developed by Karbach together with Khedher and Cox. J. He will contribute numerics in the forthcoming work.

Chapter 5 is based on the preprint [67], that Sven Karbach wrote together with Martin Friesen. Friesen and Karbach contributed equally to establishing the convergence rates in Wasserstein distances. Friesen had the idea to follow the approach in [64] to study the long-time behavior of affine processes on positive Hilbert-Schmidt operators. Karbach developed the proof based on the generalized Feller property of the affine class to circumvent proving the tightness of the transition kernels directly. Friesen helped streamlining the proofs. Karbach had the idea to consider forward-start options written on forwards and proved the long-time behavior of the forward-implied volatility smile in combination with the stationary covariance regime.

Chapter 7 & 8 are based on the preprint [17], which is a joint work of Fred Espen Benth and Sven Karbach. Benth had the initial idea to study positive multivariate CARMA processes. Karbach contributed with establishing the class of matrix-valued MCARMA, proving its equivalence with classical MCARMA under vectorization and proving the existence of unique stationary solutions to associated continuous-time linear state space models. Moreover, Karbach proved the sufficient and necessary positivity conditions, introduced positive semi-definite well-balanced OU processes and computed their second order structure. Benth helped with proofreading, streamlining arguments and giving valuable guidance throughout the preparation of the article.

Acknowledgements

First and foremost, I would like to thank my supervisors Sonja Cox and Asma Khedher for their constant support throughout my candidature. I truly admire not only their mathematical expertise, but also their open-minded, warm and humorous manner that made me feel welcome in Amsterdam from the very first day. Asma and Sonja, I am very grateful for all the advices you gave, discussions we had and laughter we shared. You made my PhD-years truly joyful!

Furthermore, I would like to thank Peter Spreij for being my promoter. Peter, your presence at the institute will surely be missed when you become emeritus next year! I wish you all the best for that.

I would also like to thank Fred Benth, my academic step-grandfather, and Dennis, my academic step-cousin of second degree (approximately) for welcoming me in Oslo, where I had the opportunity to stay for over a month and pay ten euros per beer (with key policy rate still at zero percentage points). In this regard, I also want to thank Martin for inviting me to Dublin and showing me around!

The last four years would not have been so much fun without the friends I made along the journey. Special thanks to Brendan, Rens and Ben!

Lastly, I want to say thank you to my family: My parents, who first of all made me, and secondly supported me during the whole period; Filiz for showing resilience to an increase of the "distance"-parameter in the relationship function; Lars for being a good friend and hyping every little result I found.

Bibliography

[1] Alfonsi, A., Kebaier, A., and Rey, C. Maximum likelihood estimation for Wishart processes. *Stochastic Processes Appl. 126*, 11 (2016), 3243–3282.

[2] Aliprantis, C. D., and Border, K. C. *Infinite dimensional analysis. A hitchhiker's guide.* Berlin: Springer, 2006.

[3] Arendt, W. *Generators of positive semigroups and resolvent positive operators.* PhD thesis, Mathematisches Institut der Universität Tübingen, 1984.

[4] Aström, K. J. *Introduction to stochastic control theory*, Reprint of the 1970 original edition ed. Mineola, NY: Dover Publications, 2006.

[5] Ball, K. Completely monotonic rational functions and Hall's marriage theorem. *J. Comb. Theory, Ser. B 61*, 1 (1994), 118–124.

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

[7] Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. E. D. *Ambit stochastics*, vol. 88 of *Probab. Theory Stoch. Model.* Cham: Springer, 2018.

[8] Barndorff-Nielsen, O. E., and Pérez-Abreu, V. Matrix subordinators and related upsilon transformations. *Theory Probab. Appl. 52*, 1 (2008), 1–23.

[9] Barndorff-Nielsen, O. E., and Shephard, N. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc., Ser. B, Stat. Methodol. 63*, 2 (2001), 167–241.

[10] Barndorff-Nielsen, O. E., and Stelzer, R. Positive-definite matrix processes of finite variation. *Probab. Math. Stat. 27*, 1 (2007), 3–43.

[11] Barndorff-Nielsen, O. E., and Stelzer, R. Multivariate supOU processes. *Ann. Appl. Probab. 21*, 1 (2011), 140–182.

[12] Barndorff-Nielsen, O. E., and Stelzer, R. The multivariate supOU stochastic volatility model. *Math. Finance 23*, 2 (2013), 275–296.

[13] Barski, M., and Zabczyk, J. Heath-Jarrow-Morton-Musiela equation with Lévy perturbation. *J. Differ. Equations 253*, 9 (2012), 2657–2697.

[14] Bartle, R. G., Dunford, N., and Schwartz, J. Weak compactness and vector measures. *Can. J. Math. 7* (1955), 289–305.

[15] Benth, F. E., and Benth, J. *Modeling and Pricing in Financial Markets for Weather Derivatives.* World Scientific, 2012.

[16] Benth, F. E., Christensen, T. S., and Rohde, V. Multivariate continuous-time modeling of wind indexes and hedging of wind risk. *Quant. Finance 21*, 1 (2021), 165–183.

[17] Benth, F. E., and Karbach, S. Positive multivariate continuous-time autoregressive moving-average processes, 2022. Available at https://arxiv.org/abs/2206.08782.

[18] Benth, F. E., and Krühner, P. Representation of infinite-dimensional forward price models in commodity markets. *Commun. Math. Stat. 2*, 1 (2014), 47–106.

[19] Benth, F. E., and Krühner, P. Derivatives pricing in energy markets: an infinite-dimensional approach. *SIAM J. Financ. Math. 6* (2015), 825–869.

[20] Benth, F. E., and Rohde, V. On non-negative modeling with CARMA processes. *J. Math. Anal. Appl. 476*, 1 (2019), 196–214.

[21] Benth, F. E., Rüdiger, B., and Süss, A. Ornstein-Uhlenbeck processes in Hilbert space with non-Gaussian stochastic volatility. *Stochastic Processes Appl. 128*, 2 (2018), 461–486.

[22] Benth, F. E., and Sgarra, C. A Barndorff-Nielsen and Shephard model with leverage in Hilbert space for commodity forward markets. *Available at SSRN 3835053* (2021).

[23] Benth, F. E., and Simonsen, I. C. The Heston stochastic volatility model in Hilbert space. *Stochastic Anal. Appl. 36*, 4 (2018), 733–750.

[24] Benth, F. E., and Šaltytė Benth, J. *Modeling and pricing in financial markets for weather derivatives*, vol. 17. Hackensack, NJ: World Scientific, 2012.

[25] Benth, F. E., and Šaltytė Benth, J. Dynamic pricing of wind futures. *Energy Economics 31*, 1 (2009), 16–24.

[26] Black, F. The pricing of commodity contracts. *J. financ. econ. 3*, 1-2 (1976), 167–179.

[27] Brockwell, P., and Lindner, A. Integration of CARMA processes and spot volatility modelling. *J. Time Ser. Anal. 34*, 2 (2013), 156–167.

[28] Brockwell, P. J. Recent results in the theory and applications of CARMA processes. *Ann. Inst. Stat. Math. 66*, 4 (2014), 647–685.

[29] Brockwell, P. J., Davis, R. A., and Yang, Y. Estimation for non-negative Lévy-driven CARMA processes. *J. Bus. Econ. Stat. 29*, 2 (2011), 250–259.

[30] Brockwell, P. J., and Schlemm, E. Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations. *J. Multivariate Anal. 115* (2013), 217–251.

[31] Bru, M.-F. Wishart processes. *J. Theor. Probab. 4*, 4 (1991), 725–751.

[32] Carmona, R. A., and Tehranchi, M. R. *Interest rate models: an infinite dimensional stochastic analysis perspective.* Berlin: Springer, 2006.

[33] Carr, P., and Madan, D. B. Option valuation using the fast Fourier transform. *J. Comput. Finance 2* (1999), 61–73.

[34] Caversaccio, P. M. *Pricing VIX Options with Wishart Matrix Affine Jump Diffusions while Preserving Consistency with SPX Options.* PhD thesis, University of Zurich, Switzerland, 2014.

[35] Chen, R., Xiao, H., and Yang, D. Autoregressive models for matrix-valued time series. *Journal of Econometrics 222*, 1, Part B (2021), 539–560. Annals Issue:Financial Econometrics in the Age of the Digital Economy.

[36] Chojnowska-Michalik, A. On processes of Ornstein-Uhlenbeck type in Hilbert space. *Stochastics 21* (1987), 251–286.

[37] Cont, R. Empirical properties of asset returns: stylized facts and statistical issues. *Quant. Finance 1*, 2 (2001), 223–236.

[38] Cox, S., Karbach, S., and Khedher, A. Affine pure-jump processes on positive Hilbert–Schmidt operators. *Stochastic Processes and their Applications 151* (2022), 191–229.

[39] Cox, S., Karbach, S., and Khedher, A. An infinite-dimensional affine stochastic volatility model. *Mathematical Finance 32*, 3 (2022), 878–906.

[40] Cox, S. G., and Hausenblas, E. A perturbation result for semi-linear stochastic differential equations in UMD Banach spaces. *J. Evol. Equ. 13*, 4 (2013), 795–827.

[41] Crosby, J. Pricing a class of exotic commodity options in a multi-factor jump-diffusion model. *Quant. Finance 8*, 5 (2008), 471–483.

[42] Cuchiero, C., Filipović, D., Mayerhofer, E., and Teichmann, J. Affine processes on positive semidefinite matrices. *Ann. Appl. Probab. 21*, 2 (2011), 397–463.

[43] Cuchiero, C., Keller-Ressel, M., Mayerhofer, E., and Teichmann, J. Affine processes on symmetric cones. *J. Theor. Probab. 29*, 2 (2016), 359–422.

[44] Cuchiero, C., and Teichmann, J. Path properties and regularity of affine processes on general state spaces. In *Séminaire de probabilités XLV*. Cham: Springer, 2013, pp. 201–244.

[45] Cuchiero, C., and Teichmann, J. Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. *J. Evol. Equ. 20*, 4 (2020), 1301–1348.

[46] Curtain, R. F., and Zwart, H. *An introduction to infinite-dimensional linear systems theory*, vol. 21. New York, NY: Springer-Verlag, 1995.

[47] Da Prato, G., Flandoli, F., Priola, E., and Röckner, M. Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift. *The Annals of Probability 41*, 5 (2013), 3306–3344.

[48] Da Prato, G., and Zabczyk, J. *Stochastic equations in infinite dimensions*, vol. 44. Cambridge etc.: Cambridge University Press, 1992.

[49] Dawson, D. A., and Li, Z. Skew convolution semigroups and affine Markov processes. *Ann. Probab. 34*, 3 (2006), 1103–1142.

[50] Deimling, K. *Ordinary differential equations in Banach spaces*, vol. 596. Springer, Cham, 1977.

[51] Doersek, P., and Teichmann, J. A semigroup point of view on splitting schemes for stochastic (partial) differential equations, 2010. arXiv 1011.2651.

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

[53] Eberlein, E., Glau, K., and Papapantoleon, A. Analysis of Fourier transform valuation formulas and applications. *Appl. Math. Finance 17*, 3-4 (2010), 211–240.

[54] Eisenmann, M., Emmrich, E., and Mehrmann, V. Convergence of the backward Euler scheme for the operator-valued Riccati differential equation with semi-definite data. *Evol. Equ. Control Theory 8*, 2 (2019), 315–342.

[55] Engel, K.-J., and Nagel, R. *One-parameter semigroups for linear evolution equations*, vol. 194. Berlin: Springer, 2000.

[56] Ethier, S. N., and Kurtz, T. G. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[57] Eydeland, A., and Wolyniec, K. *Energy and power risk management: New developments in modeling, pricing, and hedging*, vol. 97. John Wiley & Sons, 2002.

[58] Farina, L., and Rinaldi, S. *Positive linear systems. Theory and applications.* New York, NY: Wiley, 2000.

[59] Fasen, V. Dependence estimation for high-frequency sampled multivariate CARMA models. *Scand. J. Stat. 43*, 1 (2016), 292–320.

[60] Filipović, D. *Consistency problems for Heath-Jarrow-Morton interest rate models*, vol. 1760. Berlin: Springer, 2001.

[61] Filipovic, D. *Term-Structure Models. A Graduate Course.* Springer, 2009.

[62] Filipović, D., and Mayerhofer, E. *Affine diffusion processes: theory and applications.* Berlin: de Gruyter, 2009.

[63] FRIESEN, M., AND JIN, P. On the anisotropic stable JCIR process. *ALEA, Lat. Am. J. Probab. Math. Stat. 17*, 2 (2020), 643–674.

[64] FRIESEN, M., JIN, P., KREMER, J., AND RÜDIGER, B. Ergodicity of affine processes on the cone of symmetric positive semidefinite matrices. *Adv. Appl. Probab. 52*, 3 (2020), 825–854.

[65] FRIESEN, M., JIN, P., KREMER, J., AND RÜDIGER, B. Regularity of transition densities and ergodicity for affine jump-diffusion processes. *Math. Nach. (forthcoming)*, to appear (2022).

[66] FRIESEN, M., JIN, P., AND RÜDIGER, B. Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes. *Ann. Appl. Probab. 30*, 5 (2020), 2165–2195.

[67] FRIESEN, M., AND KARBACH, S. Stationary Covariance Regime for Affine Stochastic Covariance Models in Hilbert Spaces, 2022. Available at https://arxiv.org/abs/2203.14750.

[68] GAWARECKI, L., AND MANDREKAR, V. *Stochastic differential equations in infinite dimensions with applications to stochastic partial differential equations.* Berlin: Springer, 2011.

[69] GERMANI, A., JETTO, L., AND PICCIONI, M. Galerkin approximation for optimal linear filtering of infinite-dimensional linar systems. *SIAM J. Control Optim. 26*, 6 (1988), 1287–1305.

[70] GLASSERMAN, P., AND KIM, K.-K. Moment explosions and stationary distributions in affine diffusion models. *Math. Finance 20*, 1 (2010), 1–33.

[71] GOHBERG, I., LANCASTER, P., AND RODMAN, L. *Matrix polynomials*, reprint of the 1982 original ed., vol. 58 of *Classics Appl. Math.* Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2009.

[72] GOLDSTEIN, J. A. *Semigroups of linear operators and applications.* Oxford University Press, Oxford, 1985.

[73] GOURIEROUX, C., AND SUFANA, R. Derivative pricing with Wishart multivariate stochastic volatility. *J. Bus. Econ. Stat. 28*, 3 (2010), 438–451.

[74] GRAFENDORFER, G. *Infinite-Dimensional Affine Processes.* PhD thesis, ETH Zürich, 2016.

[75] GÖTHEL, R. Faedo-Galerkin approximations in equations of evolution. *Math. Methods Appl. Sci. 6* (1984), 41–54.

[76] Herzog, G., and Lemmert, R. On quasipositive elements in ordered Banach algebras. *Stud. Math. 129*, 1 (1998), 59–65.

[77] Heston, S. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies 6* (1993), 327–343.

[78] Horn, R. A., and Johnson, C. R. *Topics in matrix analysis.* Cambridge etc.: Cambridge University Press, 1991.

[79] Hytönen, T., van Neerven, J., Veraar, M., and Weis, L. *Analysis in Banach spaces. Volume I. Martingales and Littlewood-Paley theory*, vol. 63. Cham: Springer, 2016.

[80] Jacod, J., and Shiryaev, A. N. *Limit theorems for stochastic processes*, vol. 288. Berlin: Springer, 2003.

[81] Jacquier, A., and Roome, P. The small-maturity Heston forward smile. *SIAM J. Financ. Math. 4* (2013), 831–856.

[82] Jacquier, A., and Roome, P. Asymptotics of forward implied volatility. *SIAM J. Financ. Math. 6* (2015), 307–351.

[83] Jakubowski, A. On the Skorokhod topology. *Annales de l'I.H.P. Probabilités et statistiques 22*, 3 (1986), 263–285.

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

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

[86] Jin, P., Kremer, J., and Rüdiger, B. Existence of limiting distribution for affine processes. *J. Math. Anal. Appl. 486*, 2 (2020), 30. Id/No 123912.

[87] Joffe, A., and Métivier, M. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. Appl. Probab. 18* (1986), 20–65.

[88] Jones, R. H., and Boadi-Boateng, F. Unequally spaced longitudinal data with ar(1) serial correlation. *Biometrics 47*, 1 (1991), 161–175.

[89] Kailath, T. *Linear systems.* Prentice Hall Inform. System Sci. Ser. Prentice Hall, Englewood Cliffs, NJ, 1980.

[90] Kallsen, J., and Muhle-Karbe, J. Exponentially affine martingales, affine measure changes and exponential moments of affine processes. *Stochastic Processes Appl. 120*, 2 (2010), 163–181.

[91] Kamthan, P. K., and Gupta, M. Theory of bases and cones. Research Notes in Mathematics, 117. Boston - London - Melbourne: Pitman Advanced Publishing Program. 256 p.(1985)., 1985.

[92] Karlin, S. Positive operators. *J. Math. Mech. 8* (1959), 907–937.

[93] Keller-Ressel, M. *Affine processes: theory and applications in finance.* PhD thesis, TU Wien, 2008.

[94] Keller-Ressel, M. Moment explosions and long-term behavior of affine stochastic volatility models. *Math. Finance 21*, 1 (2011), 73–98.

[95] Keller-Ressel, M., and Mayerhofer, E. Exponential moments of affine processes. *Ann. Appl. Probab. 25*, 2 (2015), 714–752.

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

[97] Keller-Ressel, M., Schachermayer, W., and Teichmann, J. Regularity of affine processes on general state spaces. *Electron. J. Probab. 18* (2013), 17. Id/No 43.

[98] Kevei, P. Asymptotic moving average representation of high-frequency sampled multivariate CARMA processes. *Ann. Inst. Stat. Math. 70*, 2 (2018), 467–487.

[99] Konno, N., and Shiga, T. Stochastic partial differential equations for some measure-valued diffusions. *Probab. Theory Relat. Fields 79*, 2 (1988), 201–225.

[100] Krasnosel'skii, M. A., Lifshits, E. A., and Sobolev, A. V. *Positive linear systems. - The method of positive operators - Transl. from the Russian by Jürgen Appell.* Berlin: Heldermann-Verlag, 1989.

[101] Kruse, S., and Nögel, U. On the pricing of forward starting options in Heston's model on stochastic volatility. *Finance Stoch. 9*, 2 (2005), 233–250.

[102] Kurtz, T. G., and Protter, P. E. Weak convergence of stochastic integrals and differential equations. II: Infinite dimensional case. In *Probabilistic models for nonlinear partial differential equations.* Berlin: Springer, 1996, pp. 197–285.

[103] Lax, P. D. *Functional analysis.* Pure Appl. Math., Wiley-Intersci. Ser. Texts Monogr. Tracts. Chichester: Wiley, 2002.

[104] Leippold, M., and Trojani, F. Asset pricing with matrix jump diffusions. *American Finance Association Meetings (AFA)* (2008).

[105] Lemaire, V., Montes, T., and Pagès, G. Stationary Heston model: calibration and pricing of exotics using product recursive quantization. *Quantitative Finance 22*, 4 (2022), 611–629.

[106] Lemmert, R., and Volkmann, P. On the positivity of semigroups of operators. *Commentat. Math. Univ. Carol. 39*, 3 (1998), 483–489.

[107] Lewis, D. R. Integration with respect to vector measures. *Pac. J. Math. 33* (1970), 157–165.

[108] Li, C.-K., and Tsing, N.-K. Linear preserver problems: A brief introduction and some special techniques. *Linear Algebra Appl. 162-164* (1992), 217–235.

[109] Li, Z. Ergodicities and exponential ergodicities of Dawson-Watanabe type processes. *Theory Probab. Appl. 66*, 2 (2021), 276–298.

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

[111] Lions, J. L. *Optimal control of systems governed by partial differential equations*, vol. 170. Springer, Cham, 1971.

[112] Marquardt, T., and Stelzer, R. Multivariate CARMA processes. *Stochastic Processes Appl. 117*, 1 (2007), 96–120.

[113] Martin, R. H. *Nonlinear operators and differential equations in Banach spaces.* John Wiley&Sons, 1976.

[114] Mayerhofer, E. Affine processes on positive semidefinite $d \times d$ matrices have jumps of finite variation in dimension $d > 1$. *Stochastic Processes Appl. 122*, 10 (2012), 3445–3459.

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

[116] Merkle, M. J. On weak convergence of measures on Hilbert spaces. *J. Multivariate Anal. 29*, 2 (1989), 252–259.

[117] METIVIER, M. *Semimartingales: A course on stochastic processes*, vol. 2. Berlin - New York: de Gruyter, 1982.

[118] MOSSBERG, M., AND LARSSON, E. Fast and approximative estimation of continuous-time stochastic signals from discrete-time data. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing* (2004), vol. 2, pp. ii–529.

[119] NIELSEN, M., AND ROHDE, V. On nonnegative solutions of sddes with an application to carma processes. *Modern Stochastics: Theory and Applications 8*, 3 (2021), 309–328.

[120] PANCHAPAGESAN, T. V. *The Bartle-Dunford-Schwartz integral. Integration with respect to a sigma-additive vector measure*, vol. 69. Basel: Birkhäuser, 2008.

[121] PAZY, A. *Semigroups of linear operators and applications to partial differential equations*, vol. 44. Springer, Cham, 1983.

[122] PÉREZ-ABREU, V., AND ROCHA-ARTEAGA, A. On the Lévy-Khintchine representation of Lévy processes in cones of Banach spaces. *Publ. Mat. Urug. 11* (2006), 41–55.

[123] PESZAT, S., AND ZABCZYK, J. *Stochastic partial differential equations with Lévy noise. An evolution equation approach.*, vol. 113. Cambridge: Cambridge University Press, 2007.

[124] PETTIS, B. J. On integration in vector spaces. *Trans. Am. Math. Soc. 44* (1938), 277–304.

[125] PIGORSCH, C., AND STELZER, R. A multivariate Ornstein-Uhlenbeck type stochastic volatility model, 2009.

[126] PIGORSCH, C., AND STELZER, R. On the definition, stationary distribution and second order structure of positive semidefinite Ornstein-Uhlenbeck type processes. *Bernoulli 15*, 3 (2009), 754–773.

[127] RODMAN, L. On factorization of operator polynomials and analytic operator functions. *Rocky Mt. J. Math. 16* (1986), 153–162.

[128] ROSEN, I. G. Convergence of Galerkin approximations for operator Riccati equations - a nonlinear evolution equation approach. *J. Math. Anal. Appl. 155*, 1 (1991), 226–248.

[129] ROSENBLUM, M. On the operator equation $BX - XA = Q$. *Duke Math. J. 23* (1956), 263–269.

[130] Rusinek, A. Mean reversion for HJMM forward rate models. *Adv. Appl. Probab. 42*, 2 (2010), 371–391.

[131] Sato, K., and Yamazato, M. Stationary processes of Ornstein-Uhlenbeck type. Probability theory and mathematical statistics, Lect. Notes Math. 1021, 541-551 (1983)., 1983.

[132] Sato, K.-I. *Lévy processes and infinitely divisible distributions*, vol. 68 of *Camb. Stud. Adv. Math.* Cambridge: Cambridge University Press, 1999.

[133] Schlemm, E., and Stelzer, R. Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. *Bernoulli 18*, 1 (2012), 46–63.

[134] Schmidt, T., Tappe, S., and Yu, W. Infinite dimensional affine processes. *Stochastic Processes Appl. 130*, 12 (2020), 7131–7169.

[135] Schneider, H. Positive operators and an inertia theorem. *Numer. Math. 7* (1965), 11–17.

[136] Schnurr, A., and Woerner, J. H. C. Well-balanced Lévy driven Ornstein-Uhlenbeck processes. *Stat. Risk. Model. 28*, 4 (2011), 343–357.

[137] Shannon, D., and Fountas, G. Extending the Heston model to forecast motor vehicle collision rates. *Accident Analysis & Prevention 159* (2021), 106250.

[138] Tappe, S. Compact embeddings for spaces of forward rate curves. *Abstr. Appl. Anal. 2013* (2013), 6. Id/No 709505.

[139] Temam, R. Sur l'équation de Riccati associé a des opérateurs non bornés, en dimension infinie. *J. Funct. Anal. 7* (1971), 85–115.

[140] Todorov, V., and Tauchen, G. Simulation Methods for Lévy-Driven Continuous-Time Autoregressive Moving Average (CARMA) Stochastic Volatility Models. *Journal of Business & Economic Statistics 24*, 4 (2006), 455–469.

[141] Tsai, H., and Chan, K. A note on the non-negativity of continuous-time ARMA and GARCH processes. *Statistics and Computing 19* (2009), 149–153.

[142] Tsai, H., and Chan, K. S. A note on non-negative continuous time processes. *J. R. Stat. Soc., Ser. B, Stat. Methodol. 67*, 4 (2005), 589–597.

[143] Villani, C. *Optimal transport. Old and new*, vol. 338. Berlin: Springer, 2009.

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

[145] Wang, J. On the Exponential Ergodicity of Lévy-Driven Ornstein-Uhlenbeck Processes. *J. Appl. Probab. 49* (2012), 990–1004.

[146] Werner, D. *Funktionalanalysis.* Berlin: Springer, 2000.

[147] Yu, W. *Infinite dimensional affine term structure models under incomplete information.* Freiburg: Univ. Freiburg, Fakultät für Mathematik und Physik (Diss.), 2017.

[148] Zadeh, L. A., and Desoer, C. A. *Linear system theory. The state space approach.* New York, NY: McGraw-Hill, 1963.

Index