Chapter 11
Wishart Processes

The aim of this chapter is to introduce Wishart processes as tractable diffusions, which can be used to better capture dependence structures associated with multidimensional stochastic models. The focus of this chapter is on the tractability aspect. We present illustrative examples, which show that we can move beyond the dependence structures possible on the Euclidean state space. As discussed in Chap. 9, we consider a model to be tractable if we have access to its affine transform. As demonstrated in Bru (1991), Grasselli and Tebaldi (2008), Ahdida and Alfonsi (2010), Benabid et al. (2010), Laplace transforms of the Wishart process are available in closed-form and exponentially affine. The Wishart process is, in fact, an affine process. We present results on affine transforms in Sect. 11.4.

Besides computing Laplace transforms, exact simulation schemes play an important role in finance, as they allow the pricing of e.g. path-dependent options, see also Chap. 6. In Sect. 11.3, we will discuss simulation schemes for the Wishart process, where we present the approaches from Benabid et al. (2010) and Ahdida and Alfonsi (2010). The two approaches are different in nature, as they exploit different properties of Wishart processes. We hence present both approaches, as they illustrate interesting properties of Wishart processes.

Subsequently, we present an extension of the model presented in Sect. 9.7 to the case where positive factors are modeled via a Wishart process. This illustrates the additional degrees of freedom given by employing the Wishart process. We begin this chapter with a section which introduces Wishart processes and present existence results. Subsequently, we study some special cases of the Wishart process in detail to gain further insight. One of the special cases motivates immediately one of the simulation schemes to be presented in Sect. 11.3.

11.1 Definition and Existence Results

Wishart processes were introduced in Bru (1991), as a matrix generalization of squared Bessel processes. In her PhD thesis, Bru applied these processes to problems from biology. As we will show below, Wishart processes are S_d^+ or $\overline{S_d^+}$ valued,

J. Baldeaux, E. Platen, *Functionals of Multidimensional Diffusions with Applications to Finance*, Bocconi & Springer Series 5, DOI 10.1007/978-3-319-00747-2_11,

i.e. they assume values as positive definite or positive semidefinite matrices. This makes them natural candidates for the modeling of covariance matrices, as noted in Gouriéroux and Sufana (2004a). Starting with Gouriéroux and Sufana (2004a, 2004b), there is now a substantial body of literature applying Wishart processes to problems in finance, see Gouriéroux et al. (2007), Da Fonseca et al. (2007, 2008a, 2008b, 2008c), and Buraschi et al. (2008, 2010).

All of the above references study Wishart processes in a pure diffusive setting. Recently, matrix valued processes incorporating jumps have been studied, see e.g. Barndorff-Nielsen and Stelzer (2007), Leippold and Trojani (2008). These processes are all contained in the affine framework introduced in Cuchiero et al. (2011), where we direct the interested reader. Furthermore, we mention the recent paper Cuchiero et al. (2011), which extends the results from Cuchiero et al. (2011) to symmetric cones.

We introduce the Wishart process as in the work of Grasselli and collaborators. For $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$, we introduce the $\overline{\mathcal{S}_d^+}$ valued Wishart process $\boldsymbol{X}^{\boldsymbol{x}} = \boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$, which satisfies

$$d\boldsymbol{X}_t = \left(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t \boldsymbol{b}^\top\right) dt + \left(\sqrt{\boldsymbol{X}_t}\, d\boldsymbol{W}_t \boldsymbol{a} + \boldsymbol{a}^\top d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{X}_t}\right), \qquad (11.1.1)$$

where $\alpha \geq 0$, $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{M}_d$ and $\boldsymbol{X}_0 = \boldsymbol{x} \in \mathcal{M}_d$. An obvious question to ask is whether Eq. (11.1.1) admits a solution, and furthermore if such a solution is unique and strong. For results on weak solutions, we refer the reader to Cuchiero et al. (2011), and for results on strong solutions to Mayerhofer et al. (2011b). We now present a summary of their results, see Corollary 3.2 in Mayerhofer et al. (2011b) and also Theorem 1 in Ahdida and Alfonsi (2010).

Theorem 11.1.1 *Assume that $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$, and $\alpha \geq d-1$, then Eq.* (11.1.1) *admits a unique weak solution. If $\boldsymbol{x} \in \mathcal{S}_d^+$ and $\alpha \geq d+1$, then this solution is strong.*

In this book, we are primarily interested in explaining the tractability of the processes under consideration, where in this chapter, we focus on Wishart processes. In particular, we present for the Wishart process Laplace transforms and an exact simulation scheme. Weak solutions suffice for our purposes and we assume that $\alpha \geq d-1$, so that the weak solution is also unique. As in Ahdida and Alfonsi (2010), we use $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$ to denote a Wishart process and $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a}; t)$ for the value of the process at time point t.

11.2 Some Special Cases

In this section, we discuss some particular special cases of Wishart processes. Recall that we defined a Wishart process $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$ to be

$$d\boldsymbol{X}_t = \left(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t \boldsymbol{b}^\top\right) dt + \sqrt{\boldsymbol{X}_t}\, d\boldsymbol{W}_t\, \boldsymbol{a} + \boldsymbol{a}^\top d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{X}_t},$$

for $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{M}_d$, and $\alpha \geq d - 1$. In Sect. 3.2, we had already introduced Wishart processes. We recover the special case studied in Sect. 3.2 by setting $\boldsymbol{a} = \boldsymbol{I}_d$, $\boldsymbol{b} = \boldsymbol{0}$, and $\alpha = d$, to obtain

$$d\boldsymbol{X}_t = d\boldsymbol{I}_d\,dt + \sqrt{\boldsymbol{X}_t}\,d\boldsymbol{W}_t + d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{X}_t}. \tag{11.2.2}$$

Recall that Eq. (11.2.2) is the analogue of Eq. (3.1.1), which introduces the squared Bessel process as a sum of squared Brownian motions. In Eq. (11.2.2), d is also an integer. Subsequently, in Sect. 3.1 we relaxed the assumption that d is an integer. We now do the same for Wishart processes. However, in Bru (1991), the condition $\alpha \geq d - 1$ was used to establish the existence of a unique weak solution, see Theorem 2 in Bru (1991), i.e. she established the existence of a unique weak solution of a $WIS_d(\boldsymbol{x}, d, \boldsymbol{0}, \boldsymbol{I}_d)$ process.

So far, we introduced Wishart processes as squares of matrices of Brownian motions, i.e. the $WIS_d(\boldsymbol{x}, d, \boldsymbol{0}, \boldsymbol{I}_d)$ case. However, as in Bru (1991), we can also establish a connection with squared matrix-valued Ornstein-Uhlenbeck processes. This is an important observation, and will also motivate our first simulation scheme in Sect. 11.3.

Let $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$ be an $n \times d$ matrix diffusion solution of

$$d\boldsymbol{X}_t = \gamma\,d\boldsymbol{B}_t + \beta \boldsymbol{X}_t\,dt, \tag{11.2.3}$$

where $\boldsymbol{B} = \{\boldsymbol{B}_t,\ t \geq 0\}$ is an $n \times d$ Brownian motion, and $\boldsymbol{x}$ is an $n \times d$ matrix, $\gamma \in \Re$, and $\beta \in \Re^-$. We set $\boldsymbol{S}_t = \boldsymbol{X}_t^\top \boldsymbol{X}_t$, $\boldsymbol{s} = \boldsymbol{x}^\top \boldsymbol{x}$.

Lemma 11.2.1 *Assume that $\boldsymbol{X}_t$ satisfies Eq.* (11.2.3). *Then $\boldsymbol{S}_t = \boldsymbol{X}_t^\top \boldsymbol{X}_t$ satisfies the SDE*

$$\begin{aligned} d\boldsymbol{S}_t &= \gamma\left(\sqrt{\boldsymbol{S}_t}\,d\boldsymbol{B}_t + d\boldsymbol{B}_t^\top \sqrt{\boldsymbol{S}_t}\right) + 2\beta\sqrt{\boldsymbol{S}_t}\,dt + n\gamma^2 \boldsymbol{I}_d\,dt, \\ \boldsymbol{S}_0 &= \boldsymbol{s}. \end{aligned} \tag{11.2.4}$$

Proof The technique of the proof follows Theorem 4.19 in Pfaffel (2008), where the result was shown for the more general case that corresponds to Lemma 11.2.2. We define

$$\boldsymbol{S}_t = \boldsymbol{X}_t^\top \boldsymbol{X}_t, \quad t \geq 0,$$

and

$$\boldsymbol{W}_t = \int_0^t \sqrt{\boldsymbol{S}_u^{-1}}\boldsymbol{X}_u^\top\,d\boldsymbol{B}_u \in \mathcal{M}_d,$$

for all $t \geq 0$. We first show that $\boldsymbol{W} = \{\boldsymbol{W}_t,\ t \geq 0\}$ is a Brownian motion. We compute

$$\begin{aligned} &E\left(\int_0^t \left(\sqrt{\boldsymbol{S}_u^{-1}}\boldsymbol{X}_u^\top\right)^\top \left(\sqrt{\boldsymbol{S}_u^{-1}}\boldsymbol{X}_u^\top\right) du\right) \\ &= E\left(\int_0^t \boldsymbol{X}_u \boldsymbol{S}_u^{-1} \boldsymbol{X}_u^\top\,du\right) \\ &= t\boldsymbol{I}_p < \infty, \quad \text{a.s.,} \end{aligned}$$

establishing that $\boldsymbol{W}$ is a local martingale. Also,

$$dW_{t,ij} = \sum_{m=1}^{d} \left[\sqrt{\boldsymbol{S}_u^{-1}} \boldsymbol{X}_u^{\top} \right]_{i,m} dB_{t,mj}.$$

Hence

$$\begin{aligned} & d[W_{\cdot,ij} W_{\cdot,kl}]_t \\ & = \sum_{m=1}^{d} \left[\sqrt{\boldsymbol{S}_t^{-1}} \boldsymbol{X}_t^{\top} \right]_{i,m} \left[\sqrt{\boldsymbol{S}_t^{-1}} X_t^{\top} \right]_{k,m} \mathbf{1}_{j=l} \, dt \\ & = \left[\sqrt{\boldsymbol{S}_t^{-1}} \boldsymbol{X}_t^{\top} \boldsymbol{X}_t \sqrt{\boldsymbol{S}_t^{-1}} \right]_{i,k} \mathbf{1}_{j=l} \, dt \\ & = \boldsymbol{I}_{ik} \mathbf{1}_{j=l} \, dt \\ & = \mathbf{1}_{i=k} \mathbf{1}_{j=l} \, dt, \end{aligned}$$

where we used that

$$d[W_{\cdot,nj}, W_{\cdot,nl}]_t = dt \iff j = l.$$

By Theorem 10.4.5, $\boldsymbol{W}$ is a Brownian motion. Finally, we compute

$$\begin{aligned} d\boldsymbol{S}_t &= d\big(\boldsymbol{X}_t^{\top} \boldsymbol{X}_t\big) = (d\boldsymbol{X}_t)^{\top} \boldsymbol{X}_t + \boldsymbol{X}_t^{\top} d\boldsymbol{X}_t + d\big[\boldsymbol{X}^{\top}, \boldsymbol{X}\big]_t^M \\ &= \big(\beta \boldsymbol{X}_t^{\top} dt + d\boldsymbol{B}_t^{\top} \gamma\big) \boldsymbol{X}_t + \boldsymbol{X}_t^{\top} (\beta \boldsymbol{X}_t \, dt + d\boldsymbol{B}_t \gamma) + \gamma d\big[\boldsymbol{B}^{\top}, \boldsymbol{B}\big]_t^M \gamma \\ &= \beta \boldsymbol{S}_t \, dt + d\boldsymbol{B}_t^{\top} \gamma \boldsymbol{X}_t + \beta \boldsymbol{S}_t \, dt + \boldsymbol{X}_t^{\top} d\boldsymbol{B}_t \gamma + \gamma^2 n \boldsymbol{I}_d \, dt \\ &= 2\beta \boldsymbol{S}_t \, dt + \gamma \sqrt{\boldsymbol{S}_t} \, d\boldsymbol{W}_t + \gamma \, d\boldsymbol{W}_t^{\top} \sqrt{\boldsymbol{S}_t} + \gamma^2 n \boldsymbol{I}_d \, dt, \end{aligned}$$

which yields that $\boldsymbol{S}_t$ solves Eq. (11.2.4). □

The following time-change formula is reminiscent of Proposition 3.1.5, see Eq. (5.3) in Bru (1991). If $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$ is a solution of (11.2.3), then there exists a Wishart process $\boldsymbol{\Sigma} = \{\boldsymbol{\Sigma}_t,\ t \geq 0\} \in WIS_d(\boldsymbol{s}, \alpha, \mathbf{0}, \boldsymbol{I}_d)$ such that

$$\boldsymbol{S}_t = \boldsymbol{X}_t^{\top} \boldsymbol{X}_t = \exp\{2\beta t\} \boldsymbol{\Sigma}_{\gamma^2 \frac{1-\exp\{-2\beta t\}}{2\beta}}.$$

Using this time-change formula, Bru established that the Wishart process $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$, where $\boldsymbol{b} = \beta \boldsymbol{I}_d$, $\boldsymbol{a} = \gamma \boldsymbol{I}_d$, $\beta, \gamma \in \Re$, $\alpha \geq d - 1$ and $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$, with distinct eigenvalues, admits a unique weak solution, see Theorem 2′ in Bru (1991). Finally, she extended this result replacing γ and β by $d \times d$ matrices $\boldsymbol{b}$ and $\boldsymbol{a}$, where $\boldsymbol{a} \in GL(d)$. We consider the following SDE for $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$,

$$d\boldsymbol{X}_t = d\boldsymbol{B}_t \boldsymbol{a} + \boldsymbol{X}_t \boldsymbol{b} \, dt, \tag{11.2.5}$$

where $\boldsymbol{X}_0 = \boldsymbol{x}$, $\boldsymbol{B} = \{\boldsymbol{B}_t,\ t \geq 0\}$ is a $n \times d$ matrix valued Brownian motion and $\boldsymbol{X}_t$ is an $n \times d$ matrix. We set $\boldsymbol{S}_t = \boldsymbol{X}_t^{\top} \boldsymbol{X}_t$, $\boldsymbol{s} = \boldsymbol{x}^{\top} \boldsymbol{x} \in \mathcal{S}_d^+$, and in the next lemma derive the SDE for $\boldsymbol{S}_t$.

Lemma 11.2.2 *Assuming* $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$ *satisfies Eq.* (11.2.5), *we obtain the following dynamics for* $\boldsymbol{S}_t = \boldsymbol{X}_t^\top \boldsymbol{X}_t$,

$$d\boldsymbol{S}_t = \sqrt{\boldsymbol{S}_t}\, d\boldsymbol{W}_t \sqrt{\boldsymbol{a}^\top \boldsymbol{a}} + \sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\, d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{S}_t} + \big(\boldsymbol{b}^\top \boldsymbol{S}_t + \boldsymbol{S}_t \boldsymbol{b}\big)\, dt + n\boldsymbol{a}^\top \boldsymbol{a}\, dt, \tag{11.2.6}$$

where $\boldsymbol{S}_0 = \boldsymbol{s}$.

Proof The proof is completed in the same fashion as the proof of Lemma 11.2.1, and is given in Pfaffel (2008). We firstly define

$$\boldsymbol{W}_t = \int_0^t \sqrt{\boldsymbol{S}_u^{-1}} \boldsymbol{X}_u^\top\, d\boldsymbol{B}_u \boldsymbol{a} \big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1}\, du \quad \in \mathcal{M}_d.$$

Note that the matrix square-root is positive definite and hence invertible. We compute

$$\begin{aligned}
&E\bigg(\int_0^t \Big(\sqrt{\boldsymbol{S}_u^{-1}} \boldsymbol{X}_u^\top \boldsymbol{a} \big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1}\Big)^\top \Big(\sqrt{\boldsymbol{S}_u^{-1}} \boldsymbol{X}_u^\top \boldsymbol{a} \big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1}\Big)\, du\bigg) \\
&= E\bigg(\int_0^t \big(\boldsymbol{a}^\top \boldsymbol{a}\big)^{-\frac{1}{2}} \boldsymbol{a}^\top \boldsymbol{X}_u \boldsymbol{S}_u^{-1} \boldsymbol{X}_u^\top \boldsymbol{a} \big(\boldsymbol{a}^\top \boldsymbol{a}\big)^{-\frac{1}{2}}\, du\bigg) \\
&= \int_0^t \big(\boldsymbol{a}^\top \boldsymbol{a}\big)^{-\frac{1}{2}} \boldsymbol{a}^\top \boldsymbol{a} \big(\boldsymbol{a}^\top \boldsymbol{a}\big)^{-\frac{1}{2}}\, du \\
&= t\boldsymbol{I}_p < \infty \quad \text{a.s.}
\end{aligned}$$

We have that

$$dW_{t,ij} = \sum_{u=1}^n \sum_{v=1}^d \big[\boldsymbol{a}\big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1}\big]_{v,j} \Big[\sqrt{\boldsymbol{S}_t^{-1}} \boldsymbol{X}_t^\top\Big]_{i,u}\, d\boldsymbol{B}_{t,uv},$$

and we compute

$$\begin{aligned}
&d[W_{\cdot,ij}, W_{\cdot,kl}]_t \\
&= \sum_{u=1}^n \sum_{v=1}^d \Big[\sqrt{\boldsymbol{S}_t^{-1}} \boldsymbol{X}_t^\top\Big]_{i,u} \Big[\sqrt{\boldsymbol{S}_t^{-1}} \boldsymbol{X}_t^\top\Big]_{k,u} \big[\boldsymbol{a}\big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1}\big]_{v,j} \big[\boldsymbol{a}\big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1}\big]_{v,l} \\
&= \Big[\sqrt{\boldsymbol{S}_t^{-1}} \boldsymbol{X}_t^\top \boldsymbol{X}_t \sqrt{\boldsymbol{S}_t^{-1}}\Big]_{i,k} \big[\big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1} \boldsymbol{a}^\top \boldsymbol{a} \big(\sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\big)^{-1}\big]_{j,l}\, dt \\
&= \boldsymbol{I}_{d,ik} \boldsymbol{I}_{d,jl}\, dt \\
&= \mathbf{1}_{i=k} \mathbf{1}_{j=l}\, dt.
\end{aligned}$$

By Theorem 10.4.5, $\boldsymbol{W}$ is a Brownian motion. Finally, we compute

$$\begin{aligned}
d\boldsymbol{S}_t = d\big(\boldsymbol{X}_t^\top \boldsymbol{X}_t\big) &= (d\boldsymbol{X}_t)^\top \boldsymbol{X}_t + \boldsymbol{X}_t^\top\, d\boldsymbol{X}_t + d\big[\boldsymbol{X}^\top, \boldsymbol{X}\big]_t^M \\
&= \big(\boldsymbol{a}^\top\, d\boldsymbol{B}_t^\top + \boldsymbol{b}^\top \boldsymbol{X}_t^\top\, dt\big) \boldsymbol{X}_t + \boldsymbol{X}_t^\top (d\boldsymbol{B}_t \boldsymbol{a} + \boldsymbol{X}_t \boldsymbol{b})\, dt + \boldsymbol{a}^\top d\big[\boldsymbol{B}^\top, \boldsymbol{B}\big]_t^M \boldsymbol{a} \\
&= \boldsymbol{a}^\top\, d\boldsymbol{B}_t^\top \boldsymbol{X}_t + \boldsymbol{b}^\top \boldsymbol{X}_t^\top \boldsymbol{X}_t\, dt + \boldsymbol{X}_t^\top\, d\boldsymbol{B}_t \boldsymbol{a} + \boldsymbol{X}_t^\top \boldsymbol{X}_t \boldsymbol{b}\, dt + \boldsymbol{a}^\top \boldsymbol{a} n\, dt \\
&= \boldsymbol{a}^\top\, d\boldsymbol{B}_t^\top \boldsymbol{X}_t + \boldsymbol{b}^\top \boldsymbol{S}_t\, dt + \boldsymbol{X}_t^\top\, d\boldsymbol{B}_t \boldsymbol{a} + \boldsymbol{S}_t \boldsymbol{b}\, dt + n\boldsymbol{a}^\top \boldsymbol{a}\, dt \\
&= \big(\boldsymbol{b}^\top \boldsymbol{S}_t + \boldsymbol{S}_t \boldsymbol{b}\big)\, dt + \sqrt{\boldsymbol{S}_t}\, d\boldsymbol{W}_t \sqrt{\boldsymbol{a}^\top \boldsymbol{a}} + \sqrt{\boldsymbol{a}^\top \boldsymbol{a}}\, d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{S}_t} + n\boldsymbol{a}^\top \boldsymbol{a}\, dt,
\end{aligned}$$

where we used that

$$d\big[\boldsymbol{B}_{\cdot}^{\top}, \boldsymbol{B}_{\cdot}\big]_{t,ij}^{M} = \sum_{k=1}^{n} d\langle B_{\cdot,ki}, B_{\cdot,kj}\rangle_t = n\mathbf{1}_{i=j}\,dt,$$

which establishes that $\boldsymbol{S}_t$ solves Eq. (11.2.6). □

We now state Theorem 2″ from Bru (1991).

Theorem 11.2.3 *Let $\alpha \in \{1, \dots, d-1\} \cup (d-1, \infty)$, $\boldsymbol{a} \in GL(d)$, $\boldsymbol{b} \in \mathcal{S}_d^-$, $\boldsymbol{s} \in \mathcal{S}_d^+$ and all eigenvalues of $\boldsymbol{s}$ be distinct, and $\boldsymbol{B}_t$ is a $d \times d$ matrix valued Brownian motion, then on $[0, \tau)$, where τ denotes the first time that the eigenvalues of $\boldsymbol{S}_t$ collide, the stochastic differential equation*

$$d\boldsymbol{S}_t = \sqrt{\boldsymbol{S}_t}\,d\boldsymbol{W}_t\sqrt{\boldsymbol{a}^{\top}\boldsymbol{a}} + \sqrt{\boldsymbol{a}^{\top}\boldsymbol{a}}\,d\boldsymbol{W}_t^{\top}\sqrt{\boldsymbol{S}_t} + (\boldsymbol{b}\boldsymbol{S}_t + \boldsymbol{S}_t\boldsymbol{b})\,dt + \alpha\sqrt{\boldsymbol{a}^{\top}\boldsymbol{a}}\,dt,$$

where $\boldsymbol{S}_0 = \boldsymbol{s}$ has a unique weak solution if $\boldsymbol{b}$ and $\sqrt{\boldsymbol{a}^{\top}\boldsymbol{a}}$ commute.

We remind the reader that the preceding examples were all studied in the original paper on Wishart processes, Bru (1991).

Now, we turn again to the Wishart process as discussed in Sect. 11.1. We introduce the process $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\} \in WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$, given by

$$d\boldsymbol{X}_t = \big(\alpha\boldsymbol{a}^{\top}\boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t\boldsymbol{b}^{\top}\big)\,dt + \big(\sqrt{\boldsymbol{X}_t}\,d\boldsymbol{W}_t\boldsymbol{a} + \boldsymbol{a}^{\top}\,d\boldsymbol{W}_t^{\top}\sqrt{\boldsymbol{X}_t}\big),$$

and we firstly investigate the following case, which was investigated in Benabid et al. (2010). It shows how to link a Wishart process to a multidimensional square root process, see also Sect. 6.7. We assume that $\boldsymbol{a}$ and $\boldsymbol{b}$ are diagonal matrices and that the elements of $\boldsymbol{a}$ are positive, whereas the elements of $\boldsymbol{b}$ are negative. Then one can show that the diagonal elements of $\boldsymbol{X}_t$ satisfy

$$dX_{t,ii} = \big(\alpha a_{i,i}^2 + 2b_{i,i}X_{t,ii}\big) + 2a_{i,i}\sum_{k=1}^{d}[\sqrt{X_t}]_{i,k}\,dW_{t,ki}.$$

Now we define for $i \in \{1, \dots, d\}$,

$$B_{t,i} = \int_0^t \sqrt{X_{t,ii}}^{-1}\sum_{k=1}^{d}[\sqrt{\boldsymbol{X}_s}]_{k,i}\,dW_{s,ki}.$$

We have

$$\begin{aligned}
&E\left(\int_0^t (\sqrt{X_{s,ii}})^{-1}\sum_{k=1}^{d}[\sqrt{\boldsymbol{X}_s}]_{k,i}[\sqrt{\boldsymbol{X}_s}]_{k,i}[\sqrt{X_{s,ii}}]^{-1}\,ds\right)\\
&= E\left(\int_0^t (\sqrt{X_{s,ii}})^{-1}[\sqrt{\boldsymbol{X}_s}\sqrt{\boldsymbol{X}_s}]_{i,i}(\sqrt{X_{s,ii}})^{-1}\,ds\right)\\
&= E\left(\int_0^t (\sqrt{X_{s,ii}})^{-1}X_{s,ii}(\sqrt{X_{s,ii}})^{-1}\,ds\right)\\
&= t.
\end{aligned}$$

Hence, $\boldsymbol{B} = (B_1, \ldots, B_d)$ is a vector of d independent Brownian motions, and we obtain

$$dX_{t,ii} = \left(\alpha a_{i,i}^2 + 2b_{i,i} X_{t,ii}\right) dt + 2a_{i,i} \sqrt{X_{t,ii}}\, dB_{t,i}.$$

Consequently, for diagonal matrices $\boldsymbol{a}$ and $\boldsymbol{b}$, the diagonal elements of the Wishart process are square-root processes.

We now discuss how to construct a matrix-valued Wishart process from vector-valued Ornstein-Uhlenbeck processes. This construction will also motivate the first simulation scheme in Sect. 11.3. In particular, we set

$$\boldsymbol{V}_t = \sum_{k=1}^{\beta} \boldsymbol{X}_{t,k} \boldsymbol{X}_{t,k}^\top \in \mathcal{M}_d, \tag{11.2.7}$$

where

$$d\boldsymbol{X}_{t,k} = \boldsymbol{M} \boldsymbol{X}_{t,k}\, dt + \boldsymbol{Q}^\top d\boldsymbol{W}_{t,k}, \quad k = 1, \ldots, \beta, \tag{11.2.8}$$

where $\boldsymbol{M} \in \mathcal{M}_d$, $\boldsymbol{X}_t \in \Re^d$, $\boldsymbol{Q} \in \mathcal{M}_d$, $\boldsymbol{W}_k \in \Re^d$, so that $\boldsymbol{V}_t \in \mathcal{M}_d$, $\boldsymbol{V}_0 = \boldsymbol{v} \in \mathcal{S}_d^+$ and $\beta \geq d + 1$. The following lemma gives the dynamics of $\boldsymbol{V} = \{\boldsymbol{V}_t,\, t \geq 0\}$.

Lemma 11.2.4 *Assume that* $\boldsymbol{V}_t$ *is given by Eq.* (11.2.7), *where* $\boldsymbol{X}_t$ *satisfies Eq.* (11.2.8). *Then*

$$d\boldsymbol{V}_t = \left(\beta \boldsymbol{Q}^\top \boldsymbol{Q} + \boldsymbol{M} \boldsymbol{V}_t + \boldsymbol{V}_t \boldsymbol{M}^\top\right) dt + \sqrt{\boldsymbol{V}_t}\, d\boldsymbol{W}_t \boldsymbol{Q} + \boldsymbol{Q}^\top d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{V}_t},$$

where $\boldsymbol{W} = \{\boldsymbol{W}_t,\, t \geq 0\}$ *is a* $d \times d$ *matrix valued Brownian motion that is determined by*

$$\sqrt{\boldsymbol{V}_t}\, d\boldsymbol{W}_t = \sum_{k=1}^{\beta} \boldsymbol{X}_{t,k}\, d\boldsymbol{W}_{t,k}^\top.$$

Proof We compute

$$\begin{aligned} d\left(\boldsymbol{X}_{t,k} \boldsymbol{X}_{t,k}^\top\right) &= (d\boldsymbol{X}_{t,k}) \boldsymbol{X}_{t,k}^\top + \boldsymbol{X}_{t,k} (d\boldsymbol{X}_{t,k})^\top + d\left[\boldsymbol{X}_k, \boldsymbol{X}_k^\top\right]_t^M \\ &= \left(\boldsymbol{M} \boldsymbol{X}_{t,k} + \boldsymbol{Q}^\top d\boldsymbol{W}_{t,k}\right) \boldsymbol{X}_{t,k}^\top \\ &\quad + \boldsymbol{X}_{t,k} \left(\boldsymbol{X}_{t,k}^\top \boldsymbol{M}^\top dt + d\boldsymbol{W}_{t,k}^\top \boldsymbol{Q}\right) \\ &\quad + \boldsymbol{Q}^\top d\left[\boldsymbol{W}_k, \boldsymbol{W}_k^\top\right]_t^M \boldsymbol{Q} \\ &= \boldsymbol{M} \boldsymbol{X}_{t,k} \boldsymbol{X}_{t,k}^\top dt + \boldsymbol{Q}^\top d\boldsymbol{W}_{t,k} \boldsymbol{X}_{t,k}^\top + \boldsymbol{X}_{t,k} \boldsymbol{X}_{t,k}^\top \boldsymbol{M}^\top dt \\ &\quad + \boldsymbol{X}_{t,k} d\boldsymbol{W}_{t,k}^\top \boldsymbol{Q} + \boldsymbol{Q}^\top \boldsymbol{I}_d \boldsymbol{Q}\, dt, \end{aligned}$$

where we used that

$$d\left[\boldsymbol{W}_{\cdot,k}, \boldsymbol{W}_{\cdot,k}^\top\right]_t^M = \boldsymbol{I}_d\, dt.$$

Hence

$$\begin{aligned} d\boldsymbol{V}_t &= \boldsymbol{M} \sum_{k=1}^{\beta} \boldsymbol{X}_{t,k} \boldsymbol{X}_{t,k}^\top \, dt + \boldsymbol{Q}^\top \sum_{k=1}^{\beta} d\boldsymbol{W}_{t,k} \boldsymbol{X}_{t,k}^\top \\ &\quad + \sum_{k=1}^{\beta} \boldsymbol{X}_{t,k} \boldsymbol{X}_{t,k}^\top \boldsymbol{M}^\top \, dt + \sum_{k=1}^{\beta} \boldsymbol{X}_{t,k} \, d\boldsymbol{W}_{t,k}^\top \boldsymbol{Q} \\ &\quad + \beta \boldsymbol{Q}^\top \boldsymbol{Q} \, dt \\ &= \boldsymbol{M} \boldsymbol{V}_t \, dt + \boldsymbol{V}_t \boldsymbol{M}^\top \, dt + \boldsymbol{Q}^\top \, d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{V}_t} + \sqrt{\boldsymbol{V}_t} \, d\boldsymbol{W}_t \boldsymbol{Q} + \beta \boldsymbol{Q}^\top \boldsymbol{Q} \, dt, \end{aligned}$$

since

$$\sqrt{\boldsymbol{V}_t} d\boldsymbol{W}_t = \sum_{k=1}^{\beta} \boldsymbol{X}_{t,k} \, d\boldsymbol{W}_{t,k}^\top.$$

To complete the proof, we need to show that $\boldsymbol{W}_t$ is a Brownian motion. As before, we use Theorem 10.4.5. We define

$$\boldsymbol{W}_t = \int_0^t \sqrt{\boldsymbol{V}_u^{-1}} \sum_{k=1}^{\beta} \boldsymbol{X}_{t,k} \, d\boldsymbol{W}_{t,k}^\top$$

and it is easily seen that $\boldsymbol{W}$ is a local martingale. Furthermore,

$$\begin{aligned} dW_{t,ij} &= \sum_{m=1}^{d} \left[\sqrt{\boldsymbol{V}_t^{-1}} \right]_{i,m} \sum_{k=1}^{\beta} [\boldsymbol{X}_{t,k}]_m \left[d\boldsymbol{W}_{t,k}^\top \right]_j \\ &= \sum_{k=1}^{\beta} \sum_{m=1}^{d} \left[\sqrt{\boldsymbol{V}_t^{-1}} \right]_{i,m} [\boldsymbol{X}_{t,k}]_m \left[d\boldsymbol{W}_{t,k}^\top \right]_j. \end{aligned}$$

Now we have

$$\begin{aligned} & d[W_{i,j}, W_{k,l}]_t \\ &\quad = \sum_{k'=1}^{\beta} \sum_{m'=1}^{d} \left[\sqrt{\boldsymbol{V}_t^{-1}} \right]_{i,m'} [\boldsymbol{X}_{t,k'}]_{m'} \sum_{m''=1}^{d} \left[\sqrt{\boldsymbol{V}_t^{-1}} \right]_{k,m''} [\boldsymbol{X}_{t,k'}]_{m''} \mathbf{1}_{j=l} \, dt \\ &\quad = \sum_{k'=1}^{\beta} \left[\sqrt{\boldsymbol{V}_t^{-1}} \boldsymbol{X}_{t,k'} \right]_i \left[\boldsymbol{X}_{t,k'}^\top \sqrt{\boldsymbol{V}_t^{-1}} \right]_k \mathbf{1}_{j=l} \, dt \\ &\quad = \sum_{k'=1}^{\beta} \left[\sqrt{\boldsymbol{V}_t^{-1}} \boldsymbol{X}_{t,k'} \boldsymbol{X}_{t,k'}^\top \sqrt{\boldsymbol{V}_t^{-1}} \right]_{i,k} \mathbf{1}_{j=l} \, dt \\ &\quad = \left[\sqrt{\boldsymbol{V}_t^{-1}} \sum_{k'=1}^{\beta} \boldsymbol{X}_{t,k'} \boldsymbol{X}_{t,k'}^\top \sqrt{\boldsymbol{V}_t^{-1}} \right]_{i,k} \mathbf{1}_{j=l} \, dt \\ &\quad = [\boldsymbol{I}_d]_{i,k} \mathbf{1}_{j=l} \, dt \\ &\quad = \mathbf{1}_{i=k} \mathbf{1}_{j=l} \, dt. \end{aligned}$$

□

11.3 Exact and Almost Exact Simulation Schemes for Wishart Processes

In this section, we discuss two simulation schemes for Wishart processes. The first is based on Benabid et al. (2010), Sect. 1.3, the second on Ahdida and Alfonsi (2010), Sect. 2. We also alert the reader to Chap. 2 in Platen and Bruti-Liberati (2010), where the simulation of the process $WIS_d(\boldsymbol{x}, d, \boldsymbol{0}, \boldsymbol{I}_d)$ was discussed.

11.3.1 Change of Measure Approach

First we present the approach from Benabid et al. (2010). Recall that in Sect. 11.2, we showed that if α assumes integer values, we can simulate a Wishart process by simulating vectors of Ornstein-Uhlenbeck processes. The simulation of multi-dimensional Ornstein-Uhlenbeck processes was discussed in Sects. 6.7 and 10.6. Intuitively, the approach can be described as follows: starting under a probability measure P, where the Wishart process is given by its general form in Eq. (11.1.1) with $\alpha \geq d+1$, $\alpha \in \Re$, we aim to find a change of probability measure, so that under the new measure the corresponding value of α, say $\tilde{\alpha}$, assumes integer values, i.e. $\tilde{\alpha} \in \mathcal{N}$ and $\tilde{\alpha} \geq d+1$. Consequently, we can simulate the Wishart process under the new measure by using Ornstein-Uhlenbeck processes, as explained in Sect. 11.2. In particular, following Benabid et al. (2010), we represent α as follows,

$$\alpha = K + 2\nu,$$

where $K = \lfloor \alpha \rfloor \geq d + 1$, where $\lfloor a \rfloor$ denotes the largest integer less than or equal to a, and ν is a real number satisfying $0 \leq \nu \leq \frac{1}{2}$. The next result, Theorem 2 in Benabid et al. (2010), shows how to introduce a new measure, say P^*, under which the Wishart process can be simulated via Ornstein-Uhlenbeck processes.

Theorem 11.3.1 *Let $q = K + \nu - d - 1$. If*

$$\Lambda_T = \frac{dP^*}{dP}\bigg|_{\mathcal{A}_T}$$

defines the Radon-Nikodym derivative of dP^* with respect to dP, then

$$\Lambda_T = \left(\frac{det(\boldsymbol{X}_T)}{det(\boldsymbol{X}_0)}\right)^{-\frac{\nu}{2}} \exp\{\nu T(Tr(\boldsymbol{b}))\} \exp\left\{\frac{\nu q}{2}\int_0^T Tr\big(\boldsymbol{X}_s^{-1}\boldsymbol{a}^\top \boldsymbol{a}\big)\,ds\right\}.$$

Proof The proof is given in Benabid et al. (2010), see Theorem 2. We present here only the basic ideas of the proof. As in Definition 10.5.4 and Theorem 10.5.5, we specify the new measure via

$$\frac{dP^*}{dP}\bigg|_{\mathcal{A}_T} = \exp\left\{-\nu\int_0^T Tr\Big(\sqrt{\boldsymbol{X}_s^{-1}}\,d\boldsymbol{W}_s\boldsymbol{a}\Big) - \frac{\nu^2}{2}\int_0^T Tr\big(\boldsymbol{X}_s^{-1}\boldsymbol{a}^\top\boldsymbol{a}\big)\,ds\right\},$$

where the Wishart process $\boldsymbol{X} \in WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$ satisfies under P

$$d\boldsymbol{X}_t = \left(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t \boldsymbol{b}^\top\right) dt + \sqrt{\boldsymbol{X}_t}\, d\boldsymbol{W}_t \boldsymbol{a} + \boldsymbol{a}^\top d\boldsymbol{W}^\top \sqrt{\boldsymbol{X}_t}.$$

Under the new measure P^*,

$$\boldsymbol{W}_t^* = \nu \int_0^t \sqrt{\boldsymbol{X}_t^{-1}} \boldsymbol{a}^\top dt + \boldsymbol{W}_t,$$

is a Brownian motion, see Benabid et al. (2010). Consequently, the dynamics of $\boldsymbol{X}_t$ are given by

$$\begin{aligned} d\boldsymbol{X}_t &= \left(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t \boldsymbol{b}^\top\right) dt + \sqrt{\boldsymbol{X}_t}\, d\boldsymbol{W}_t \boldsymbol{a} + \boldsymbol{a}^\top d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{X}_t} \\ &= \left(K \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t \boldsymbol{b}^\top\right) dt + \sqrt{\boldsymbol{X}_t}\, d\boldsymbol{W}_t^* \boldsymbol{a} + \boldsymbol{a}^\top \left(d\boldsymbol{W}_t^*\right)^\top \sqrt{\boldsymbol{X}_t}. \end{aligned}$$

As shown in Benabid et al. (2010), Sect. 1.3.1, the dynamics of the determinant of $\boldsymbol{X}_t$ satisfy

$$\begin{aligned} &\log\left(\frac{det(\boldsymbol{X}_T)}{det(\boldsymbol{X}_0)}\right) \\ &= 2T\left(Tr(\boldsymbol{b})\right) + (K - d - 1) \int_0^T Tr\left(\boldsymbol{X}_t^{-1} \boldsymbol{a}^\top \boldsymbol{a}\right) dt + 2 \int_0^T Tr\left(\sqrt{\boldsymbol{X}_t^{-1}}\, d\boldsymbol{W}_t^* \boldsymbol{a}\right). \end{aligned}$$

Substituting, we get

$$\begin{aligned} &\left.\frac{dP^*}{dP}\right|_{\mathcal{A}_T} \\ &= \exp\left\{-\nu \int_0^T Tr\left(\sqrt{\boldsymbol{X}_t^{-1}} d\boldsymbol{W}_t \boldsymbol{a}\right) - \frac{\nu^2}{2} \int_0^T Tr\left(\boldsymbol{X}_t^{-1} \boldsymbol{a}^\top \boldsymbol{a}\right) dt\right\} \\ &= \exp\left\{-\nu \int_0^T Tr\left(\sqrt{\boldsymbol{X}_t^{-1}} d\boldsymbol{W}_t^* \boldsymbol{a}\right) + \frac{\nu^2}{2} \int_0^T Tr\left(\sqrt{\boldsymbol{X}_t^{-1}} \boldsymbol{a}^\top \boldsymbol{a}\right) dt\right\} \\ &= \exp\left\{-\frac{\nu}{2}\left(\log\left(\frac{det(\boldsymbol{X}_T)}{det(\boldsymbol{X}_0)}\right) - 2T\left(Tr(\boldsymbol{b})\right) - (K - d - 1) \int_0^T Tr\left(\boldsymbol{X}_t^{-1} \boldsymbol{a}^\top \boldsymbol{a}\right) dt\right) \right. \\ &\quad \left. + \frac{\nu^2}{2} \int_0^T Tr\left(\sqrt{\boldsymbol{X}_t^{-1}} \boldsymbol{a}^\top \boldsymbol{a}\right) dt\right\} \\ &= \left(\frac{det(\boldsymbol{X}_T)}{det(\boldsymbol{X}_0)}\right)^{-\frac{\nu}{2}} \exp\left\{T\nu Tr(\boldsymbol{b}) + \frac{\nu}{2}(K - d - 1 + \nu) \int_0^T Tr\left(\sqrt{\boldsymbol{X}_t^{-1}} \boldsymbol{a}^\top \boldsymbol{a}\right) dt\right\}. \end{aligned}$$

□

Consequently, if we are interested in computing

$$E_P\left(f(\boldsymbol{X}_T)\right),$$

for a suitable function $f(\cdot)$, we use Theorem 11.3.1 and obtain

$$
\begin{aligned}
&E\big(f(\boldsymbol{X}_T)\big)\\
&\quad = \exp\big\{-\nu T\big(Tr(\boldsymbol{b})\big)\big\} E_{P^*}\left(\left(\frac{det(\boldsymbol{X}_T)}{det(\boldsymbol{X}_0)}\right)^{\frac{\nu}{2}}\right.\\
&\qquad \left.\times \exp\left\{-\frac{\nu}{2}(K+\nu-d-1)\int_0^T Tr\big(\boldsymbol{X}_t^{-1}\boldsymbol{a}^\top\boldsymbol{a}\big)\,dt\right\} f(\boldsymbol{X}_T)\right).
\end{aligned}
$$

The simulation of the integral, which appears in the Radon-Nikodym derivative, can be discretized and approximated as follows:

$$
\int_t^{t+\Delta t} Tr\big(\boldsymbol{X}_s^{-1}\boldsymbol{a}^\top\boldsymbol{a}\big)\,ds \sim \frac{1}{2}\Delta t Tr\big(\big(\boldsymbol{X}_t^{-1}+\boldsymbol{X}_{t+\Delta t}^{-1}\big)\boldsymbol{a}^\top\boldsymbol{a}\big).
$$

11.3.2 An Exact Simulation Method

We now discuss an exact simulation scheme for Wishart processes, which is based on Ahdida and Alfonsi (2010). To produce the result, we firstly recall the characteristic function associated with $WIS_d(\boldsymbol{x},\alpha,\boldsymbol{b},\boldsymbol{a})$. The result was presented in Gouriéroux and Sufana (2004a), see also Gouriéroux and Sufana (2004b), and we point out that this result led to the realization that there are affine processes that do not assume values on the Euclidean state space. Furthermore, we point out that additional Laplace transform identities are presented in Sect. 11.4. We now follow the presentation in Ahdida and Alfonsi (2010).

Proposition 11.3.2 *Let* $\boldsymbol{X}_t \sim WIS_d(\boldsymbol{x},\alpha,\boldsymbol{b},\boldsymbol{a};t)$,

$$
\boldsymbol{q}_t = \int_0^t \exp(s\boldsymbol{b})\boldsymbol{a}^\top\boldsymbol{a}\exp\big(s\boldsymbol{b}^\top\big)\,ds
$$

and $\boldsymbol{m}_t = \exp\{t\boldsymbol{b}\}$. *The Laplace transform of* $\boldsymbol{X}_t$, *for* $\boldsymbol{v}\in\mathcal{D}_{\boldsymbol{b},\boldsymbol{a};t}$, *is given by*

$$
E\big(\exp\big\{Tr(\boldsymbol{v}\boldsymbol{X}_t)\big\}\big) = \frac{\exp\{Tr(\boldsymbol{v}(\boldsymbol{I}_d-2\boldsymbol{q}_t\boldsymbol{v})^{-1}\boldsymbol{m}_t\boldsymbol{x}\boldsymbol{m}_t^\top)\}}{det(\boldsymbol{I}_d-2\boldsymbol{q}_t\boldsymbol{v})^{\frac{\alpha}{2}}}, \tag{11.3.9}
$$

where $\mathcal{D}_{\boldsymbol{b},\boldsymbol{a};t} = \{\boldsymbol{v}\in\mathcal{S}_d, E(\exp\{Tr(\boldsymbol{v}\boldsymbol{X}_t)\})<\infty\}$ *is the set of convergence of the Laplace transform, which is given explicitly by*

$$
\mathcal{D}_{\boldsymbol{b},\boldsymbol{a};t} = \big\{\boldsymbol{v}\in\mathcal{S}_d,\ \forall s\in[0,t], \boldsymbol{I}_d-2\boldsymbol{q}_s\boldsymbol{v}\in GL(d)\big\}.
$$

We remark that for $\boldsymbol{v}=\boldsymbol{v}_{\Re}+\imath\boldsymbol{v}_I$, $\boldsymbol{v}_{\Re}\in\mathcal{D}_{\boldsymbol{b},\boldsymbol{a};t}$ *and* $\boldsymbol{v}_I\in\mathcal{S}_d$, *the Laplace transform in Eq.* (11.3.9) *is well-defined. For* $\boldsymbol{X}_t\sim WIS_d(\boldsymbol{x},\alpha,\boldsymbol{0},\boldsymbol{I}_d^n;t)$, *we have*

$$
E\big(\exp\big\{Tr(\boldsymbol{v}\boldsymbol{X}_T)\big\}\big) = \frac{\exp\{Tr(\boldsymbol{v}(\boldsymbol{I}_d-2t\boldsymbol{I}_d^n\boldsymbol{v})^{-1}\boldsymbol{x})\}}{det(\boldsymbol{I}_d-2t\boldsymbol{I}_d^n\boldsymbol{v})^{\frac{\alpha}{2}}}.
$$

For a proof of Proposition 11.3.2, we refer the reader to Gouriéroux and Sufana (2004a), and also to Ahdida and Alfonsi (2010). We remark that in Lemma 11.4.3

and Corollary 11.4.5, we will discuss the special cases $\boldsymbol{v}_I = \mathbf{0}$, $\boldsymbol{b} = \mathbf{0}$, and $\boldsymbol{a} = \boldsymbol{I}_d$ and $\boldsymbol{a} \in GL(d)$, respectively.

Regarding the exact simulation procedure, we need the following result from linear algebra, which shows how to perform an *extended Cholesky decomposition*.

Lemma 11.3.3 *Let $\boldsymbol{q} \in \overline{\mathcal{S}_d^+}$ be a matrix with rank r. Then there is a permutation matrix $\boldsymbol{p}$, an invertible lower triangular matrix $\boldsymbol{c}_r \in GL(r)$ and $\boldsymbol{k}_r \in \mathcal{M}_{d-r\times r}$ such that*

$$\boldsymbol{p}\boldsymbol{q}\boldsymbol{p}^\top = \boldsymbol{c}\boldsymbol{c}^\top, \quad \boldsymbol{c} = \begin{pmatrix} \boldsymbol{c}_r & \mathbf{0} \\ \boldsymbol{k}_r & \mathbf{0} \end{pmatrix}.$$

The triplet $(\boldsymbol{c}_r, \boldsymbol{k}_r, \boldsymbol{p})$ is called extended Cholesky decomposition of $\boldsymbol{q}$. Besides,

$$\tilde{\boldsymbol{c}} = \begin{pmatrix} \boldsymbol{c}_r & \mathbf{0} \\ \boldsymbol{k}_r & \boldsymbol{I}_{d-r} \end{pmatrix} \in GL(d),$$

and we have

$$\boldsymbol{q} = \left(\tilde{\boldsymbol{c}}^\top \boldsymbol{p}\right)^\top \boldsymbol{I}_d^r \tilde{\boldsymbol{c}}^\top \boldsymbol{p},$$

where $\boldsymbol{I}_d^r = [\mathbf{1}_{i=j\le r}]_{1\le i,j\le d}$ and $r \le d$.

Proof The lemma appeared in this form in Ahdida and Alfonsi (2010), Lemma 33, which refers to Golub and Van Loan (1996), Algorithm 4.2.4. □

We point out that a numerical procedure to obtain such a decomposition can be found in Golub and Van Loan (1996), see Algorithm 4.2.4. When $r = d$, then we can choose $\boldsymbol{p} = \boldsymbol{I}_d$, and $\boldsymbol{c}_r$ is the usual Cholesky decomposition.

The following proposition, which is Proposition 9 in Ahdida and Alfonsi (2010), sits at the heart of the approach. Essentially, it shows that by rescaling, we can represent a general Wishart process $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a}; t)$ as one which satisfies $\boldsymbol{b} = \mathbf{0}$ and $\boldsymbol{a} = \boldsymbol{I}_d^n$, where $n = Rank(\boldsymbol{q}_t)$. This is crucial, as the law of $WIS_d(\boldsymbol{x}, \alpha, \mathbf{0}, \boldsymbol{I}_d^n; t)$ can be simulated exactly, as we demonstrate below. As in Ahdida and Alfonsi (2010), we remark that one can exactly compute $\boldsymbol{\theta}_t$, which appears in Proposition 11.3.4, using Lemma 11.3.3.

Proposition 11.3.4 *Let $t > 0$, $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{M}_d$, and $\alpha \ge d-1$. Then $\boldsymbol{m}_t = \exp\{t\boldsymbol{b}\}$, $\boldsymbol{q}_t = \int_0^t \exp\{s\boldsymbol{b}\}\boldsymbol{a}^\top \boldsymbol{a} \exp\{s\boldsymbol{b}^\top\}\, ds$ and $n = rank(\boldsymbol{q}_t)$, and there is a $\boldsymbol{\theta}_t \in GL(d)$ such that*

$$\boldsymbol{q}_t = t\boldsymbol{\theta}_t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top,$$

and we have

$$WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a}; t) \overset{d}{=} \boldsymbol{\theta}_t WIS_d\left(\boldsymbol{\theta}_t^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top \left(\boldsymbol{\theta}_t^{-1}\right)^\top, \alpha, \mathbf{0}, \boldsymbol{I}_d^n; t\right)\boldsymbol{\theta}_t^\top.$$

Proof We present the proof from Ahdida and Alfonsi (2010), due to the importance of the result. We apply Lemma 11.3.3 to $\boldsymbol{q}_t/t \in \overline{\mathcal{S}_d^+}$ and obtain the extended Cholesky decomposition $(\boldsymbol{c}_n, \boldsymbol{k}_n, \boldsymbol{p})$. Also, we obtain from Lemma 11.3.3 that

$$\tilde{\boldsymbol{c}} = \begin{pmatrix} \boldsymbol{c}_n & \boldsymbol{0} \\ \boldsymbol{k}_n & \boldsymbol{I}_{d-n} \end{pmatrix}.$$

We define

$$\boldsymbol{\theta}_t = \boldsymbol{p}^{-1}\tilde{\boldsymbol{c}},$$

which by Lemma 11.3.3 is invertible. We now get that

$$\boldsymbol{q}_t = t\boldsymbol{\theta}_t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top.$$

Next, we recall that for $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathcal{M}_d$, the following equalities hold

$$det(\boldsymbol{ab}) = det(\boldsymbol{ba}), \qquad Tr(\boldsymbol{ab}) = Tr(\boldsymbol{ba}),$$

and also

$$(\boldsymbol{abc})^{-1} = \boldsymbol{c}^{-1}\boldsymbol{b}^{-1}\boldsymbol{a}^{-1},$$

assuming that $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are invertible. We hence obtain the following string of equalities,

$$\begin{aligned} det(\boldsymbol{I}_d - 2\imath \boldsymbol{q}_t \boldsymbol{v}) &= det\big(\boldsymbol{I}_d - 2\imath t \boldsymbol{\theta}_t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\big) \\ &= det\big(\boldsymbol{\theta}_t\big(\boldsymbol{\theta}_t^{-1} - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\big)\big) \\ &= det\big(\big(\boldsymbol{\theta}_t^{-1} - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\big)\boldsymbol{\theta}_t\big) \\ &= det\big(\boldsymbol{I}_d - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t\big). \end{aligned}$$

Furthermore,

$$\begin{aligned} &Tr\big(\imath \boldsymbol{v}(\boldsymbol{I}_d - 2\imath \boldsymbol{q}_t \boldsymbol{v})^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top\big) \\ &\quad = Tr\big(\imath\big(\boldsymbol{\theta}_t^{-1}\big)^\top \boldsymbol{\theta}_t^\top \boldsymbol{v}\big(\boldsymbol{\theta}_t\boldsymbol{\theta}_t^{-1} - 2\imath t \boldsymbol{\theta}_t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t\boldsymbol{\theta}_t^{-1}\big)^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top\big) \\ &\quad = Tr\big(\imath\big(\boldsymbol{\theta}_t^{-1}\big)^\top \boldsymbol{\theta}_t^\top \boldsymbol{v}\big(\boldsymbol{\theta}_t\big(\boldsymbol{I}_d - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t\big)\boldsymbol{\theta}_t^{-1}\big)^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top\big) \\ &\quad = Tr\big(\imath\big(\boldsymbol{\theta}_t^{-1}\big)^\top \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t\big(\boldsymbol{I}_d - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t\big)^{-1}\boldsymbol{\theta}_t^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top\big) \\ &\quad = Tr\big(\imath \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t\big(\boldsymbol{I}_d - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t\big)^{-1}\boldsymbol{\theta}_t^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top\big(\boldsymbol{\theta}_t^{-1}\big)^\top\big). \end{aligned}$$

We now let $\boldsymbol{X}_t \sim WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a}; t)$ and $\tilde{\boldsymbol{X}}_t \sim WIS_d(\boldsymbol{\theta}_t^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top(\boldsymbol{\theta}_t^{-1})^\top, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n; t)$ and apply Proposition 11.3.2 to obtain

$$\begin{aligned} &E\big(\exp\big\{\imath Tr(\boldsymbol{v}\boldsymbol{X}_t)\big\}\big) \\ &\quad = \frac{\exp\{Tr(\imath \boldsymbol{v}(\boldsymbol{I}_d - 2\boldsymbol{q}_t \imath \boldsymbol{v})^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top)\}}{det(\boldsymbol{I}_d - 2\boldsymbol{q}_t \imath \boldsymbol{v})^{\frac{\alpha}{2}}} \\ &\quad = \frac{\exp\{Tr(\imath \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t(\boldsymbol{I}_d - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t)^{-1}\boldsymbol{\theta}_t^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top(\boldsymbol{\theta}_t^{-1})^\top)\}}{det(\boldsymbol{I}_d - 2\imath t \boldsymbol{I}_d^n \boldsymbol{\theta}_t^\top \boldsymbol{v}\boldsymbol{\theta}_t)^{\frac{\alpha}{2}}} \end{aligned}$$

$$= E\big(\exp\big\{Tr\big(\imath \boldsymbol{\theta}_t^\top \boldsymbol{v} \boldsymbol{\theta}_t \tilde{\boldsymbol{X}}_t\big)\big\}\big)$$
$$= E\big(\exp\big\{Tr\big(\imath \boldsymbol{v} \boldsymbol{\theta}_t \tilde{\boldsymbol{X}}_t \boldsymbol{\theta}_t^\top\big)\big\}\big)$$

completing the proof. □

We remark that Lemma 11.3.3 generalizes the well-known one-dimensional link between a square-root and a squared Bessel process. For $d = 1$, Lemma 11.3.3 gives

$$WIS_1(x, \alpha, b, a; t) = \frac{a^2(\exp\{2b\} - 1)}{2bt} WIS_1\left(\frac{2btx}{a^2(1 - \exp\{-2bt\})}, \alpha, 0, 1; t\right). \tag{11.3.10}$$

This identity can easily be obtained from the results in Sect. 3.1. Let $X = \{X_t,\ t \geq 0\}$ be a $WIS_1(x, \alpha, b, a)$, then

$$dX_t = \big(\alpha a^2 + 2bX_t\big)\,dt + 2a\sqrt{X_t}\,dW_t.$$

From Proposition 3.1.5, it follows that

$$X_t \stackrel{d}{=} \exp\{2bt\}\tilde{X}_{c(t)},$$

where

$$c(t) = \frac{a^2(1 - \exp\{-2bt\})}{2b}$$

and $\tilde{X}$ is a squared Bessel process, i.e. a $WIS_1(x, \alpha, 0, 1)$ process. We hence have established that

$$WIS_1(x, \alpha, b, a; t) \stackrel{d}{=} \exp\{2bt\} WIS_1\big(x, \alpha, 0, 1; c(t)\big).$$

Now we apply the linear time-change, see Proposition 3.1.2,

$$WIS_1\left(x, \alpha, 0, 1; \frac{c(t)}{t} t\right) \stackrel{d}{=} \frac{c(t)}{t} WIS_1\left(\frac{xt}{c(t)}, \alpha, 0, 1; t\right).$$

Hence

$$\begin{aligned} WIS_1(x, \alpha, 0, 1; t) &= \exp\{2bt\} WIS_1\big(x, \alpha, 0, 1; c(t)\big) \\ &= \frac{\exp\{2bt\}c(t)}{t} WIS_1\left(\frac{xt}{c(t)}, \alpha, 0, 1; t\right), \end{aligned}$$

which is Eq. (11.3.10).

We now proceed as follows: from Proposition 11.3.4, it is clear that we can focus on the $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n)$ case. For the generator of $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n)$, we recall a remarkable splitting property from Ahdida and Alfonsi (2010). Having split the operator, we show that each of these operators correspond to an SDE which can be solved explicitly.

11.3.3 A Remarkable Splitting Property

The infinitesimal generator of the Wishart process, or rather the splitting property thereof, plays an important role in the development of an exact simulation scheme. We hence recall this generator, which is a special case e.g. of Corollary 4 in Ahdida and Alfonsi (2010).

Lemma 11.3.5 *On $\mathcal{M}_d$, we associate with $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$ the infinitesimal generator*

$$\boldsymbol{L} = Tr\big(\big[\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\big]\boldsymbol{D}\big) + 2Tr\big(\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\big),$$

where $\boldsymbol{D} = (\frac{\partial}{\partial x_{i,j}})$, $1 \le i, j \le d$.

The following result, which is Proposition 10 in Ahdida and Alfonsi (2010), gives the splitting property of the operator $\boldsymbol{L}$, for the special case $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n)$. Recall that by Proposition 11.3.4, the study of the simulation of Wishart processes can be reduced to this case. We use $\boldsymbol{e}_d^i$ to denote the matrix

$$\boldsymbol{e}_d^n = [\boldsymbol{1}_{i=j=n}]_{1 \le i,j \le d}.$$

We clearly have, $\boldsymbol{I}_d^n = \sum_{i=1}^n \boldsymbol{e}_d^i$.

Theorem 11.3.6 *Let $\boldsymbol{L}$ be the generator associated with the Wishart process $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n)$ and $\boldsymbol{L}_i$ the generator associated with $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^i)$ for $i \in \{1, \dots, d\}$. Then we have*

$$\boldsymbol{L} = \sum_{i=1}^n \boldsymbol{L}_i \quad \text{and } \forall i, j \in \{1, \dots, d\},\ \boldsymbol{L}_i \boldsymbol{L}_j = \boldsymbol{L}_j \boldsymbol{L}_i. \tag{11.3.11}$$

Proof The first part of the proof follows immediately from Lemma 11.3.5, noting that $\boldsymbol{I}_d^n = \sum_{i=1}^n \boldsymbol{e}_d^i$. The commutativity property is established in Appendix C.1 in Ahdida and Alfonsi (2010). □

As stated in Ahdida and Alfonsi (2010), two features of Eq. (11.3.11) are important:

- the operators $\boldsymbol{L}_i$ and $\boldsymbol{L}_j$ are the same up to the exchange of coordinates i and j;
- the processes $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^i)$ and $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n)$ are well defined on $\overline{\mathcal{S}_d^+}$ under the same hypothesis, namely that $\alpha \ge d-1$ and $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$.

The latter property motivates the simulation scheme:

$$\begin{aligned}
\boldsymbol{X}_t^{1,\boldsymbol{x}} &\sim WIS_d\big(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^1; t\big) \\
\boldsymbol{X}_t^{2,\boldsymbol{X}_t^{1,\boldsymbol{x}}} &\sim WIS_d\big(\boldsymbol{X}_t^{1,\boldsymbol{x}}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^2; t\big) \\
&\vdots \\
\boldsymbol{X}_t^{n,\dots \boldsymbol{X}_t^{1,\boldsymbol{x}}} &\sim WIS_d\big(\boldsymbol{X}_t^{n-1,\dots \boldsymbol{X}_t^{1,\boldsymbol{x}}}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^n; t\big).
\end{aligned}$$

Thus, one samples $X_t^{i,\dots X_t^{1,x}}$ according to the distribution at time t of a Wishart process starting from $X_t^{i-1,\dots X_t^{1,x}}$ and with parameters α, $\boldsymbol{a} = \boldsymbol{e}_d^i$ and $\boldsymbol{b} = \boldsymbol{0}$.

Proposition 11.3.7 *Let $X_t^{n,\dots X_t^{1,x}}$ be defined as above. Then*

$$X_t^{n,\dots X_t^{1,x}} \sim WIS_d\big(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n; t\big).$$

Proof For a formal proof, we refer the reader to Ahdida and Alfonsi (2010), here we just present the main ideas of the proof. Consider a smooth function f on $\overline{\mathcal{S}_d^+}$. Then by iterating Itô's formula, one can establish that

$$E\big(f\big(X_t^{\boldsymbol{x}}\big)\big) = \sum_{k=0}^{\infty} t^k \boldsymbol{L}^k f(\boldsymbol{x})/k!.$$

Next, we employ the tower property, to get

$$\begin{aligned} E\big(f\big(X_t^{n,\dots X_t^{1,x}}\big)\big) &= E\big(E\big(f\big(X_t^{n,\dots X_t^{1,x}}\big) | X_t^{n-1,\dots X_t^{1,x}}\big)\big) \\ &= \sum_{k_n=0}^{\infty} \frac{t^{k_n}}{k_n!} E\big(\boldsymbol{L}_n^{k_n} f\big(X_t^{n-1,\dots X_t^{1,x}}\big)\big). \end{aligned}$$

Repeating this argument, we obtain

$$\begin{aligned} E\big(f\big(X_t^{n,\dots X_t^{1,x}}\big)\big) &= \sum_{k_1,\dots,k_n=0}^{\infty} \frac{t^{\sum_{i=1}^n k_i}}{k_1!\dots k_n!} \boldsymbol{L}_1^{k_1} \dots \boldsymbol{L}_n^{k_n} f(\boldsymbol{x}) \\ &= \sum_{k=0}^{\infty} \frac{t^k}{k!} (\boldsymbol{L}_1 + \dots + \boldsymbol{L}_n)^k = E\big(f\big(X_t^{\boldsymbol{x}}\big)\big). \end{aligned} \quad (11.3.12)$$

Equality (11.3.12) relies on the identification of a Cauchy product and one uses the fact that the operators commute. For example, for $n = 2$,

$$\sum_{k_1,k_2=0}^{\infty} \frac{t^{k_1+k_2}}{k_1!k_2!} \boldsymbol{L}_1^{k_1} \boldsymbol{L}_2^{k_2} f(\boldsymbol{x}) = \sum_{k=0}^{\infty} \boldsymbol{c}_k,$$

where

$$\begin{aligned} \boldsymbol{c}_k &= \sum_{l=0}^{k} \boldsymbol{a}_l \boldsymbol{b}_{k-l} \\ &= \sum_{l=0}^{k} \frac{t^l}{l!} \frac{t^{k-l}}{(k-l)!} \boldsymbol{L}_1^l \boldsymbol{L}_2^{k-l} \\ &= \frac{t^k}{k!} \sum_{l=0}^{k} \frac{k!}{l!(k-l)!} \boldsymbol{L}_1^l \boldsymbol{L}_2^{k-l} \end{aligned}$$

$$= \frac{t^k}{k!} \sum_{l=0}^{k} \binom{k}{l} \boldsymbol{L}_1^l \boldsymbol{L}_2^{k-l}$$

$$= \frac{t^k}{k!} (\boldsymbol{L}_1 + \boldsymbol{L}_2)^k. \qquad \square$$

Proposition 11.3.7 shows that if we can simulate $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^k; t)$, for $k \in \{1, \ldots, d\}$, then we can simulate $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n; t)$, which according to Proposition 11.3.4 means that we can simulate $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a}; t)$. The next lemma shows that we can simulate $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^i; t)$ by simulating $WIS_d(\boldsymbol{p}_k \boldsymbol{x} \boldsymbol{p}_k, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^1; t)$, and subsequently changing the first and the kth coordinates, where we use $\boldsymbol{p}_k$ to denote the matrix which changes the first and the kth coordinate. The following lemma formalizes this.

Lemma 11.3.8 *Construct a matrix $\boldsymbol{p}_k \in \mathcal{S}_d$, so that $p_{k,1} = p_{1,k} = p_{i,i} = 1$, for $i \notin \{1, k\}$, and $p_{i,j} = 0$ otherwise. Let the law of $\boldsymbol{X}_t$ be given by $WIS_d(\boldsymbol{p}_k \boldsymbol{x} \boldsymbol{p}_k, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^1; t)$ and the law of $\tilde{\boldsymbol{X}}_t$ by $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^k; t)$. Then*

$$WIS_d\big(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{e}_d^k; t\big) \stackrel{d}{=} \boldsymbol{p}_k WIS_d\big(\boldsymbol{p}_k \boldsymbol{x} \boldsymbol{p}_k, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^1; t\big) \boldsymbol{p}_k.$$

Proof This result is proven in the same way as Proposition 11.3.4. In particular, we use the characteristic function given in Proposition 11.3.2 for the case $\boldsymbol{b} = \boldsymbol{0}$ and $\boldsymbol{a} = \boldsymbol{I}_d^n$. The proof is now easily completed by using the facts

$$\boldsymbol{p}_k \boldsymbol{I}_d^1 \boldsymbol{p}_k = \boldsymbol{e}_d^k \quad \text{and} \quad \boldsymbol{p}_k \boldsymbol{p}_k = \boldsymbol{I}_d,$$

which then allows us to establish that

$$E\big(\exp\{\imath Tr(\boldsymbol{v} \boldsymbol{p}_k \boldsymbol{X}_t \boldsymbol{p}_k)\}\big) = E\big(\exp\{\imath Tr(\boldsymbol{v} \tilde{\boldsymbol{X}}_t)\}\big). \qquad \square$$

11.3.4 Exact Simulation for Wishart Processes

In this subsection, we discuss how to simulate a $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^1)$ process, with $\alpha \geq d - 1$ and $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$. Due to Proposition 11.3.7 and Lemma 11.3.8, this allows us to sample from the distribution of $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n; t)$. As the presentation is easier, we start with the case $d = 2$.

From Lemma 11.3.5, we obtain the following infinitesimal generator of $WIS_2(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_2^1)$. For $\boldsymbol{x} \in \overline{S_2^+}$,

$$\boldsymbol{L} f(\boldsymbol{x}) = \alpha \partial_{1,1} f(\boldsymbol{x}) + 2x_{1,1} \partial_{1,1}^2 f(\boldsymbol{x}) + 2x_{1,2} \partial_{1,1} \partial_{1,2} f(\boldsymbol{x}) + \frac{x_{2,2}}{2} \partial_{1,2}^2 f(\boldsymbol{x}). \tag{11.3.13}$$

This generator is associated with an SDE that can be solved explicitly. As in Ahdida and Alfonsi (2010), we denote by $Z^1 = (Z_t^1, t \geq 0)$ and $Z^2 = \{Z_t^2, t \geq 0\}$

two independent standard Brownian motions. We distinguish the two cases when $x_{2,2} = 0$ and $x_{2,2} > 0$. For $x_{2,2} = 0$, we have that $x_{1,2} = 0$, since $\boldsymbol{x} \in \overline{\mathcal{S}_2^+}$. In fact, one has

$$dX_{t,11} = \alpha\, dt + 2\sqrt{X_{t,11}}dZ_t^1, \qquad dX_{t,12} = 0, \qquad dX_{t,22} = 0, \tag{11.3.14}$$

where $\boldsymbol{X}_0 = \boldsymbol{x}$, has the infinitesimal generator given in Eq. (11.3.13). Clearly, $X_{t,11}$ is a squared Bessel process of dimension α, that can be sampled as discussed in Sect. 3.1.

We now turn to the case where $x_{2,2} > 0$. The SDE

$$dX_{t,11} = \alpha\, dt + 2\sqrt{X_{t,11} - \frac{X_{t,12}^2}{X_{t,22}}}\, dZ_t^1 + 2\frac{X_{t,12}}{X_{t,22}}\, dZ_t^2 \tag{11.3.15}$$

$$dX_{t,12} = \sqrt{X_{t,22}}\, dZ_t^2 \tag{11.3.16}$$

$$dX_{t,22} = 0, \tag{11.3.17}$$

started at $\boldsymbol{X}_0 = \boldsymbol{x}$ has an infinitesimal generator as given in Eq. (11.3.13). This system can be solved explicitly. We introduce auxiliary variables

$$U_{t,11} = X_{t,11} - \frac{(X_{t,12})^2}{X_{t,22}}, \qquad U_{t,12} = \frac{X_{t,12}}{\sqrt{x_{2,2}}}, \qquad U_{t,22} = x_{2,2}, \tag{11.3.18}$$

where $\boldsymbol{U}_0 = \boldsymbol{u}$. An application of the Itô formula confirms that

$$dU_{t,11} = (\alpha - 1)\, dt + 2\sqrt{U_{t,11}}\, dZ_t^1, \qquad dU_{t,12} = dZ_t^2, \qquad U_{t,22} = 0.$$

Hence, $U_{t,11}$ is a squared Bessel process of dimension $\alpha - 1$, and $U_{t,12}$ is a Brownian motion. Consequently, we simulate $X_{t,11}$, $X_{t,12}$ and $X_{t,22}$ by inverting Eq. (11.3.18) to yield

$$X_{t,11} = U_{t,11} + (U_{t,12})^2, \qquad X_{t,12} = U_{t,12}\sqrt{U_{t,22}}, \qquad X_{t,22} = U_{t,22}. \tag{11.3.19}$$

The following proposition summarizes the discussion in this subsection.

Proposition 11.3.9 *Let $\boldsymbol{x} \in \overline{\mathcal{S}_2^+}$. Then the process defined by either Eq.* (11.3.14) *or Eq.* (11.3.16) *when $x_{2,2} = 0$ or $x_{2,2} > 0$ respectively, has its infinitesimal generator given by* (11.3.13)*. Moreover, the SDE given by Eq.* (11.3.16) *has a unique strong solution that is given by Eq.* (11.3.19) *starting from $u_{1,1} = x_{1,1} - \frac{x_{1,2}^2}{x_{2,2}} \geq 0$, $u_{1,2} = \frac{x_{1,2}}{\sqrt{x_{2,2}}}$, $u_{2,2} = x_{2,2}$.*

Proof This result is a special case of Theorem 13 in Ahdida and Alfonsi (2010). □

As noted in Ahdida and Alfonsi (2010), an interesting property of the result in Proposition 11.3.9 is that the squared Bessel process is well-defined when its dimension $\alpha - 1$ satisfies $\alpha - 1 \geq 0$, which is the same condition under which the Wishart process $WIS_2(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_2^1)$ is well defined, $\alpha \geq d - 1$, since $d = 2$. Lastly, we point out that the process $\boldsymbol{U} = \{\boldsymbol{U}_t,\ t \geq 0\}$ has a squared Bessel process on its diagonal and a Brownian motion on the off-diagonal.

We now discuss how to sample from the distribution $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n; t)$, where $d \geq 2$. It is easy to check that for $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n)$, for $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$, the infinitesimal generator is given by

$$\begin{aligned}\boldsymbol{L} f(\boldsymbol{x}) &= \alpha \partial_{1,1} f(\boldsymbol{x}) + 2x_{1,1} \partial_{1,1}^2 f(\boldsymbol{x}) \\ &\quad + 2 \sum_{\substack{1 \leq m \leq d \\ m \neq 1}} x_{1,m} \partial_{1,m} \partial_{1,1} f(\boldsymbol{x}) + \frac{1}{2} \sum_{\substack{1 \leq m,l \leq d \\ m \neq 1, l \neq 1}} x_{m,l} \partial_{1,m} \partial_{1,l} f(\boldsymbol{x}).\end{aligned} \tag{11.3.20}$$

The next theorem, which is Theorem 13 in Ahdida and Alfonsi (2010), shows how to construct an SDE with the same infinitesimal generator as Eq. (11.3.20) and that it can be solved explicitly. Recall that for the case $d = 2$, we distinguished two cases depending on whether $x_{2,2} = 0$ or $x_{2,2} > 0$. For the general case, the SDE depends on the rank of the submatrix $[x_{i,j}]_{2 \leq i,j \leq d}$. We set

$$r = Rank\big([x_{i,j}]_{2 \leq i,j \leq d}\big) \in \{0, \ldots, d-1\}.$$

First we consider the case $\exists \boldsymbol{c}_r \in \mathcal{G}_r$ that is lower triangular, $\boldsymbol{k}_r \in \mathcal{M}_{d-1-r \times r}$, so that

$$[x_{i,j}]_{2 \leq i,j \leq d} = \begin{pmatrix} \boldsymbol{c}_r & \boldsymbol{0} \\ \boldsymbol{k}_r & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{c}_r^\top & \boldsymbol{k}_r^\top \\ v0 & \boldsymbol{0} \end{pmatrix} =: \boldsymbol{c}\boldsymbol{c}^\top. \tag{11.3.21}$$

The following theorem formally applies to the case where $\boldsymbol{X}_0 = \boldsymbol{x}$ satisfies (11.3.21). However, the subsequent Lemma 11.3.11 shows that such a decomposition can always be obtained by permuting the coordinates $\{2, \ldots, d\}$. As in Ahdida and Alfonsi (2010), we also abuse the notation as follows: when $r = 0$, we still assume that Eq. (11.3.21) holds, in particular with $\boldsymbol{c} = \boldsymbol{0}$. When $r = d - 1$, we recover the usual Cholesky decomposition of $[x_{i,j}]_{2 \leq i,j \leq d}$.

Theorem 11.3.10 *Let us consider $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$ such that Eq. (11.3.21) holds. Let $\boldsymbol{Z} = \{\boldsymbol{Z}_t = (Z_t^1, Z_t^2, \ldots, Z_t^{r+1}),\ t \geq 0\}$ be a vector valued standard Brownian motion. Then the following SDE, where $\sum_{k=1}^r = 0$, for $r = 0$,*

$$\begin{aligned} dX_{t,11} &= \alpha\, dt + 2\sqrt{X_{t,11} - \sum_{k=1}^r \left(\sum_{l=1}^r [\boldsymbol{c}_r^{-1}]_{k,l} X_{t,1(l+1)} \right)^2}\, dZ_t^1 \\ &\quad + 2 \sum_{k=1}^r \sum_{l=1}^r [\boldsymbol{c}_r^{-1}]_{k,l} X_{t,1(l+1)}\, dZ_t^{k+1} \\ dX_{t,1i} &= \sum_{k=1}^r c_{i-1,k} dZ_t^{k+1}, \quad i = 2 \ldots, d \\ dX_{t,lk} &= 0, \quad k, l = 2, \ldots, d, \end{aligned}$$

has a unique strong solution $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$ *starting from* $\boldsymbol{x}$. *It assumes values in* $\overline{\mathcal{S}_d^+}$ *and has the infinitesimal generator* $\boldsymbol{L}$ *given in Eq.* (11.3.20). *Moreover, the explicit solution is given by*

$$\boldsymbol{X}_t = \boldsymbol{C} \begin{pmatrix} U_{t,11} + \sum_{k=1}^{r} (U_{t,1(k+1)})^2 & [U_{t,1(l+1)}]_{1\leq l \leq r}^{\top} & \boldsymbol{0} \\ [U_{t,1(l+1)}]_{1\leq l\leq r} & \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{C}^{\top}, \qquad (11.3.22)$$

where

$$dU_{t,11} = (\alpha - r)\, dt + 2\sqrt{U_{t,11}}\, dZ_t^1, \qquad u_{1,1} = x_{1,1} - \sum_{k=1}^{r} u_{1,k+1}^2 \geq 0$$

$$dU_{t,1(l+1)} = dZ_t^{l+1}, \quad 1 \leq l \leq r, \qquad [u_{1,l+1}]_{1\leq l\leq r} = \boldsymbol{c}_r^{-1} [x_{1,l+1}]_{1\leq l\leq r},$$

and

$$\boldsymbol{C} = \begin{pmatrix} 1 & \boldsymbol{0} & \boldsymbol{0} \\ 0 & \boldsymbol{c}_r & \boldsymbol{0} \\ 0 & \boldsymbol{k}_r & \boldsymbol{I}_{d-r-1} \end{pmatrix}.$$

When $r = 0$, then Eq. (11.3.22) should simply be read as

$$\boldsymbol{X}_t = \begin{pmatrix} U_{t,11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Regarding the matrix $\boldsymbol{U}_t$, we point out that the algorithm only accesses the first row and column of this matrix. As in Ahdida and Alfonsi (2010), $\boldsymbol{X}_t$ can be seen as a function of $\boldsymbol{U}_t$ by setting

$$U_{t,ij} = u_{i,j} = x_{i,j}, \quad i, j \geq 2, \qquad U_{t,1i} = u_{1,i} = 0, \quad r + 1 \leq i \leq d.$$

For a proof of Theorem 11.3.10, we refer the reader to Ahdida and Alfonsi (2010). We point out that sampling from the $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^1; t)$ distribution amounts to sampling a non-central chi-squared random variable and a Gaussian random variable. As for the $d = 2$ case, we note that the condition ensuring that the squared Bessel process $U_{1,1}$ is well-defined for all $r \in \{0, \ldots, d-1\}$ is $\alpha - d - 1 \geq 0$, the same as for the Wishart process. We now recall that the procedure in Theorem 11.3.10 assumed that $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$ satisfied Eq. (11.3.21). This assumption can be relaxed using the extended Cholesky decomposition from Lemma 11.3.3.

Lemma 11.3.11 *Let* $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$ *be a* $WIS(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^1)$ *process and* $(\boldsymbol{c}_r, \boldsymbol{k}_r, \boldsymbol{p})$ *be an extended Cholesky decomposition of* $[x_{i,j}]_{2\leq i,j\leq d}$ *obtained from Lemma* 11.3.3. *Then*

$$\boldsymbol{\pi} = \begin{pmatrix} 1 & \boldsymbol{0} \\ 0 & \boldsymbol{p} \end{pmatrix}$$

is a permutation matrix, and

$$\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq\} \stackrel{d}{=} \boldsymbol{\pi}^\top WIS_d\big(\boldsymbol{\pi}\boldsymbol{x}\boldsymbol{\pi}^\top, \alpha, \mathbf{0}, \boldsymbol{I}_d^1\big)\boldsymbol{\pi},$$

and

$$\big[(\boldsymbol{\pi}\boldsymbol{x}\boldsymbol{\pi}^\top)_{i,j}\big]_{2\leq i,j\leq d} = \begin{pmatrix} \boldsymbol{c}_r & \mathbf{0} \\ \boldsymbol{k}_r & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{c}_r^\top & \boldsymbol{k}_r^\top \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

satisfies (11.3.21).

Proof The first part of the proof can be completed in the same way as the proof of Proposition 11.3.4, namely using characteristic functions. The second part of the proof is an immediate consequence of Lemma 11.3.3. □

Hence, using Theorem 11.3.10 and Lemma 11.3.11, we have a simple way of constructing an SDE that has the generator $\boldsymbol{L}$ from (11.3.13) for any initial condition $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$. It means that we can sample exactly the Wishart distribution $WIS_d(\boldsymbol{x}, \alpha, \mathbf{0}, \boldsymbol{I}_d^1; t)$, which we summarize in Algorithm 11.1.

As discussed in Ahdida and Alfonsi (2010), the computational cost of Algorithm 11.1 is $\mathcal{O}(d^3)$, as this is the computational cost of performing the extended Cholesky decomposition.

We recall that the splitting property established in Theorem 11.3.6 means that if we can sample $WIS_d(\boldsymbol{x}, \alpha, \mathbf{0}, \boldsymbol{e}_d^i; t)$, for $i = 1, \ldots, n$, we can sample $WIS_d(\boldsymbol{x}, \alpha, \mathbf{0}, \boldsymbol{I}_d^n; t)$. However, Lemma 11.3.8 established that sampling $WIS_d(\boldsymbol{x}, \alpha, \mathbf{0}, \boldsymbol{e}_d^i; t)$ amounts to sampling $WIS_d(\boldsymbol{x}, \alpha, \mathbf{0}, \boldsymbol{I}_d^1; t)$, which we discussed in Theorem 11.3.10 and Algorithm 11.1. This is illustrated in Algorithm 11.2.

Algorithm 11.2 performs Algorithm 11.1 n times, resulting in a computational complexity of $\mathcal{O}(nd^3)$, which is bounded by $\mathcal{O}(d^4)$. Concluding this section, we present Algorithm 11.3, which shows how to sample $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a}; t)$, which uses Proposition 11.3.4 to reformulate the problem into the one solved by Algorithm 11.2. We remind the reader that this algorithm is applicable if $\alpha \geq d - 1$, which is also the requirement for the existence of a unique weak solution of the SDE (11.1.1) describing the Wishart process.

11.4 Affine Transforms of the Wishart Process

In this section, we discuss the explicit computation of affine transforms associated with Wishart processes. These results are crucial, as they make Wishart processes useful for practical applications. We present two approaches to this problem: the first is based on the linearization procedure presented in Chap. 9. As discussed below, it turns out that the Riccati equations associated with affine transforms of the Wishart process can be linearized allowing us to compute the affine transform. Consequently, we present an alternative approach, which generalizes a result from Bru (1991), and, following Bru, we refer to it as *Cameron-Martin formula*. The section concludes with a comparison of the two approaches.

Algorithm 11.1 Exact Simulation for the operator L_1

Require: $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$, $\alpha \geq d-1$ and $t > 0$

1: Compute the extended Cholesky decomposition $(\boldsymbol{c}_r, \boldsymbol{k}_r, \boldsymbol{p})$ of $[x_{i,j}]_{2\leq i,j\leq d}$ given by Lemma 11.3.3, $r \in \{0, \ldots, d-1\}$

2: Set

$$\boldsymbol{\pi} = \begin{pmatrix} 1 & \mathbf{0} \\ 0 & \boldsymbol{p} \end{pmatrix},$$
$$\tilde{\boldsymbol{x}} = \boldsymbol{\pi}\boldsymbol{x}\boldsymbol{\pi}^\top,$$
$$[u_{1,l+1}]_{1\leq l\leq r} = \boldsymbol{c}_r^{-1}[\tilde{\boldsymbol{x}}_{1,l+1}]_{1\leq l\leq r},$$
$$u_{1,1} = \tilde{x}_{1,1} - \sum_{k=1}^{r}(u_{1,k+1})^2 \geq 0.$$

3: Sample independently r normal variates $G_2, \ldots, G_{r+1} \sim N(0,1)$ and a non-central chi-square random variate $\chi^2_{\alpha-r}(\frac{u_{1,1}}{t})$, i.e. a non-central chi-square distributed random variable with $\alpha - r$ degrees of freedom and non-centrality parameter $\frac{u_{1,1}}{t}$.

4: Set $U_{t,1(l+1)} = u_{1,l+1} + \sqrt{t}G_{l+1}$

5: Set $U_{t,11} = t\chi^2_{\alpha-r}(\frac{u_{1,1}}{t})$

6: **return** $X =$

$$\boldsymbol{\pi}^\top \boldsymbol{C} \begin{pmatrix} U_{t,11} + \sum_{k=1}^{r}(U_{t,1(k+1)})^2 & [U_{t,1(l+1)}]_{1\leq l\leq r}^\top & \mathbf{0} \\ [U_{t,1(l+1)}]_{1\leq l\leq r} & \boldsymbol{I}_r & \mathbf{0} \\ 0 & \mathbf{0} & \mathbf{0} \end{pmatrix} \boldsymbol{C}^\top \boldsymbol{\pi},$$

where

$$\boldsymbol{C} = \begin{pmatrix} 1 & \mathbf{0} & \mathbf{0} \\ 0 & \boldsymbol{c}_r & \mathbf{0} \\ 0 & \boldsymbol{k}_r & \boldsymbol{I}_{d-r-1} \end{pmatrix}.$$

Algorithm 11.2 Exact Simulation for $WIS_d(\boldsymbol{x}, \alpha, \mathbf{0}, \boldsymbol{I}_d^n; t)$

Require: $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$, $n \leq d$, $\alpha \geq d-1$ and $t > 0$

1: Set $\boldsymbol{y} = \boldsymbol{x}$.

2: **for** $k = 1$ to n **do**

3: Construct the permutation matrix $\boldsymbol{p}$ by setting $p_{k,1} = p_{1,k} = p_{i,i} = 1$ for $i \notin \{1,k\}$, and $p_{i,j} = 0$ otherwise.

4: Set $\boldsymbol{y} = \boldsymbol{p}\boldsymbol{Y}\boldsymbol{p}$, where $\boldsymbol{Y}$ is sampled according to $WIS_d(\boldsymbol{p}\boldsymbol{y}\boldsymbol{p}, \alpha, \mathbf{0}, \boldsymbol{I}_d^1; t)$ by using Algorithm 11.1.

5: **end for**

6: **return** $X = \boldsymbol{y}$.

Algorithm 11.3 Exact Simulation for $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a}; t)$

Require: $\boldsymbol{x} \in \overline{\mathcal{S}_d^+}$, $\alpha \geq d-1$, $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{M}_d$ and $t > 0$

1: Calculate $\boldsymbol{q}_t = \int_0^t \exp\{s\boldsymbol{b}\}\boldsymbol{a}^\top \boldsymbol{a} \exp\{s\boldsymbol{b}^\top\}\, ds$ and $(\boldsymbol{c}_n, \boldsymbol{k}_n, \boldsymbol{p})$ an extended Cholesky decomposition of $\boldsymbol{q}_t/t$.

2: Set

$$\boldsymbol{\theta}_t = \boldsymbol{p}^{-1} \begin{pmatrix} \boldsymbol{c}_n & \boldsymbol{0} \\ \boldsymbol{k}_n & \boldsymbol{I}_{d-n} \end{pmatrix}$$

and $\boldsymbol{m}_t = \exp\{t\boldsymbol{b}\}$.

3: **return** $\boldsymbol{X} = \boldsymbol{\theta}_t \boldsymbol{Y} \boldsymbol{\theta}_t^\top$, where $\boldsymbol{Y} \sim WIS_d(\boldsymbol{\theta}_t^{-1}\boldsymbol{m}_t \boldsymbol{x} \boldsymbol{m}_t^\top (\boldsymbol{\theta}_t^{-1})^\top,\ \alpha,\ \boldsymbol{\theta},\ \boldsymbol{I}_d^n; t)$ is sampled by Algorithm 11.2.

11.4.1 Linearization Applied to Wishart Processes

We assume the following dynamics for the Wishart process,

$$d\boldsymbol{X}_t = \left(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t \boldsymbol{b}^\top\right) dt + \left(\sqrt{\boldsymbol{X}_t}\, d\boldsymbol{W}_t \boldsymbol{a} + \boldsymbol{a}^\top\, d\boldsymbol{W}_t \sqrt{\boldsymbol{X}_t}\right), \qquad (11.4.23)$$

and the infinitesimal generator is given by

$$\boldsymbol{L} = Tr\left(\left[\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\right]\boldsymbol{D}\right) + 2Tr\left(\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\right),$$

see Lemma 11.3.5. As in Chap. 9, we study the discounted conditional characteristic function,

$$\begin{aligned} &\Psi_{\boldsymbol{X}}(\boldsymbol{u}, \boldsymbol{x}, t, \tau) \\ &= E\left(\exp\left\{-\int_t^T \left(\eta_0 + Tr(\boldsymbol{\eta}\boldsymbol{X}_s)\right) ds\right\} \exp\left\{Tr(\imath \boldsymbol{u}\boldsymbol{X}_T)\right\} | \mathcal{A}_t\right) \\ &= \exp\left\{V^0(\tau, \imath\boldsymbol{u}) - Tr\left(\mathbf{V}(\tau, \imath\boldsymbol{u})\boldsymbol{X}_t\right)\right\}, \end{aligned}$$

where $\tau = T - t$. The Feynman-Kac argument now yields, where we use $\Psi = \Psi_{\boldsymbol{X}}(\boldsymbol{u}, \boldsymbol{x}, t, \tau)$,

$$\begin{aligned} \frac{\partial \Psi}{\partial \tau} &= \boldsymbol{L}\Psi - \left(\eta_0 + Tr(\boldsymbol{\eta}\boldsymbol{x})\right) \\ &= Tr\left(\left(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\right)\boldsymbol{D}\Psi + 2\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\Psi\right) \\ &\quad - \left(\eta_0 + Tr(\boldsymbol{\eta}\boldsymbol{x})\right). \end{aligned}$$

On the other hand,

$$\frac{\partial \Psi}{\partial \tau} = \frac{d}{d\tau} V^0(\tau) - Tr\left(\frac{d}{d\tau}\mathbf{V}(\tau)\boldsymbol{x}\right),$$

which yields

$$\begin{aligned}
&\frac{dV^0(\tau)}{d\tau} - Tr\left(\frac{d}{d\tau}\mathbf{V}(\tau)\boldsymbol{x}\right)\\
&\quad = Tr\big(\big(\alpha \boldsymbol{a}\boldsymbol{a}^\top + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\big)\boldsymbol{D}\Psi + 2\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{D}\Psi\big)\\
&\qquad - \big(\eta_0 + Tr(\boldsymbol{\eta}\boldsymbol{x})\big)\\
&\quad = Tr\big(\big(\alpha \boldsymbol{a}^\top\boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\ \big)(-\mathbf{V}) + 2\boldsymbol{x}\mathcal{V}\boldsymbol{a}^\top\boldsymbol{a}\mathbf{V}\big)\\
&\qquad - \big(\eta_0 + Tr(\boldsymbol{\eta}\boldsymbol{x})\big),
\end{aligned}$$

subject to the initial conditions

$$V^0(0) = 0, \qquad \mathbf{V}(0) = -\imath\boldsymbol{u}.$$

By identifying the coefficients of $\boldsymbol{x}$, we obtain the matrix Riccati ODE satisfied by $\mathbf{V}(\tau)$:

$$-\frac{d}{d\tau}\mathbf{V}(\tau) = -\mathbf{V}(\tau)\boldsymbol{b} - \boldsymbol{b}^\top\mathbf{V}(\tau) + 2\mathbf{V}(\tau)\boldsymbol{a}^\top\boldsymbol{a}\mathbf{V}(\tau) - \boldsymbol{\eta}, \tag{11.4.24}$$

and

$$\frac{dV^0(\tau)}{d\tau} = Tr\big(\alpha\boldsymbol{a}^\top\boldsymbol{a}\big(-\mathbf{V}(\tau)\big)\big) - \eta_0. \tag{11.4.25}$$

From Eq. (11.4.24) we get

$$\frac{d\mathbf{V}(\tau)}{d\tau} = \mathbf{V}(\tau)\boldsymbol{b} + \boldsymbol{b}^\top\mathbf{V}(\tau) - 2\mathbf{V}(\tau)\boldsymbol{a}^\top\boldsymbol{a}\mathbf{V}(\tau) + \boldsymbol{\eta}. \tag{11.4.26}$$

We now employ the linearization idea from Chap. 9, and set

$$\mathbf{V}(\tau) = \mathbf{F}(\tau)^{-1}\mathbf{G}(\tau),$$

where $\mathbf{F}(\tau) \in GL(d)$ and $\mathbf{G}(\tau) \in \mathcal{M}_d$. Now

$$\frac{d}{d\tau}\big(\mathbf{F}(\tau)\mathbf{V}(\tau)\big) - \left(\frac{d}{d\tau}\mathbf{F}(\tau)\right)\mathbf{V}(\tau) = \mathbf{F}(\tau)\frac{d}{d\tau}\mathbf{V}(\tau),$$

and substituting (11.4.26), we get

$$\begin{aligned}
\frac{d}{d\tau}\big(\mathbf{F}(\tau)\mathbf{V}(\tau)\big) - \frac{d}{d\tau}\mathbf{F}(\tau)\mathbf{V}(\tau) &= \mathbf{F}(\tau)\mathbf{V}(\tau)\boldsymbol{b} + \mathbf{F}(\tau)\boldsymbol{b}^\top\mathbf{V}\\
&\quad - 2\mathbf{F}(\tau)\mathbf{V}(\tau)\boldsymbol{a}^\top\boldsymbol{a}\mathbf{V}(\tau) + \mathbf{F}(\tau)\boldsymbol{\eta}.
\end{aligned}$$

Matching coefficients, we obtain

$$\begin{aligned}
\frac{d}{d\tau}\mathbf{G}(\tau) &= \mathbf{G}(\tau)\boldsymbol{b} + \mathbf{F}(\tau)\boldsymbol{\eta}\\
\frac{d}{d\tau}\mathbf{F}(\tau) &= -\mathbf{F}(\tau)\boldsymbol{b}^\top + 2\mathbf{G}(\tau)\boldsymbol{a}^\top\boldsymbol{a},
\end{aligned} \tag{11.4.27}$$

which can be written as

$$\frac{d}{d\tau}\big[\mathbf{G}(\tau) \quad \mathbf{F}(\tau)\big] = \big[\mathbf{G}(\tau) \quad \mathbf{F}(\tau)\big]\begin{bmatrix} \boldsymbol{b} & 2\boldsymbol{a}^\top\boldsymbol{a}\\ \boldsymbol{\eta} & -\boldsymbol{b}^\top \end{bmatrix}.$$

The solution is obtained through exponentiation,

$$\begin{aligned}\begin{bmatrix}\mathbf{G}(\tau) & \mathbf{F}(\tau)\end{bmatrix} &= \begin{bmatrix}\mathbf{G}(0) & \mathbf{F}(0)\end{bmatrix}\exp\left\{\tau\begin{bmatrix}\boldsymbol{b} & 2\boldsymbol{a}^\top\boldsymbol{a}\\ \boldsymbol{\eta} & -\boldsymbol{b}^\top\end{bmatrix}\right\}\\ &= \begin{bmatrix}\mathbf{V}(0) & \boldsymbol{I}_d\end{bmatrix}\exp\left\{\tau\begin{bmatrix}\boldsymbol{b} & 2\boldsymbol{a}^\top\boldsymbol{a}\\ \boldsymbol{\eta} & -\boldsymbol{b}^\top\end{bmatrix}\right\}\\ &= \begin{bmatrix}\big(\mathbf{V}(0)\boldsymbol{A}_{11}(\tau)+\boldsymbol{A}_{21}(\tau)\big) & \big(\mathbf{V}(0)\boldsymbol{A}_{12}(\tau)+\boldsymbol{A}_{22}(\tau)\big)\end{bmatrix},\end{aligned}$$

where we use the notation

$$\begin{bmatrix}\boldsymbol{A}_{11}(\tau) & \boldsymbol{A}_{12}(\tau)\\ \boldsymbol{A}_{21}(\tau) & \boldsymbol{A}_{22}(\tau)\end{bmatrix} := \exp\left\{\tau\begin{pmatrix}\boldsymbol{b} & 2\boldsymbol{a}^\top\boldsymbol{a}\\ \boldsymbol{\eta} & -\boldsymbol{b}^\top\end{pmatrix}\right\}$$

for the matrix exponential. Hence

$$\begin{aligned}\mathbf{V}(\tau) &= \big[\mathbf{V}(0)\boldsymbol{A}_{12}(\tau)+\boldsymbol{A}_{22}(\tau)\big]^{-1}\big[\mathbf{V}(0)\boldsymbol{A}_{11}(\tau)+\boldsymbol{A}_{21}(\tau)\big]\\ &= \big[-\imath\boldsymbol{u}\boldsymbol{A}_{12}(\tau)+\boldsymbol{A}_{22}(\tau)\big]^{-1}\big[-\imath\boldsymbol{u}\boldsymbol{A}_{11}(\tau)+\boldsymbol{A}_{21}(\tau)\big],\end{aligned}$$

since $\mathbf{V}(0) = -\imath\boldsymbol{u}$. As usual, a direct integration allows us to compute

$$\frac{d}{d\tau}V^0(\tau) = -Tr\big(\alpha\boldsymbol{a}^\top\boldsymbol{a}\mathbf{V}(\tau)\big) - \eta_0, \tag{11.4.28}$$

which implies that

$$V^0(\tau) = -\int_0^\tau Tr\big(\alpha\boldsymbol{a}^\top\boldsymbol{a}\mathbf{V}(s)\big)\,ds - \eta_0\tau. \tag{11.4.29}$$

Performing the integration in (11.4.29) can be cumbersome, hence we employ the following technique from Da Fonseca et al. (2008c). Equation (11.4.27) can be rewritten as

$$\frac{1}{2}\left(\frac{d}{d\tau}\mathbf{F}(\tau)+\mathbf{F}(\tau)\boldsymbol{b}^\top\right)\big(\boldsymbol{a}^\top\boldsymbol{a}\big)^{-1} = \mathbf{G}(\tau)$$

and from

$$\mathbf{V}(\tau) = \mathbf{F}^{-1}(\tau)\mathbf{G}(\tau),$$

we obtain

$$\mathbf{F}(\tau)\mathbf{V}(\tau) = \frac{1}{2}\left(\frac{d}{d\tau}\mathbf{F}(\tau)+\mathbf{F}(\tau)\boldsymbol{b}^\top\right)\big(\boldsymbol{a}^\top\boldsymbol{a}\big)^{-1},$$

which is equivalent to

$$\mathbf{V}(\tau) = \frac{1}{2}\left(\mathbf{F}^{-1}(\tau)\frac{d}{d\tau}\mathbf{F}(\tau)+\boldsymbol{b}^\top\right)\big(\boldsymbol{a}^\top\boldsymbol{a}\big)^{-1},$$

which we substitute into (11.4.28) to obtain,

$$\frac{d}{d\tau}V^0(\tau) = -Tr\left(\alpha \boldsymbol{a}^\top \boldsymbol{a}\frac{1}{2}\left(\mathbf{F}^{-1}(\tau)\frac{d}{d\tau}\mathbf{F}(\tau) + \boldsymbol{b}^\top\right)\left(\boldsymbol{a}^\top \boldsymbol{a}\right)^{-1}\right) - \eta_0$$
$$= -\frac{\alpha}{2}Tr\left(\mathbf{F}^{-1}(\tau)\frac{d}{d\tau}\mathbf{F}(\tau) + \boldsymbol{b}^\top\right) - \eta_0,$$

which gives

$$V^0(\tau) = -\frac{\alpha}{2}Tr\left(\log\left(\mathbf{F}(\tau)\right) + \boldsymbol{b}^\top \tau\right) - \eta_0 \tau.$$

We conclude that the solution can be explicitly represented in terms of blocks of a matrix exponential. Before discussing this solution further, we present a competing method from Gnoatto and Grasselli (2011) and conclude this section with a comparison.

11.4.2 Cameron-Martin Formula

In this subsection, we present an alternative derivation of the Laplace transform. The result is presented in Gnoatto and Grasselli (2011), and it generalizes a result from Bru (1991), namely Eq. (4.7) in Bru (1991). We first state the result and then compare it with the one from the preceding subsection.

Theorem 11.4.1 *Let* $\boldsymbol{X} \in WIS_d(\boldsymbol{X}_0, \alpha, \boldsymbol{b}, \boldsymbol{a})$, *assume that* $\boldsymbol{a} \in GL(d)$,

$$\boldsymbol{b}^\top\left(\boldsymbol{a}^\top \boldsymbol{a}\right)^{-1} = \left(\boldsymbol{a}^\top \boldsymbol{a}\right)^{-1}\boldsymbol{b},$$

let $\alpha \geq d+1$, *and define the set of convergence of the Laplace transform*

$$\mathcal{D}_t = \left\{\boldsymbol{w}, \boldsymbol{v} \in \mathcal{S}_d \colon E\left(\exp\left\{-Tr\left(\boldsymbol{w}\boldsymbol{X}_t + \int_0^t \boldsymbol{v}\boldsymbol{X}_u\, du\right)\right\}\right) < \infty\right\}.$$

Then for all $\boldsymbol{w}, \boldsymbol{v} \in \mathcal{D}_t$ *the joint moment generating function of the process and its integral is given by*:

$$E\left(\exp\left\{-Tr\left(\boldsymbol{w}\boldsymbol{X}_t + \int_0^t \boldsymbol{v}\boldsymbol{X}_u\, du\right)\right\}\right)$$
$$= det\left(\exp\{-\boldsymbol{b}t\}\left(\cosh(\sqrt{\bar{\boldsymbol{v}}}t) + \sinh(\sqrt{\bar{\boldsymbol{v}}}t)\boldsymbol{k}\right)\right)^{\frac{\alpha}{2}}$$
$$\times \exp\left\{Tr\left(\left(\frac{\boldsymbol{a}^{-1}\sqrt{\bar{\boldsymbol{v}}}\boldsymbol{k}(\boldsymbol{a}^\top)^{-1}}{2} - \frac{(\boldsymbol{a}^\top \boldsymbol{a})^{-1}\boldsymbol{b}}{2}\right)\boldsymbol{X}_0\right)\right\},$$

where the matrices $\boldsymbol{k}$, $\bar{\boldsymbol{v}}$, $\bar{\boldsymbol{w}}$ *are given by*

$$\boldsymbol{k} = -\left(\sqrt{\bar{\boldsymbol{v}}}\cosh(\sqrt{\bar{\boldsymbol{v}}}t) + \bar{\boldsymbol{w}}\sinh(\sqrt{\bar{\boldsymbol{v}}}t)\right)^{-1}\left(\sqrt{\bar{\boldsymbol{v}}}\sinh(\sqrt{\bar{\boldsymbol{v}}}t) + \bar{\boldsymbol{w}}\cosh(\sqrt{\bar{\boldsymbol{v}}}t)\right),$$
$$\bar{\boldsymbol{v}} = \boldsymbol{a}\left(2\boldsymbol{v} + \boldsymbol{b}^\top \boldsymbol{a}^{-1}\left(\boldsymbol{a}^\top\right)^{-1}\boldsymbol{b}\right)\boldsymbol{a}^\top,$$
$$\bar{\boldsymbol{w}} = \boldsymbol{a}\left(2\boldsymbol{w} - \left(\boldsymbol{a}^\top \boldsymbol{a}\right)^{-1}\boldsymbol{b}\right)\boldsymbol{a}^\top.$$

The proof of Theorem 11.4.1 can be found in Gnoatto and Grasselli (2011), it consists of three parts. The first establishes the result for $WIS_d(\boldsymbol{X}_0, \alpha, \boldsymbol{0}, \boldsymbol{I}_d)$, and consequently the conclusion of the first part is extended to $WIS_d(\boldsymbol{X}_0, \alpha, \boldsymbol{0}, \boldsymbol{a})$. The final part establishes the result for $WIS_d(\boldsymbol{X}_0, \alpha, \boldsymbol{b}, \boldsymbol{a})$. Here, we present the first two parts of the proof, for the third part, we refer the reader to Gnoatto and Grasselli (2011). Before proceeding with the proof, we recall two results from Bru (1991). The next result is Proposition 2 in Bru (1991).

Lemma 11.4.2 *If* $\boldsymbol{\Phi} : \Re^+ \to \overline{\mathcal{S}_d^+}$ *is continuous, constant on* $[t, \infty)$ *and such that its right derivative (in the distribution sense)* $\boldsymbol{\Phi}' : \Re^+ \to \overline{\mathcal{S}_d^-}$ *is continuous, with* $\boldsymbol{\Phi}(0) = \boldsymbol{I}_d$, *and* $\boldsymbol{\Phi}'(t) = \boldsymbol{0}$, *then for every Wishart process* $\boldsymbol{X} \in WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d^n)$, *we have*

$$E\left(\exp\left\{-\frac{1}{2}Tr\left(\int_0^t \boldsymbol{\Phi}''(s)\boldsymbol{\Phi}^{-1}(s)\boldsymbol{X}_s\,ds\right)\right\}\right)$$
$$= \left(det\,\boldsymbol{\Phi}(t)\right)^{\frac{\alpha}{2}} \exp\left\{\frac{1}{2}Tr\left(\boldsymbol{X}_0\boldsymbol{\Phi}^+(0)\right)\right\},$$

where

$$\boldsymbol{\Phi}^+(0) = \lim_{t \searrow t} \boldsymbol{\Phi}'(t).$$

Also, we recall Theorem 3 from Bru (1991), see also Proposition 11.3.2.

Lemma 11.4.3 *Let* $\boldsymbol{X}$ *be a* $WIS_d(\boldsymbol{X}_0, \alpha, \boldsymbol{0}, \boldsymbol{I}_d)$ *process, where* $\alpha \geq d + 1$, *and* $\boldsymbol{u} \in \overline{\mathcal{S}_d^+}$, *then*

$$E\left(\exp\{-Tr(\boldsymbol{u}\boldsymbol{X}_t)\}\right) = \left(det(\boldsymbol{I}_d + 2t\boldsymbol{u})\right)^{\frac{\alpha}{2}} \exp\left\{-Tr\left(\boldsymbol{X}_0(\boldsymbol{I}_d + 2t\boldsymbol{u})^{-1}\boldsymbol{u}\right)\right\}.$$

We now establish the result from Theorem 11.4.1 for a $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d)$ process, which is Proposition 2 in Gnoatto and Grasselli (2011).

Proposition 11.4.4 *Let* $\boldsymbol{\Sigma} \in WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d)$, *then*

$$E\left(\exp\left\{-\frac{1}{2}Tr\left(\boldsymbol{w}\boldsymbol{\Sigma}_t + \int_0^t \boldsymbol{v}\boldsymbol{\Sigma}_s\,ds\right)\right\}\right)$$
$$= det\left(\cosh(\sqrt{\boldsymbol{v}}t) + \sinh(\sqrt{\boldsymbol{v}}t)\boldsymbol{k}\right)^{\frac{\alpha}{2}} \exp\left\{\frac{1}{2}Tr(\boldsymbol{\Sigma}_0\sqrt{\boldsymbol{v}}\boldsymbol{k})\right\},$$

where $\boldsymbol{k}$ *is given by*

$$\boldsymbol{k} = -\left(\sqrt{\boldsymbol{v}}\cosh(\sqrt{\boldsymbol{v}}t) + \boldsymbol{w}\sinh(\sqrt{\boldsymbol{v}}t)\right)^{-1}\left(\sqrt{\boldsymbol{v}}\sinh(\sqrt{\boldsymbol{v}}t) + \boldsymbol{w}\cosh(\sqrt{\boldsymbol{v}}t)\right). \tag{11.4.30}$$

Proof By Lemma 11.4.2, we have to solve the ODE:

$$\begin{aligned}\boldsymbol{\Phi}''(s) &= \boldsymbol{v}\boldsymbol{\Phi}(s), \quad s \in (0,t),\\ \boldsymbol{\Phi}'^{-}(t) &= -\boldsymbol{w}\boldsymbol{\Phi}(t),\\ \boldsymbol{\Phi}(0) &= \boldsymbol{I}_d.\end{aligned} \tag{11.4.31}$$

The general solution of (11.4.31) is given by

$$\boldsymbol{\Phi}(s) = \cosh(\sqrt{\boldsymbol{v}}s)\boldsymbol{k}_1 + \sinh(\sqrt{\boldsymbol{v}}s)\boldsymbol{k}.$$

From the condition $\boldsymbol{\Phi}(0) = \boldsymbol{I}_d$, we get $\boldsymbol{k}_1 = \boldsymbol{I}_d$. In order to determine $\boldsymbol{k}$, we look at the boundary condition at $\boldsymbol{\Phi}'^{-}(t)$ and hence obtain

$$\sqrt{\boldsymbol{v}}\sinh(\sqrt{\boldsymbol{v}}t) + \sqrt{\boldsymbol{v}}\cosh(\sqrt{\boldsymbol{v}}t)\boldsymbol{k} = -\boldsymbol{w}\big(\cosh(\sqrt{\boldsymbol{v}}t) + \sinh(\sqrt{\boldsymbol{v}}t)\boldsymbol{k}\big).$$

This yields the value of $\boldsymbol{k}$ given in Eq. (11.4.30). Next, we compute the derivative of $\boldsymbol{\Phi}$,

$$\boldsymbol{\Phi}'(s) = \sqrt{\boldsymbol{v}}\sinh(\sqrt{\boldsymbol{v}}s) + \sqrt{\boldsymbol{v}}\cosh(\sqrt{\boldsymbol{v}}s)\boldsymbol{k},$$

which yields

$$\lim_{s\searrow 0}\boldsymbol{\Phi}'(s) = \sqrt{\boldsymbol{v}}\boldsymbol{k}.$$

Since $\boldsymbol{\Phi}$ is constant on $[t,\infty)$, we obtain that $\boldsymbol{\Phi}(\infty) = \boldsymbol{\Phi}(t)$, which completes the proof. □

Now we attend to the second part.

Corollary 11.4.5 *Let $\boldsymbol{X} \in WIS_d(\boldsymbol{x}, \alpha, 0, \boldsymbol{a})$, where $\alpha \geq d+1$ and $\boldsymbol{a} \in GL(d)$, and let $\boldsymbol{u} \in \overline{\mathcal{S}_d^+}$. Then*

$$\begin{aligned}&E\big(\exp\{-Tr(\boldsymbol{u}\boldsymbol{X}_t)\}\big)\\ &\quad = \big(det\big(\boldsymbol{I}_d + 2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big)\big)^{-\frac{\alpha}{2}}\exp\big\{-Tr\big(\boldsymbol{u}\big(\boldsymbol{I}_d + 2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big)^{-1}\boldsymbol{x}\big)\big\}.\end{aligned}$$

Proof Firstly, we note that since $\boldsymbol{a} \in GL(d)$, $\boldsymbol{a}^\top\boldsymbol{a} \in \mathcal{S}_d^+$, and since $\boldsymbol{u} \in \overline{\mathcal{S}_d^+}$, we have that $\boldsymbol{I}_d + 2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u} \in \mathcal{S}_d^+$. Furthermore, as demonstrated in Sect. 11.2, for $\boldsymbol{\Sigma} \in WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{0}, \boldsymbol{I}_d)$, we can set

$$\boldsymbol{X}_t = \boldsymbol{a}^\top\boldsymbol{\Sigma}_t\boldsymbol{a}, \quad t \geq 0$$

to obtain

$$d\boldsymbol{X}_t = \sqrt{\boldsymbol{X}_t}\,d\tilde{\boldsymbol{W}}_t\boldsymbol{a} + \boldsymbol{a}^\top d\tilde{\boldsymbol{W}}_t^\top\sqrt{\boldsymbol{X}_t} + \alpha\boldsymbol{a}^\top\boldsymbol{a}\,dt,$$

where $d\tilde{\boldsymbol{W}}_t = \sqrt{\boldsymbol{X}_t}^{-1}\boldsymbol{Q}^\top\sqrt{\boldsymbol{\Sigma}_t}\,d\boldsymbol{W}_t$ is a Brownian motion, and $\boldsymbol{W}$ denotes the Brownian motion driving $\boldsymbol{\Sigma}$. We apply Lemma 11.4.3 and use the fact that $\boldsymbol{\Sigma}_0 = (\boldsymbol{a}^\top)^{-1}\boldsymbol{X}_0\boldsymbol{a}^{-1}$ to obtain

$$\begin{aligned}E\big(\exp\big\{-Tr(\boldsymbol{u}\boldsymbol{X}_t)\big\}\big) &= E\big(\exp\big\{-Tr\big(\boldsymbol{u}\big(\boldsymbol{a}^\top\boldsymbol{\Sigma}_t\boldsymbol{a}\big)\big)\big\}\big)\\ &= \big(det\big(\boldsymbol{I}_d+2t\boldsymbol{a}\boldsymbol{u}\boldsymbol{a}^\top\big)\big)^{-\frac{\alpha}{2}}\\ &\quad\times\exp\big\{-Tr\big(\big(\boldsymbol{a}^\top\big)^{-1}\boldsymbol{X}_0\boldsymbol{a}^{-1}\big(\boldsymbol{I}_d+2t\boldsymbol{a}\boldsymbol{u}\boldsymbol{a}^\top\big)^{-1}\boldsymbol{a}\boldsymbol{u}\boldsymbol{a}^\top\big)\big\}\\ &= \big(det\big(\boldsymbol{I}_d+2t\boldsymbol{a}\boldsymbol{u}\boldsymbol{a}^\top\big)\big)^{-\frac{\alpha}{2}}\\ &\quad\times\exp\big\{-Tr\big(\boldsymbol{X}_0\boldsymbol{a}^{-1}\big(\boldsymbol{I}_d+2t\boldsymbol{a}\boldsymbol{u}\boldsymbol{a}^\top\big)^{-1}\boldsymbol{a}\boldsymbol{u}\big)\big\}.\end{aligned}$$

We now use Sylvester's determinant theorem,

$$det(\boldsymbol{I}_d+\boldsymbol{A}\boldsymbol{B}) = det(\boldsymbol{I}_d+\boldsymbol{B}\boldsymbol{A}),$$

to obtain

$$det\big(\boldsymbol{I}_d+2t\boldsymbol{a}\boldsymbol{u}\boldsymbol{a}^\top\big) = det\big(\boldsymbol{I}_d+2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big).$$

Since

$$\boldsymbol{a}^{-1}\big(\boldsymbol{I}_d+2t\boldsymbol{a}\boldsymbol{u}\boldsymbol{a}^\top\big)^{-1}\boldsymbol{a}\boldsymbol{u} = \boldsymbol{u}\big(\boldsymbol{I}_d+2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big)^{-1},$$

we compute

$$\begin{aligned}&E\big(\exp\big\{-Tr(\boldsymbol{u}\boldsymbol{X}_t)\big\}\big)\\ &\quad= \big(det\big(\boldsymbol{I}_d+2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big)\big)^{-\frac{\alpha}{2}}\exp\big\{-Tr\big(\boldsymbol{X}_0\boldsymbol{u}\big(\boldsymbol{I}_d+2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big)^{-1}\big)\big\}\\ &\quad= \big(det\big(\boldsymbol{I}_d+2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big)\big)^{-\frac{\alpha}{2}}\exp\big\{-Tr\big(\boldsymbol{u}\big(\boldsymbol{I}_d+2t\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{u}\big)^{-1}\boldsymbol{X}_0\big)\big\}.\end{aligned}$$ □

We remark that Corollary 11.4.5 is a special case of Proposition 11.3.2. The third step, where one incorporates the drift in Eq. (11.4.23), is completed by employing the Girsanov theorem, we refer the reader to Gnoatto and Grasselli (2011).

11.4.3 A Comparison of the Two Approaches

In this subsection, we recall the discussion in Sect. 3.4 of Gnoatto and Grasselli (2011), which compares the linearization approach to the Cameron-Martin formula. First, in terms of precision and execution speed, the two methods produce identical results. However, the disadvantage of the linearization method is that the functions $\boldsymbol{F}(\tau)$ and $\boldsymbol{G}(\tau)$ are expressed in terms of matrix exponentials, and the matrix exponential depends on the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ of the Wishart process. Furthermore, to obtain the function $V^0(\tau)$, one multiplies the remaining parameter α by the logarithm of $\boldsymbol{F}(\tau)$, and $\boldsymbol{F}(\tau)$ depends on the matrix exponential. As the matrix exponential is a symbolic expression, it means that the linearization method might be less useful if we want to understand the implications of the various model parameters, which is particularly important in applications. The result in Theorem 11.4.1 is strictly explicit, and furthermore it involves exponentials of $d \times d$ matrices, as opposed to the linearization method, which doubles the dimensionality of the problem, resulting in a $2d \times 2d$ matrix. Also, the Cameron-Martin formula does not require the computation of the matrix logarithm. Finally, with regards to the computation of sensitivities, which play an important role in finance, we can expect the Cameron-Martin formula to be more useful.

11.5 Two Heston Multifactor Volatility Models

In this section, we discuss two Heston multifactor volatility models, firstly a single-asset and secondly a multi-asset model, which were presented in Da Fonseca et al. (2007) and Da Fonseca et al. (2008c), respectively. The aim of this section is to illustrate how to exploit the tractability of the Wishart process. For each of the two models, we firstly discuss how to correlate the Brownian motion driving the asset, or assets respectively, and the Brownian motion driving the Wishart process, to retain the affinity of the model. Finally, we find that once we have an explicit representation of the infinitesimal generator, we can immediately employ the approach from Sect. 11.4 to compute the characteristic function. We remark that we employ linearization, as it follows easily from the presentation. However, instead the Cameron-Martin formula could have been used, see Gnoatto and Grasselli (2011), where the two models were studied using the Cameron-Martin formula.

11.5.1 A Single Asset Heston Multifactor Volatility Model

In this subsection, we present a single-asset model, in which we describe the stochastic volatility via a Wishart process. This model can be considered to be the natural extension of the Heston model, as discussed in Sect. 9.5. Following Da Fonseca et al. (2008c), we model the risky asset under an assumed risk-neutral measure via the SDE,

$$\frac{dS_t}{S_t} = r\,dt + Tr(\sqrt{\boldsymbol{X}_t}\,d\boldsymbol{Z}_t), \tag{11.5.32}$$

where r denotes the risk-free interest rate which, for ease of presentation, is assumed to be constant. The process $\boldsymbol{Z} = \{\boldsymbol{Z}_t,\ t \geq 0\}$ is a matrix-valued Brownian motion, $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$ is a $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$ process, given by

$$d\boldsymbol{X}_t = \left(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{b}^\top \boldsymbol{X}_t\right) dt + \left(\sqrt{\boldsymbol{X}_t}\, d\boldsymbol{W}_t \boldsymbol{a} + \boldsymbol{a}^\top\, d\boldsymbol{W}_t^\top \sqrt{\boldsymbol{X}_t}\right), \tag{11.5.33}$$

where $\alpha \geq d - 1$, $\boldsymbol{b} \in \mathcal{M}_d$, and $\boldsymbol{a} \in GL(d)$. Following Da Fonseca et al. (2008c), we assume $\boldsymbol{b} \in \overline{\mathcal{S}_d^-}$, to obtain the mean-reverting behavior of $\boldsymbol{X}$. We now turn to the correlation structure of the Brownian motions $\boldsymbol{Z}$ and $\boldsymbol{W}$. In particular, Da Fonseca et al. (2008c) introduce a correlation matrix $\boldsymbol{R} \in \mathcal{M}_d$ to obtain the following correlation structure,

$$\boldsymbol{Z}_t = \boldsymbol{W}_t \boldsymbol{R}^\top + \boldsymbol{B}_t \sqrt{\boldsymbol{I} - \boldsymbol{R}\boldsymbol{R}^\top}, \quad t \geq 0, \tag{11.5.34}$$

where $\boldsymbol{B} = \{\boldsymbol{B}_t,\ t \geq 0\}$ is a Brownian motion independent of $\boldsymbol{W}$. The next proposition establishes that $\boldsymbol{Z}$ is a Brownian motion.

Proposition 11.5.1 *The process $\boldsymbol{Z} = \{\boldsymbol{Z}_t,\ t \geq 0\}$ defined in Eq. (11.5.34) is a Brownian motion.*

Proof We use Theorem 10.4.5 to obtain the proof. Clearly, the process $\boldsymbol{Z}$ is a local martingale. Furthermore,

$$dZ_{t,ij} = \sum_{k=1}^{d} dW_{t,ik}\, R_{j,k} + \sum_{k=1}^{d} dB_{t,ik}\big(\sqrt{\boldsymbol{I} - \boldsymbol{R}\boldsymbol{R}^\top}\big)_{k,j}.$$

Hence we have

$$\begin{aligned} &d[Z_{\cdot,ij}, Z_{\cdot,kl}]_t \\ &= \left(\sum_{m=1}^{d} R_{j,m} R_{l,m} + \big(\sqrt{\boldsymbol{I} - \boldsymbol{R}\boldsymbol{R}^\top}\big)_{m,j}\big(\sqrt{\boldsymbol{I} - \boldsymbol{R}\boldsymbol{R}^\top}\big)_{m,l}\right) \mathbf{1}_{i=k}\, dt \\ &= \boldsymbol{I}_{j,l} \mathbf{1}_{i=k}\, dt \\ &= \mathbf{1}_{i=k} \mathbf{1}_{j=l}\, dt, \end{aligned}$$

which completes the proof. □

The next result discusses the correlation structure of $\boldsymbol{Z}_t$ and $\boldsymbol{W}_t$.

Proposition 11.5.2 *The covariance of $\boldsymbol{Z}_t$ and $\boldsymbol{W}_t$ is given by*

$$Cov(\boldsymbol{Z}_t,\ \boldsymbol{W}_t) = t\,\boldsymbol{I}_d \otimes \boldsymbol{R}. \tag{11.5.35}$$

Proof From Definition 10.3.7, we have

$$\begin{aligned} Cov(\boldsymbol{Z}_t, \boldsymbol{W}_t) &= E\big(vec\big(\boldsymbol{Z}_t^\top\big)vec\big(\boldsymbol{W}_t^\top\big)^\top\big) - E\big(vec\big(\boldsymbol{Z}_t^\top\big)\big)E\big(vec\big(\boldsymbol{W}_t^\top\big)\big)^\top \\ &= E\big(vec\big(\boldsymbol{R}\boldsymbol{W}_t^\top\big)vec\big(\boldsymbol{W}_t^\top\big)^\top\big). \end{aligned}$$

We find it convenient to denote the i-th row of $\boldsymbol{W}_t$ by $\boldsymbol{w}_i$, and regarding the matrix $\boldsymbol{R}\boldsymbol{W}_t^\top$, we denote its j-th column by $\boldsymbol{r}_j$, so that

$$\boldsymbol{r}_j = \begin{bmatrix} \sum_{k=1}^d R_{1,k} W_{j,k} \\ \sum_{k=1}^d R_{2,k} W_{j,k} \\ \vdots \\ \sum_{k=1}^d R_{d,k} W_{j,k} \end{bmatrix}.$$

Hence

$$
\begin{aligned}
E\big(vec(\boldsymbol{R}\boldsymbol{W}_t^\top)vec(\boldsymbol{W}_t^\top)^\top\big) &= E\left(\begin{bmatrix}\boldsymbol{r}_1\\ \vdots\\ \boldsymbol{r}_d\end{bmatrix}[\boldsymbol{w}_1\cdots\boldsymbol{w}_d]\right)\\
&= E\left(\begin{bmatrix}\boldsymbol{r}_1\boldsymbol{w}_1 & \cdots & \boldsymbol{r}_1\boldsymbol{w}_d\\ \vdots & \ddots & \vdots\\ \boldsymbol{r}_d\boldsymbol{w}_1 & \cdots & \boldsymbol{r}_d\boldsymbol{w}_d\end{bmatrix}\right)\\
&= \begin{bmatrix}tI_{1,1}\boldsymbol{R} & \cdots & tI_{1,d}\boldsymbol{R}\\ \vdots & \ddots & \vdots\\ tI_{d,1}\boldsymbol{R} & \cdots & t_{d,d}\boldsymbol{R}\end{bmatrix} \qquad (11.5.36)\\
&= t\boldsymbol{I}_d\otimes\boldsymbol{R}.
\end{aligned}
$$

To see equality (11.5.36), we consider an element of the matrix $\boldsymbol{r}_l\boldsymbol{w}_m$, say the element $[\boldsymbol{r}_l\boldsymbol{w}_m]_{i,j}$, where $i, j, l, m \in \{1,\dots,d\}$. This element admits the representation

$$\sum_{k=1}^{d} R_{i,k}W_{t,lk}W_{t,mj}.$$

Consequently,

$$E\left(\sum_{k=1}^{d} R_{i,k}W_{t,lk}W_{t,mj}\right) = \begin{cases}0 & \text{for } l\neq m\\ tR_{i,j} & \text{for } l=m.\end{cases} \qquad \square$$

Hence, $\boldsymbol{R}$, which is a $d\times d$ matrix, summarizes the covariance structure, which is, in principle, a matrix of size $d^2\times d^2$. We choose to summarize the covariance structure by $\boldsymbol{R}$, as it preserves the affine structure of the model, which is crucial for analytical tractability.

We now turn to option pricing. It is convenient to work with the logarithm of the stock price, i.e. $Y_t = \log(S_t)$, which satisfies the SDE

$$dY_t = \left(r - \frac{1}{2}Tr(\boldsymbol{X}_t)\right)dt + Tr\Big(\sqrt{\boldsymbol{X}_t}\big(d\boldsymbol{W}_t\,\boldsymbol{R}^\top + d\boldsymbol{B}_t\sqrt{\boldsymbol{I}_d - \boldsymbol{R}\boldsymbol{R}^\top}\big)\Big).$$

As in Da Fonseca et al. (2008c), we work with the infinitesimal generator of the process, which then allows us to employ linearization to compute the Laplace transform. Alternatively, the Cameron-Martin formula could have been employed, we refer the reader to Gnoatto and Grasselli (2011) for this approach. Recall that the Laplace transform is given by

$$
\begin{aligned}
\Psi_{\gamma,t}(\tau) &= E\big(\exp\{\gamma Y_{t+\tau}\}\big)\\
&= \exp\big\{Tr\big(\boldsymbol{A}(\tau)\boldsymbol{X}_t\big) + b(\tau)Y_t + c(\tau)\big\}, \qquad (11.5.37)
\end{aligned}
$$

where $\gamma\in\Re$, $\boldsymbol{A}(\tau)\in\mathcal{M}_d$, $b(\tau)\in\Re$ and $c(\tau)\in\Re$. We use $L_{\boldsymbol{X}}$ to denote the infinitesimal generator of $\boldsymbol{X}$, and $L_{Y,\boldsymbol{X}}$ to denote the infinitesimal generator of $(Y, \boldsymbol{X})$.

Recall from Lemma 11.3.5 that the infinitesimal generator of $\boldsymbol{X}$ is given by

$$L_{\boldsymbol{X}} = Tr\big(\big[\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\big]\boldsymbol{D} + 2\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\big),$$

where $\boldsymbol{D}$ is a matrix differential operator with $D_{i,j} = (\frac{\partial}{\partial x_{i,j}})$, and from Da Fonseca et al. (2008c), Proposition 3.1, we obtain the infinitesimal generator of $(Y_t, \boldsymbol{X}_t)$, which is given by

$$\begin{aligned} L_{Y,\boldsymbol{X}} = {} & \left(r - \frac{1}{2}Tr(\boldsymbol{x})\right)\frac{\partial}{\partial y} + \frac{1}{2}Tr(\boldsymbol{x})\frac{\partial^2}{\partial y^2} \\ & + Tr\big(\big(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\big)\boldsymbol{D} + 2\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\big) \\ & + 2Tr(\boldsymbol{x}\boldsymbol{R}\boldsymbol{Q}\boldsymbol{D})\frac{\partial}{\partial y}. \end{aligned} \tag{11.5.38}$$

Using the Feynman-Kac argument, we have

$$\frac{\partial \Psi_{\gamma,t}}{\partial \tau} = L_{Y,\boldsymbol{X}}\Psi_{\gamma,t}$$

and

$$\Psi_{\gamma,t}(0) = \exp\{\gamma Y_t\}.$$

Using Eq. (11.5.38), we obtain that

$$\begin{aligned} \frac{\partial \Psi_{\gamma,t}}{\partial \tau} = {} & \left(r - \frac{1}{2}Tr(\boldsymbol{x})\right)\frac{\partial \Psi_{\gamma,t}}{\partial y} + \frac{1}{2}Tr(\boldsymbol{x})\frac{\partial^2 \Psi_{\gamma,t}}{\partial y^2} \\ & + Tr\big(\big(\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\big)\boldsymbol{D}\Psi_{\gamma,t} \\ & + 2\big(\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\big)\Psi_{\gamma,t}\big) \\ & + 2Tr(\boldsymbol{x}\boldsymbol{R}\boldsymbol{a}\boldsymbol{D})\frac{\partial \Psi_{\gamma,t}}{\partial \tau}, \end{aligned}$$

subject to $\boldsymbol{A}(0) = \boldsymbol{0}$, $b(0) = \gamma$, and $c(0) = 0$. From Eq. (11.5.37), we obtain that

$$\frac{\partial \Psi_{\gamma,t}}{\partial \tau} = Tr\left(\frac{d}{d\tau}\boldsymbol{A}(\tau)\boldsymbol{x}\right) + \frac{d}{d\tau}b(\tau)y + \frac{d}{d\tau}c(\tau).$$

Identifying the coefficients of y, we obtain

$$\frac{d}{d\tau}b(\tau) = 0,$$

hence $b(\tau) = \gamma$, for $\tau \geq 0$. The remaining part of the argument is identical to the linearization procedure employed in Sect. 11.4. We obtain the following matrix Riccati ODE satisfied by $\boldsymbol{A}(\tau)$,

$$\frac{d}{d\tau}\boldsymbol{A}(\tau) = \boldsymbol{A}(\tau)\boldsymbol{b} + \big(\boldsymbol{b}^\top + 2\gamma \boldsymbol{R}\boldsymbol{a}\big)\boldsymbol{A}(\tau) + 2\boldsymbol{A}(\tau)\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{A}(\tau) + \frac{\gamma(\gamma-1)}{2}\boldsymbol{I}_d,$$

subject to the condition $\boldsymbol{A}(0) = \boldsymbol{0}$. Again, we compute $c(\tau)$ by direct integration,

$$\frac{d}{d\tau}c(\tau) = Tr\big(\alpha \boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{A}(\tau)\big) + \gamma r,$$

subject to $c(0) = 0$. As in Sect. 11.4, we double the dimension of the problem, by setting

$$\boldsymbol{A}(\tau) = \mathbf{F}^{-1}(\tau)\mathbf{G}(\tau),$$

where $\mathbf{F}(\tau) \in GL(d)$, $\mathbf{G}(\tau) \in \mathcal{M}_d$. Hence we conclude that

$$\begin{bmatrix}\mathbf{G}(\tau) & \mathbf{F}(\tau)\end{bmatrix} = \begin{bmatrix}\big(\mathbf{A}(0)\mathbf{A}_{11}(\tau) + \mathbf{A}_{21}(\tau)\big) & \big(\mathbf{A}(0)\mathbf{A}_{12}(\tau) + \mathbf{A}_{22}(\tau)\big)\end{bmatrix},$$

where

$$\begin{bmatrix}\mathbf{A}_{11}(\tau) & \mathbf{A}_{12}(\tau) \\ \mathbf{A}_{21}(\tau) & \mathbf{A}_{22}(\tau)\end{bmatrix} := \exp\left\{\tau\begin{pmatrix}\boldsymbol{b} & -2\boldsymbol{a}^\top\boldsymbol{a} \\ \frac{\gamma(\gamma-1)}{2}\boldsymbol{I}_d & -(\boldsymbol{b}^\top + 2\gamma\boldsymbol{R}\boldsymbol{a})\end{pmatrix}\right\}.$$

We conclude that

$$\begin{aligned}\mathbf{A}(\tau) &= \big(\mathbf{A}(0)\mathbf{A}_{12}(\tau) + \mathbf{A}_{22}(\tau)\big)^{-1}\big(\mathbf{A}(0)\mathbf{A}_{11}(\tau) + \mathbf{A}_{21}(\tau)\big) \\ &= \big(\mathbf{A}_{22}(\tau)\big)^{-1}\mathbf{A}_{21}(\tau),\end{aligned}$$

since $\mathbf{A}(0) = \mathbf{0}$. Lastly, we conclude that

$$c(\tau) = -\frac{\alpha}{2}Tr\big(\log\big(\mathbf{F}(\tau)\big) + \big(\boldsymbol{b}^\top + 2\gamma\boldsymbol{R}\boldsymbol{a}\big)\tau\big) + \gamma r\tau,$$

which, as in Sect. 11.4, avoids a numerical integration to compute $c(\tau)$.

11.5.2 A Heston Multi-asset Multifactor Volatility Model

We now discuss Wishart processes in a multi-asset framework. The model presented in this subsection first appeared in Da Fonseca et al. (2007) and extends the models presented in Sect. 6.7. Under an assumed risk-neutral measure, we use the following model for the vector of risky assets,

$$d\boldsymbol{S}_t = Diag(\boldsymbol{S}_t)(r\mathbf{1}\,dt + \sqrt{\boldsymbol{X}_t}\,d\boldsymbol{Z}_t), \tag{11.5.39}$$

where $\mathbf{1} = (1, \ldots, 1)^\top$, and $\boldsymbol{Z} = \{\boldsymbol{Z}_t,\ t \geq 0\} \in \Re^d$ is a vector-valued Brownian motion. The process $\boldsymbol{X} = \{\boldsymbol{X}_t,\ t \geq 0\}$ is a $WIS_d(\boldsymbol{x}, \alpha, \boldsymbol{b}, \boldsymbol{a})$ process with dynamics

$$d\boldsymbol{X}_t = \big(\alpha\boldsymbol{a}^\top\boldsymbol{a} + \boldsymbol{b}\boldsymbol{X}_t + \boldsymbol{X}_t\boldsymbol{b}^\top\big)\,dt + \sqrt{\boldsymbol{X}_t}\,d\boldsymbol{W}_t\boldsymbol{a} + \boldsymbol{a}^\top\,d\boldsymbol{W}_t^\top\sqrt{\boldsymbol{X}_t},$$

where $\alpha \geq d-1$, $\boldsymbol{b} \in \mathcal{M}_d$ and $\boldsymbol{a} \in GL(d)$. We now make the following assumptions, cf. Da Fonseca et al. (2007):

Assumption 11.5.3 The following assumptions are in force in this subsection:

1. the continuous-time diffusion model for $\boldsymbol{S}$ is a linear-affine stochastic factor model with respect to the log-returns and variance-covariance factors $X_{\cdot,kl}$;
2. the stochastic covariance matrix is given by the Wishart process $\boldsymbol{X}$;
3. the Brownian motion driving the assets' returns and those driving the instantaneous covariance matrix are linearly correlated.

Now we discuss how the Brownian motions $\boldsymbol{Z} = \{\boldsymbol{Z}_t,\ t \geq 0\}$ and $\boldsymbol{W} = \{\boldsymbol{W}_t,\ t \geq 0\}$ can be correlated in order to satisfy Assumptions 1–3 above.

First we introduce d real-valued matrices $\boldsymbol{R}_k \in \mathcal{M}_d$, $k = 1, \dots, d$, so that

$$dZ_t^k = \sqrt{1 - Tr\big(\boldsymbol{R}_k \boldsymbol{R}_k^\top\big)}\, dB_t^k + Tr\big(\boldsymbol{R}_k\, d\boldsymbol{W}_t^\top\big), \quad k = 1, \dots, d,$$

where the vector Brownian motion $\boldsymbol{B} = (B^1, \dots, B^d)$ is independent of $\boldsymbol{W}$. We point out that for a generic choice of $\boldsymbol{R}_k$ the model in Eq. (11.5.39) need not remain affine. Instead, we show the following result from Da Fonseca et al. (2007), which explains how the Brownian motions can be correlated. For a proof, we refer to Da Fonseca et al. (2007).

Proposition 11.5.4 *Assumptions* 1 *and* 2 *imply that for* $k = 1, \dots, d$, *the correlation matrix* $\boldsymbol{R}_k$ *is given by*

$$\boldsymbol{R}_k = \begin{pmatrix} 0 & 0 & 0 \\ \rho_1 & \cdots & \rho_d \\ 0 & 0 & 0 \end{pmatrix} \leftarrow k\text{-th row}, \tag{11.5.40}$$

where $\rho_i \in [-1, 1]$, $i = 1, \dots, d$ *and* $\boldsymbol{\rho}^\top \boldsymbol{\rho} \le 1$.

Equation (11.5.40) implies that the Brownian motion driving the asset vector has to satisfy

$$d\boldsymbol{Z}_t = \sqrt{1 - \boldsymbol{\rho}^\top \boldsymbol{\rho}}\, d\boldsymbol{B}_t + d\boldsymbol{W}_t\, \boldsymbol{\rho}.$$

In particular, for $d = 2$, this means that

$$dZ_{t,1} = \sqrt{1 - \big(\rho_1^2 + \rho_2^2\big)}\, dB_{t,1} + (dW_{t,11}\, \rho_1 + dW_{t,12}\, \rho_2)$$
$$dZ_{t,2} = \sqrt{1 - \big(\rho_1^2 + \rho_2^2\big)}\, dB_{t,2} + (dW_{t,21} \rho_1 + dW_{t,22}\, \rho_2).$$

So all elements of the correlation vector $\boldsymbol{\rho} = (\rho_1, \rho_2)$ feature in both Brownian motions, Z_1 and Z_2.

We now turn to derivative pricing. Recall from Lemma 11.3.5 that the infinitesimal generator of the Wishart process $\boldsymbol{X}$ is given by

$$L_{\boldsymbol{X}} = Tr\big(\big[\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\big]\boldsymbol{D} + 2\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\big),$$

and furthermore, the infinitesimal generator of the asset returns, $\boldsymbol{Y}_t = \log(\boldsymbol{S}_t)$, is given by

$$\begin{aligned} L_{\boldsymbol{Y}} &= \nabla_{\boldsymbol{y}} \left(r\mathbf{1} - \frac{1}{2} Vec(x_{ii}) \right) + \frac{1}{2} \nabla_{\boldsymbol{y}} \boldsymbol{x} \nabla_{\boldsymbol{y}}^\top \\ &= \sum_{i=1}^{d} \left(r - \frac{1}{2} x_{ii} \right) \frac{\partial}{\partial y_i} + \frac{1}{2} \sum_{i,j=1}^{d} x_{ij} \frac{\partial^2}{\partial y_i \partial y_j}, \end{aligned}$$

where $\nabla_{\boldsymbol{y}}$ denotes the gradient operator, $\nabla_{\boldsymbol{y}} = (\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_d})$. Lastly, from Proposition 4 in Da Fonseca et al. (2007), we have

$$L_{Y,X} = Tr\big((\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{bx} + \boldsymbol{xb}^\top)\boldsymbol{D} + 2\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\big)$$
$$+ \nabla_{\boldsymbol{y}}\left(r\mathbf{1} - \frac{1}{2}Vec(x_{ii})\right) + \frac{1}{2}\nabla_{\boldsymbol{y}}\boldsymbol{x}\nabla_{\boldsymbol{y}}^\top$$
$$+ 2Tr\big(\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{\rho}\nabla_{\boldsymbol{y}}\boldsymbol{x}\big),$$

where $\boldsymbol{D}$ is a matrix differential operator with elements

$$D_{i,j} = \left(\frac{\partial}{\partial x_{i,j}}\right),$$

and $Vec(x_{ii})$ is the vector comprised of the elements x_{ii}, $i = 1, \ldots, d$. We now attend to the computation of the affine transform of the log-returns under the assumed risk-neutral measure,

$$\Psi_{\boldsymbol{\gamma},t} = E\big(\exp\{\langle \boldsymbol{\gamma}, \boldsymbol{Y}_{t+\tau}\rangle\}|\mathcal{A}_t\big).$$

As before, we apply the Feynman-Kac argument,

$$\frac{\partial \Psi_{\boldsymbol{\gamma},t}}{\partial \tau} = L_{Y,X}\Psi_{\boldsymbol{\gamma},t}. \tag{11.5.41}$$

We guess that $\Psi_{\boldsymbol{\gamma},t}$ is exponentially affine in $\boldsymbol{X}_t$ and $\boldsymbol{Y}_t$, so we assume that

$$\Psi_{\boldsymbol{\gamma},t} = \exp\big\{Tr\big(\mathbf{A}(\tau)\boldsymbol{X}_t\big) + \boldsymbol{\beta}^\top(\tau)\boldsymbol{Y}_t + c(\tau)\big\}, \tag{11.5.42}$$

where $\mathbf{A}(\tau) \in \mathcal{M}_d$, $\boldsymbol{\beta}(\tau) \in \Re^d$, and $c(\tau) \in \Re$. From Eq. (11.5.41), we compute

$$\frac{\partial \Psi_{\boldsymbol{\gamma},t}}{\partial \tau} = Tr\big((\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{bx} + \boldsymbol{xb}^\top)\boldsymbol{D} + 2\boldsymbol{x}\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{a}\boldsymbol{D}\big)\Psi_{\boldsymbol{\gamma},t}$$
$$+ \nabla_{\boldsymbol{y}}\left(r\mathbf{1} - \frac{1}{2}Vec\big(Tr(\boldsymbol{e}_{ii}\boldsymbol{x})\big)\right)\Psi_{\boldsymbol{\gamma},t}$$
$$+\frac{1}{2}\nabla_{\boldsymbol{y}}\boldsymbol{x}\nabla_{\boldsymbol{y}}^\top \Psi_{\boldsymbol{\gamma},t}$$
$$+ 2Tr\big(\boldsymbol{D}\boldsymbol{a}^\top \boldsymbol{\rho}\nabla_{\boldsymbol{y}}\boldsymbol{x}\big)\Psi_{\boldsymbol{\gamma},t},$$

where $\boldsymbol{e}_{ii} = (\delta_{i,j,k})_{j,k=1,\ldots,d}$ denotes the canonical basis of $\mathcal{M}_d$. Replacing $\frac{\partial \Psi_{\boldsymbol{\gamma},t}}{\partial \tau}$, we get

$$0 = -Tr\left(\frac{d}{d\tau}\boldsymbol{A}(\tau)\boldsymbol{x}\right) - \frac{d}{d\tau}\boldsymbol{\beta}^\top(\tau)\boldsymbol{y} - \frac{d}{d\tau}c(\tau)$$
$$+ \boldsymbol{\beta}^\top(\tau)\left(r\mathbf{1} - \frac{1}{2}Vec\big(Tr(\boldsymbol{e}_{ii}\boldsymbol{x})\big)\right) + \frac{1}{2}\boldsymbol{\beta}^\top(\tau)\boldsymbol{x}\boldsymbol{\beta}(\tau)$$
$$+ Tr\big((\alpha \boldsymbol{a}^\top \boldsymbol{a} + \boldsymbol{bx} + \boldsymbol{xb}^\top)\mathbf{A}(\tau) + 2\boldsymbol{x}\mathbf{A}(\tau)\boldsymbol{a}^\top \boldsymbol{a}\mathbf{A}(\tau)\big)$$
$$+ 2Tr\big(\mathbf{A}(\tau)\boldsymbol{a}^\top \boldsymbol{\rho}\boldsymbol{\beta}^\top(\tau)\boldsymbol{x}\big),$$

that is

$$0 = -Tr\left(\frac{d}{d\tau}\mathbf{A}(\tau)\boldsymbol{x} + \frac{\partial}{\partial\tau}\boldsymbol{\beta}(\tau)\boldsymbol{y}^\top\right) - \frac{\partial}{\partial\tau}c(\tau)$$
$$+ Tr\left(r\mathbf{1}\boldsymbol{\beta}^\top(\tau) - \frac{1}{2}\sum_{i=1}^{d}\boldsymbol{\beta}_i(\tau)\boldsymbol{e}_{ii}\boldsymbol{x} + \frac{1}{2}\boldsymbol{\beta}(\tau)\boldsymbol{\beta}^\top(\tau)\boldsymbol{x}\right)$$
$$+ Tr\big(\big(\alpha\boldsymbol{a}^\top\boldsymbol{a} + \boldsymbol{b}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{b}^\top\big)\mathbf{A}(\tau) + 2\boldsymbol{x}\mathbf{A}(\tau)\boldsymbol{a}^\top\boldsymbol{a}\mathbf{A}(\tau) + 2\mathbf{A}(\tau)\boldsymbol{a}^\top\boldsymbol{\rho}\boldsymbol{\beta}^\top(\tau)\boldsymbol{x}\big),$$

subject to the boundary conditions

$$\mathbf{A}(0) = \mathbf{0}, \qquad \boldsymbol{\beta}(0) = \boldsymbol{\gamma}, \qquad c(0) = 0.$$

Identifying the coefficients of $\boldsymbol{y}$ we deduce

$$\frac{d}{d\tau}\boldsymbol{\beta}(\tau) = \mathbf{0},$$

hence $\boldsymbol{\beta}(\tau) = \boldsymbol{\gamma}$, for $\tau \geq 0$. As in Sect. 11.4, by identifying the coefficients of $\boldsymbol{X}$, we obtain the matrix Riccati ODE satisfied by $\boldsymbol{A}(\tau)$,

$$\frac{d}{d\tau}\boldsymbol{A}(\tau) = \boldsymbol{A}(\tau)\boldsymbol{b} + \boldsymbol{b}^\top\boldsymbol{A}(\tau) - \frac{1}{2}\sum_{i=1}^{d}\boldsymbol{\gamma}_i\boldsymbol{e}_{ii} + 2\boldsymbol{A}(\tau)\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{A}(\tau) + \frac{1}{2}\boldsymbol{\gamma}\boldsymbol{\gamma}^\top$$
$$+ \boldsymbol{A}(\tau)\boldsymbol{a}^\top\boldsymbol{\rho}\boldsymbol{\gamma}^\top + \boldsymbol{\gamma}\boldsymbol{\rho}^\top\boldsymbol{a}\boldsymbol{A}(\tau)$$
$$= \boldsymbol{A}(\tau)\big(\boldsymbol{b} + \boldsymbol{a}^\top\boldsymbol{\rho}\boldsymbol{\gamma}^\top\big) + \big(\boldsymbol{b}^\top + \boldsymbol{\gamma}\boldsymbol{\rho}^\top\boldsymbol{a}\big)\boldsymbol{A}(\tau) + 2\boldsymbol{A}(\tau)\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{A}(\tau)$$
$$- \frac{1}{2}\sum_{i=1}^{d}\boldsymbol{\gamma}_i\boldsymbol{e}_{ii} + \frac{1}{2}\boldsymbol{\gamma}\boldsymbol{\gamma}^\top,$$

subject to $\boldsymbol{A}(\tau) = \mathbf{0}$. Doubling the dimension of the problem, as in Sect. 11.4, we obtain

$$\boldsymbol{A}(\tau) = \big(\boldsymbol{A}(0)\boldsymbol{A}_{12}(\tau) + \boldsymbol{A}_{22}(\tau)\big)^{-1}\big(\boldsymbol{A}(0)\boldsymbol{A}_{11}(\tau) + \boldsymbol{A}_{21}(\tau)\big),$$

where

$$\begin{bmatrix} \boldsymbol{A}_{11}(\tau) & \boldsymbol{A}_{12}(\tau) \\ \boldsymbol{A}_{21}(\tau) & \boldsymbol{A}_{22}(\tau) \end{bmatrix} := \exp\left\{\tau\begin{pmatrix} \boldsymbol{b} + \boldsymbol{a}^\top\boldsymbol{\rho}\boldsymbol{\gamma}^\top & -2\boldsymbol{a}^\top\boldsymbol{a} \\ \frac{1}{2}(\boldsymbol{\gamma}\boldsymbol{\gamma}^\top - \sum_{i=1}^{d}\boldsymbol{\gamma}_i\boldsymbol{e}_{ii}) & -(\boldsymbol{b}^\top + \boldsymbol{\gamma}\boldsymbol{\rho}^\top\boldsymbol{a}) \end{pmatrix}\right\}.$$

For the function $c(\tau)$, we have

$$\frac{d}{d\tau}c(\tau) = Tr\big(r\mathbf{1}\boldsymbol{\gamma}^\top + \alpha\boldsymbol{a}^\top\boldsymbol{a}\boldsymbol{A}(\tau)\big),$$

subject to the initial condition $c(0) = 0$. We can solve the above equation to yield

$$c(\tau) = -\frac{\alpha}{2}Tr\big(\log\big(\boldsymbol{A}_{22}(\tau)\big) + \tau\boldsymbol{b}^\top + \tau\boldsymbol{\gamma}\boldsymbol{\rho}^\top\boldsymbol{a}\big) + \tau r\boldsymbol{\gamma}^\top\mathbf{1}.$$

Consequently, we price derivatives as discussed in Chap. 8.