The Annals of Applied Probability
2013, Vol. 23, No. 3, 1025–1073
DOI: 10.1214/12-AAP863

EXACT AND HIGH-ORDER DISCRETIZATION SCHEMES FOR WISHART PROCESSES AND THEIR AFFINE EXTENSIONS[1]

BY ABDELKODDOUSSE AHDIDA AND AURÉLIEN ALFONSI

Université Paris-Est

This work deals with the simulation of Wishart processes and affine diffusions on positive semidefinite matrices. To do so, we focus on the splitting of the infinitesimal generator in order to use composition techniques as did Ninomiya and Victoir [*Appl. Math. Finance* **15** (2008) 107–121] or Alfonsi [*Math. Comp.* **79** (2010) 209–237]. Doing so, we have found a remarkable splitting for Wishart processes that enables us to sample exactly Wishart distributions without any restriction on the parameters. It is related but extends existing exact simulation methods based on Bartlett's decomposition. Moreover, we can construct high-order discretization schemes for Wishart processes and second-order schemes for general affine diffusions. These schemes are, in practice, faster than the exact simulation to sample entire paths. Numerical results on their convergence are given.

Introduction. This paper focuses on simulation methods for Wishart processes and more generally for affine diffusions on positive semidefinite matrices. Before explaining our motivations and our main results, we start with a short introduction to these processes. Even though we use rather standard notation for matrices, they are recalled at the end of the Introduction, and we invite the reader to first give a quick look at it. Wishart processes have been initially introduced by Bru [4, 5]. They are also named because their marginal laws follow Wishart distributions. Very recently, Cuchiero et al. [7] have introduced a general framework for affine processes on positive semidefinite matrices $\mathcal{S}_d^+(\mathbb{R})$ that embeds Wishart processes and includes possible jumps. In this paper, we only consider continuous processes of this kind. Such processes solve the following SDE:

$$X_t^x = x + \int_0^t (\bar{\alpha} + B(X_s^x))\, ds + \int_0^t \left(\sqrt{X_s^x}\, dW_s a + a^T\, dW_s^T \sqrt{X_s^x}\right). \tag{1}$$

Here, and throughout the paper, $(W_t, t \geq 0)$ denotes a d-by-d square matrix made of independent standard Brownian motions and

$$x, \bar{\alpha} \in \mathcal{S}_d^+(\mathbb{R}), \qquad a \in \mathcal{M}_d(\mathbb{R}) \quad \text{and} \quad B \in \mathcal{L}(\mathcal{S}_d(\mathbb{R})) \tag{2}$$

Received April 2011; revised March 2012.

[1]Supported by the Credinext project from Finance Innovation, of the Eurostars E!5144-TFM project, and of the "Chaire Risques Financiers" of Fondation du Risque.

MSC2010 subject classifications. 65C30, 60H35, 91B70.

Key words and phrases. Wishart processes, affine processes, exact simulation, discretization schemes, weak error, Bartlett's decomposition.

is a linear mapping on $\mathcal{S}_d(\mathbb{R})$. Wishart processes correspond to the case where

$$\exists\alpha \ge 0, \qquad \bar{\alpha} = \alpha a^T a \quad \text{and} \qquad (3)$$
$$\exists b \in \mathcal{M}_d(\mathbb{R}), \forall x \in \mathcal{S}_d(\mathbb{R}) \qquad B(x) = bx + xb^T.$$

When $d = 1$, (1) is simply the SDE of the Cox–Ingersoll–Ross (CIR) process that has been broadly studied, and we will implicitly assume that $d \ge 2$ throughout the paper. Weak and strong uniqueness of SDE (1) has been studied by Bru [5], Cuchiero et al. [7] and Mayerhofer, Pfaffel and Stelzer [22]. Here we sum up their results.

THEOREM 1. *If $x \in \mathcal{S}_d^+(\mathbb{R})$, $\bar{\alpha} - (d-1)a^T a \in \mathcal{S}_d^+(\mathbb{R})$ and B satisfies the following condition*:

$$(4) \qquad \forall x_1, x_2 \in \mathcal{S}_d^+(\mathbb{R}) \qquad \operatorname{Tr}(x_1 x_2) = 0 \implies \operatorname{Tr}(B(x_1)x_2) \ge 0,$$

there is a unique weak solution to the SDE (1) *in $\mathcal{S}_d^+(\mathbb{R})$. We denote by $\mathrm{AFF}_d(x, \bar{\alpha}, B, a)$ the law of $(X_t^x)_{t\ge 0}$ and $\mathrm{AFF}_d(x, \bar{\alpha}, B, a; t)$ the marginal law of X_t^x. If we assume, moreover, that $\bar{\alpha} - (d+1)a^T a \in \mathcal{S}_d^+(\mathbb{R})$ and $x \in \mathcal{S}_d^{+,*}(\mathbb{R})$, there is a unique strong solution to the SDE* (1).

Under the parametrization of Wishart processes (3), *condition* (4) *is satisfied and weak uniqueness holds as soon as $\alpha \ge d-1$. In that case, we denote by $\mathrm{WIS}_d(x, \alpha, b, a)$ the law of the Wishart process $(X_t^x)_{t\ge 0}$ and $\mathrm{WIS}_d(x, \alpha, b, a; t)$ the law of X_t^x.*

Throughout the paper, when we use the notation $\mathrm{AFF}_d(x, \bar{\alpha}, B, a)$ or $\mathrm{AFF}_d(x, \bar{\alpha}, B, a; t)$ [resp., $\mathrm{WIS}_d(x, \alpha, b, a)$ or $\mathrm{WIS}_d(x, \alpha, b, a; t)$], we implicitly assume that $\bar{\alpha} - (d-1)a^T a \in \mathcal{S}_d^+(\mathbb{R})$ (resp., $\alpha \ge d-1$) and B satisfies (4) so that weak uniqueness holds.

In her Ph.D. thesis [4], Bru introduced Wishart processes and used them in biology to study perturbed experimental data. Recently, great attention has been paid to Wishart processes for applications in finance. Namely, Gourieroux and Sufana [14] and Da Fonseca, Grasselli and Tebaldi [8] have suggested the use of these processes to model the instantaneous covariance matrix of d assets. It naturally extends stochastic volatility models for only one asset like the Heston model [16]. Obviously, processes on positive semidefinite matrices are really interesting to model the evolution of a dependence structure because they can describe a covariance matrix. However, when dealing with applications, it is, in general, crucial to be able to sample paths of such processes and make Monte Carlo algorithms.

To the best of our knowledge, there is minimal literature on simulation methods for Wishart and general affine processes (1). Wishart distributions have been intensively studied in statistics when $\alpha \in \mathbb{N}$. In this case, exact simulation methods have been proposed by Odell and Feiveson [25], Smith and Hocking [26] and

Gleser [12], to mention a few. Concerning discretization schemes, the usual Euler–Maruyama scheme is not well defined because of the square-root. This already happens for the CIR process ($d = 1$). One has then to find specific schemes. Recently, Benabid et al. [3] and Gauthier and Possamai [10] have proposed numerical approximations for Wishart processes that are well defined under some restrictions on the parameters. However, there is no result on the accuracy of their methods. Currently, Teichmann [29] is working on dedicated schemes for general affine processes by approximating their characteristic functions. Our study here is only dedicated to the diffusion (1).

Initially, our goal was to find high-order discretization schemes for Wishart processes by splitting operators and using scheme compositions. Indeed, this approach has already proved to be very efficient for other affine diffusions (see [2]). The main difficulty here was to find a splitting that involves infinitesimal generators of diffusions that are well defined on $\mathcal{S}_d^+(\mathbb{R})$ and that can be simulated. Doing so, we incidentally have found a remarkable splitting for some canonical Wishart processes: the infinitesimal generator of $\mathrm{WIS}_d(x, \alpha, 0, I_d^n)$ is the sum of commuting operators that are associated to elementary SDEs that can be sampled exactly. With the help of a simple but useful law identity, this enables us to sample exactly Wishart processes for any admissible parameter. In particular, our result extends the Bartlett's decomposition that is commonly used to sample central Wishart distributions. This splitting is not only interesting for the exact simulation method. It is also useful to construct high-order discretization schemes for Wishart processes that are, in practice, faster to generate full paths. In fact, it allows us to get a high-order scheme that preserves the domain $\mathcal{S}_d^+(\mathbb{R})$. We provide a rigorous analysis of the weak error in this framework. Still, by using the splitting technique, we also get a second-order scheme for any affine diffusion (1) without any restriction on the parameters.

This paper is structured as follows. First, we present some general results on affine diffusions. We calculate their infinitesimal generator and obtain interesting identities in law that are intensively used next for the different simulation methods. Section 2 is devoted to the exact simulation of Wishart processes. It exhibits the remarkable splitting of the infinitesimal generator and shows how it can be used to sample exactly any Wishart distribution. Section 3 deals with high-order schemes for affine diffusions. Thanks to the remarkable splitting, we are able to construct a third-order scheme for Wishart processes and second-order schemes for affine diffusions. Last, we give numerical illustrations of our convergence results in Section 4. We compare the time required by each method and also give a possible application of our results in finance.

Notation for real matrices.

- For $d, d' \in \mathbb{N}^*$, $\mathcal{M}_d(\mathbb{R})$ denotes the real d square matrices and $\mathcal{M}_{d\times d'}(\mathbb{R})$ the real matrices with d rows and d' columns.

- $\mathcal{S}_d(\mathbb{R})$, $\mathcal{S}_d^+(\mathbb{R})$, $\mathcal{S}_d^{+,*}(\mathbb{R})$ and $\mathcal{G}_d(\mathbb{R})$ denote, respectively, the set of symmetric, symmetric positive semidefinite, symmetric positive definite and nonsingular matrices.
- For $x \in \mathcal{M}_d(\mathbb{R})$, x^T, $\mathrm{adj}(x)$, $\det(x)$, $\mathrm{Tr}(x)$ and $\mathrm{Rk}(x)$ are, respectively, the transpose, the adjugate, the determinant, the trace and the rank of x.
- For $x \in \mathcal{S}_d^+(\mathbb{R})$, $\sqrt{x}$ denotes the unique symmetric positive semidefinite matrix such that $(\sqrt{x})^2 = x$.
- The identity matrix is denoted by I_d and we set for $n \le d$, $I_d^n = (\mathbb{1}_{i=j\le n})_{1\le i,j\le d}$ and $e_d^n = (\mathbb{1}_{i=j=n})_{1\le i,j\le d}$, so that $I_d^n = \sum_{i=1}^n e_d^i$. We also set for $1 \le i, j \le d$, $e_d^{i,j} = (\mathbb{1}_{k=i,l=j})_{1\le k,l\le d}$.
- For $x \in \mathcal{S}_d(\mathbb{R})$, we denote by $x_{\{i,j\}}$ the value of $x_{i,j}$, so that

$$x = \sum_{1\le i\le j\le d} x_{\{i,j\}}(e_d^{i,j} + \mathbb{1}_{i\ne j} e_d^{j,i}).$$

  We use both notation in the paper: notation $(x_{i,j})_{1\le i,j\le d}$ is more convenient for matrix calculations while $(x_{\{i,j\}})_{1\le i\le j\le d}$ is preferred to emphasize that we work on symmetric matrices.
- For $\lambda_1, \dots, \lambda_d \in \mathbb{R}$, $\mathrm{diag}(\lambda_1, \dots, \lambda_d)$ denotes the diagonal matrix such that $\mathrm{diag}(\lambda_1, \dots, \lambda_d)_{i,i} = \lambda_i$.

1. Some properties of affine processes on positive semidefinite matrices.

1.1. *The infinitesimal generator on $\mathcal{M}_d(\mathbb{R})$ and $\mathcal{S}_d(\mathbb{R})$.* We start with a simple lemma. It is useful to calculate the infinitesimal generator of processes on matrices.

LEMMA 2. *Let $(\mathcal{F}_t)_{t\ge 0}$ denote the filtration generated by $(W_t, t \ge 0)$. We consider continuous $(\mathcal{F}_t)$-adapted processes $(A_t)_{t\ge0}$, $(B_t)_{t\ge0}$ and $(C_t)_{t\ge0}$, respectively, valued in $\mathcal{M}_d(\mathbb{R})$, $\mathcal{M}_d(\mathbb{R})$ and $\mathcal{S}_d(\mathbb{R})$, and a process $(Y_t)_{t\ge0}$ that admits the following semimartingale decomposition*:

$$dY_t = C_t\,dt + B_t\,dW_t A_t + A_t^T\,dW_t^T B_t^T. \tag{5}$$

Then, for $i, j, m, n \in \{1, \dots, d\}$, the quadratic covariation of $(Y_t)_{i,j}$ and $(Y_t)_{m,n}$ is

$$\begin{aligned} &d\langle (Y_t)_{i,j}, (Y_t)_{m,n}\rangle \\ &\quad= [(B_tB_t^T)_{i,m}(A_t^TA_t)_{j,n} + (B_tB_t^T)_{i,n}(A_t^TA_t)_{j,m} \\ &\qquad + (B_tB_t^T)_{j,m}(A_t^TA_t)_{i,n} + (B_tB_t^T)_{j,n}(A_t^TA_t)_{i,m}]\,dt. \end{aligned} \tag{6}$$

It is worth noticing that the quadratic covariation given by (5) depends on A_t and B_t only through the matrices $A_t^TA_t$ and $B_tB_t^T$. Lemma 2 enables us to easily calculate the infinitesimal generator for the affine process (1) which is defined by

$$x \in \mathcal{S}_d^+(\mathbb{R}), \qquad L^{\mathcal{M}} f(x) = \lim_{t\to 0^+} \frac{\mathbb{E}[f(X_t^x)] - f(x)}{t}$$

for $f \in \mathcal{C}^2(\mathcal{M}_d(\mathbb{R}), \mathbb{R})$ with bounded derivatives.

In fact, we get that the generator of $\mathrm{AFF}_d(x,\bar{\alpha},B,a)$ is given by

$$L^{\mathcal{M}} = \mathrm{Tr}([\bar{\alpha}+B(x)]D^{\mathcal{M}}) + \tfrac{1}{2}\{2\,\mathrm{Tr}(xD^{\mathcal{M}}a^TaD^{\mathcal{M}}) + \mathrm{Tr}(x(D^{\mathcal{M}})^Ta^TaD^{\mathcal{M}}) + \mathrm{Tr}(xD^{\mathcal{M}}a^Ta(D^{\mathcal{M}})^T)\}, \tag{7}$$

where $D^{\mathcal{M}} = (\partial_{i,j})_{1\le i,j\le d}$. Since we know that the affine process $(X_t^x)_{t\ge 0}$ takes values in $\mathcal{S}_d^+(\mathbb{R}) \subset \mathcal{S}_d(\mathbb{R})$, we can also look at the infinitesimal generator of this diffusion on $\mathcal{S}_d(\mathbb{R})$, which is defined by

$$x \in \mathcal{S}_d^+(\mathbb{R}), \qquad L^{\mathcal{S}} f(x) = \lim_{t\to 0^+} \frac{\mathbb{E}[f(X_t^x)] - f(x)}{t}$$

for $f \in \mathcal{C}^2(\mathcal{S}_d(\mathbb{R}),\mathbb{R})$ with bounded derivatives.

For $x \in \mathcal{S}_d(\mathbb{R})$, we denote by $x_{\{i,j\}} = x_{i,j} = x_{j,i}$ the value of the coordinates (i,j) and (j,i), so that $x = \sum_{1\le i\le j\le d} x_{\{i,j\}}(e_d^{i,j} + \mathbb{1}_{i\neq j}e_d^{j,i})$. For $f \in \mathcal{C}^2(\mathcal{S}_d(\mathbb{R}),\mathbb{R})$, we then denote by $\partial_{\{i,j\}}f$ its derivative with respect to $x_{\{i,j\}}$. For $x \in \mathcal{M}_d(\mathbb{R})$, we set $\pi(x) = (x+x^T)/2$. It is such that $\pi(x) = x$ for $x \in \mathcal{S}_d(\mathbb{R})$, and we have

$$L^{\mathcal{S}} f(x) = L^{\mathcal{M}} f \circ \pi(x).$$

By the chain rule, we have for $x \in \mathcal{S}_d(\mathbb{R})$, $\partial_{i,j} f\circ\pi(x) = (\mathbb{1}_{i=j} + \frac{1}{2}\mathbb{1}_{i\neq j})\partial_{\{i,j\}}f(x)$ and get from (7) the following result.

PROPOSITION 3. *The infinitesimal generator on $\mathcal{S}_d(\mathbb{R})$ associated to* $\mathrm{AFF}_d(x,\bar{\alpha},B,a)$ *is given by*

$$L^{\mathcal{S}} = \mathrm{Tr}\big([\bar{\alpha}+B(x)]D^{\mathcal{S}}\big) + 2\,\mathrm{Tr}(xD^{\mathcal{S}}a^TaD^{\mathcal{S}}), \tag{8}$$

where $D^{\mathcal{S}}$ is defined by $D^{\mathcal{S}}_{i,j} = (\mathbb{1}_{i=j} + \frac{1}{2}\mathbb{1}_{i\neq j})\partial_{\{i,j\}}$, for $1 \le i,j \le d$.

Of course, the generators $L^{\mathcal{M}}$ and $L^{\mathcal{S}}$ are equivalent; one can be deduced from the other. However, $L^{\mathcal{S}}$ already embeds the fact that the process lies in $\mathcal{S}_d(\mathbb{R})$, which reduces the dimension from d^2 to $d(d+1)/2$ and gives, in practice, shorter formulas. This is why we will mostly work in the sequel with infinitesimal generators on $\mathcal{S}_d(\mathbb{R})$. Unless it is necessary to make the distinction with $L^{\mathcal{M}}$, we will simply denote $L = L^{\mathcal{S}}$.

1.2. *The characteristic function of Wishart processes.* As for other affine processes, the characteristic function of affine processes on positive semidefinite matrices can be obtained by solving two ODEs. In the case of Wishart processes, it is possible to solve explicitly these ODEs by solving a matrix Riccati equation (see Levin [20]). Here, we give the closed formula for the Laplace transform and a precise description of its set of convergence.

PROPOSITION 4. *Let* $X_t^x \sim \mathrm{WIS}_d(x, \alpha, b, a; t)$, $q_t = \int_0^t \exp(sb) a^T a \exp(s \times b^T)\, ds$ *and* $m_t = \exp(tb)$. *We introduce the set of convergence of the Laplace transform of* X_t^x, $\mathcal{D}_{b,a;t} = \{v \in \mathcal{S}_d(\mathbb{R}), \mathbb{E}[\exp(\mathrm{Tr}(vX_t^x))] < \infty\}$. *This is a convex open set that is given explicitly by*

$$\mathcal{D}_{b,a;t} = \{v \in \mathcal{S}_d(\mathbb{R}), \forall s \in [0,t], I_d - 2q_s v \in \mathcal{G}_d(\mathbb{R})\}. \tag{9}$$

Besides, the Laplace transform of X_t^x *is well defined for* $v = v_R + i v_I$ *with* $v_R \in \mathcal{D}_{b,a;t}$, $v_I \in \mathcal{S}_d(\mathbb{R})$ *and is given by*

$$\mathbb{E}[\exp(\mathrm{Tr}(vX_t^x))] = \frac{\exp(\mathrm{Tr}[v(I_d - 2q_t v)^{-1} m_t x m_t^T])}{\det(I_d - 2q_t v)^{\alpha/2}}. \tag{10}$$

The characteristic function corresponds to the case $v_R = 0$ that clearly belongs to $\mathcal{D}_{b,a;t}$. The proof of this result is given in Appendix B.1. The formula (10) is well known in the literature, and our contribution is to characterize precisely the set of convergence. In particular, let us observe that $\rho I_d \in \mathcal{D}_{b,a;t}$ when $\rho > 0$ is small enough, which will help us to study the Cauchy problem (Proposition 14).

Last, let us remark here that for $\tilde{X}_t^x \sim \mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$, the formula above becomes even simpler and we have for $v = v_R + i v_I$ such that $v_R \in \mathcal{D}_{b,a;t}$, $v_I \in \mathcal{S}_d(\mathbb{R})$,

$$\mathbb{E}[\exp(\mathrm{Tr}(v\tilde{X}_t^x))] = \frac{\exp(\mathrm{Tr}[v(I_d - 2tI_d^n v)^{-1} x])}{\det(I_d - 2tI_d^n v)^{\alpha/2}}. \tag{11}$$

1.3. *Some identities in law for affine processes.* This section gives simple but interesting identities in law for affine processes. First, we observe that their infinitesimal generator (8) only depends on a through $a^T a$ and get

$$\mathrm{AFF}_d(x, \bar{\alpha}, B, a) \underset{\mathrm{Law}}{=} \mathrm{AFF}_d\big(x, \bar{\alpha}, B, \sqrt{a^T a}\big). \tag{12}$$

Also, it is natural to look at linear transformations of affine processes. Let $q \in \mathcal{G}_d(\mathbb{R})$ and define $B_q \in \mathcal{L}(\mathcal{S}_d(\mathbb{R}))$ by $B_q(x) = (q^T)^{-1} B(q^T x q) q^{-1}$. One easily has that B satisfies (4) iff B_q satisfies (4), and we get

$$\mathrm{AFF}_d(x, \bar{\alpha}, B, a) \underset{\mathrm{Law}}{=} q^T \,\mathrm{AFF}_d((q^{-1})^T x q^{-1}, (q^{-1})^T \bar{\alpha} q^{-1}, B_q, a q^{-1}) q, \tag{13}$$

since both processes solve the same martingale problem. An interesting consequence is given in the following proposition: any affine process can be obtained as a linear transformation of an affine process for which $\bar{\alpha}$ is a diagonal matrix and $a = I_d^n$. Since our main goal here is to sample paths of such processes, this says to us that it is sufficient to focus on this special case.

PROPOSITION 5. *Let* $n = \mathrm{Rk}(a)$ *be the rank of* $a^T a$. *Then, there exist a diagonal matrix* $\bar{\delta}$ *and a nonsingular matrix* $u \in \mathcal{G}_d(\mathbb{R})$ *such that* $\bar{\alpha} = u^T \bar{\delta} u$ *and* $a^T a = u^T I_d^n u$ *and we have*

$$\mathrm{AFF}_d(x, \bar{\alpha}, B, a) \underset{Law}{=} u^T \,\mathrm{AFF}_d((u^{-1})^T x u^{-1}, \bar{\delta}, B_u, I_d^n) u,$$

where $\forall y \in \mathcal{S}_d(\mathbb{R})$, $B_u(y) = (u^{-1})^T B(u^T y u) u^{-1}$.

The proof of this result consists of algebraic arguments and is found in Appendix B.2. It gives, in particular, a general way to compute u and $\bar{\delta}$. Let us notice, however, that in the case of Wishart processes, u can directly be obtained by using a single extended Cholesky decomposition (Lemma 23).

Up to now, we have stated identities for the law of affine processes. Thanks to the explicit characteristic function of Wishart processes, we are also able to get another interesting identity on the marginal laws.

PROPOSITION 6. *Let* $t > 0$, $a, b \in \mathcal{M}_d(\mathbb{R})$ *and* $\alpha \geq d-1$. *Let* $m_t = \exp(tb)$, $q_t = \int_0^t \exp(sb) a^T a \exp(sb^T)\,ds$ *and* $n = \mathrm{Rk}(q_t)$. *Then there is* $\theta_t \in \mathcal{G}_d(\mathbb{R})$ *such that* $q_t = t\theta_t I_d^n \theta_t^T$, *and we have*

$$\mathrm{WIS}_d(x, \alpha, b, a; t) \underset{Law}{=} \theta_t \,\mathrm{WIS}_d(\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T, \alpha, 0, I_d^n; t) \theta_t^T. \tag{14}$$

This proposition plays a crucial role for the exact simulation of Wishart processes. Thanks to (14), we can sample any Wishart distribution if we are able to simulate exactly the distribution $\mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$ for any $x \in \mathcal{S}_d^+(\mathbb{R})$. In Section 2, we focus on this and give a way to sample exactly $\mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$. Let us stress here that we can compute the matrix θ_t by using the extended Cholesky decomposition of q_t/t, as it is explained in the proof below.

PROOF OF PROPOSITION 6. We apply Lemma 23 to $q_t/t \in \mathcal{S}_d^+(\mathbb{R})$ and consider (p, c_n, k_n) an extended Cholesky decomposition of q_t/t. We set $\theta_t = p^{-1}\left(\begin{smallmatrix} c_n & 0 \\ k_n & I_{d-n} \end{smallmatrix}\right)$. Then θ_t is invertible and it is easy to check that $q_t = t\theta_t I_d^n \theta_t^T$. Now, let us observe that for $v \in \mathcal{S}_d(\mathbb{R})$,

$$\begin{aligned}
&\det(I_d - 2iq_t v) = \det\bigl(\theta_t(\theta_t^{-1} - 2itI_d^n\theta_t^T v)\bigr) = \det(I_d - 2itI_d^n \theta_t^T v \theta_t),\\
&\quad \mathrm{Tr}[iv(I_d - 2iq_t v)^{-1} m_t x m_t^T]\\
&\qquad = \mathrm{Tr}[i(\theta_t^{-1})^T \theta_t^T v(\theta_t\theta_t^{-1} - 2it\theta_t I_d^n \theta_t^T v \theta_t \theta_t^{-1})^{-1} m_t x m_t^T]\\
&\qquad = \mathrm{Tr}[i\theta_t^T v \theta_t (I_d - 2itI_d^n \theta_t^T v\theta_t)^{-1} \theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T].
\end{aligned}$$

Let $X_t^x \sim \mathrm{WIS}_d(x, \alpha, b, a; t)$ and $\tilde{X}_t^x \sim \mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$. Then, from (10) and (11), we get that

$$\begin{aligned}
\mathbb{E}[\exp(i\,\mathrm{Tr}(vX_t^x))] &= \mathbb{E}\bigl[\exp\bigl(i\,\mathrm{Tr}\bigl(\theta_t^T v \theta_t \tilde{X}_t^{\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T}\bigr)\bigr)\bigr]\\
&= \mathbb{E}\bigl[\exp\bigl(i\,\mathrm{Tr}\bigl(v\theta_t \tilde{X}_t^{\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T} \theta_t^T\bigr)\bigr)\bigr]. \qquad \square
\end{aligned}$$

Last, let us mention that (14) extends a usual identity between CIR and squared Bessel distribution. It gives when $d = 1$,

$$\mathrm{WIS}_1(x, \alpha, b, a; t) \underset{\text{Law}}{=} a^2 \frac{e^{2bt} - 1}{2bt} \mathrm{WIS}_1\left(\frac{2btx}{a^2(1 - e^{-2bt})}, \alpha, 0, 1; t\right).$$

In that case, this identity can also be obtained directly from the SDE. Let $(X_t^x)_{t\geq 0} \sim \mathrm{WIS}_1(x,\alpha,b,a)$. Then, $Y_t = e^{-2bt}X_t^x/a^2$ is a time-changed Bessel squared process since $dY_t = \alpha(e^{-2bt}\,dt) + 2\sqrt{Y_t}(e^{-bt}\,dW_t)$. We obtain $\mathrm{WIS}_1(x,\alpha, b,a;t) \underset{\mathrm{Law}}{=} a^2e^{2bt}\,\mathrm{WIS}_1(x/a^2,\alpha,0,1;\frac{1-e^{-2bt}}{2b})$. A linear time-change also gives that $\mathrm{WIS}_1(x,\alpha,0,1;\lambda t) \underset{\mathrm{Law}}{=} \lambda\,\mathrm{WIS}_1(x/\lambda,\alpha,0,1;t)$, which leads to (14) by taking $\lambda = (1-e^{-2bt})/(2bt)$.

2. Exact simulation of Wishart processes. In this section, we present a new method to simulate exactly a Wishart process. To the best of our knowledge, this is the first exact simulation method for noncentral Wishart distributions that works for any $\alpha \geq d-1$. Wishart distributions have been thoroughly studied in statistics when $\alpha \in \mathbb{N}$ (which is then called the number of degrees of freedom). Exact simulation methods have already been proposed in that case. For instance, Odell and Feiveson [25] and Smith and Hocking [26] have proposed an exact simulation method for central Wishart distributions based on the Bartlett's decomposition. Gleser [12] extends it to any (noncentral) Wishart distribution. Bru [5] also explains, when $\alpha \in \mathbb{N}$, how Wishart processes can be obtained as a square of Ornstein–Uhlenbeck processes on matrices.

Here, our method relies on the identity in law (14) that enables us to focus on the case $b=0$, $a=I_d^n$. Then we show a remarkable splitting of the infinitesimal generator as the sum of commuting operators. These operators are associated to SDE that can be solved explicitly on $\mathcal{S}_d^+(\mathbb{R})$, which enables us to sample any Wishart distribution.

2.1. *A remarkable splitting for* $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$. The following theorem explains how to split the infinitesimal generator of $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$ as the sum of commutative infinitesimal generators. This result is the keystone of the paper and will play a crucial role in the sequel both for the exact and discretization schemes.

THEOREM 7. *Let L be the generator associated to the Wishart process $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$ and $L_{e_d^i}$ be the generator associated to $\mathrm{WIS}_d(x,\alpha,0,e_d^i)$ for $i\in\{1,\dots,d\}$. Then, we have*

$$L = \sum_{i=1}^n L_{e_d^i} \quad \text{and} \quad \forall i,j \in \{1,\dots,d\} \qquad L_{e_d^i}L_{e_d^j} = L_{e_d^j}L_{e_d^i}. \tag{15}$$

PROOF. From (8), we easily get that $L=\sum_{i=1}^n L_{e_d^i}$ since $I_d^n = \sum_{i=1}^n e_d^i$. The commutativity property comes from a tedious but simple calculation. □

Beyond the commutativity property, two other features of (15) are important to notice:

- The operators $L_{e_d^i}$ and $L_{e_d^j}$ are the same up to the exchange of coordinates i and j.
- The processes $\mathrm{WIS}_d(x,\alpha,0,e_d^i)$ and $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$ are well defined on $\mathcal{S}_d^+(\mathbb{R})$ under the same hypothesis, namely, $\alpha \ge d-1$ and $x \in \mathcal{S}_d^+(\mathbb{R})$.

This second property makes possible the composition that we explain now. Let us consider $t>0$ and $x \in \mathcal{S}_d^+(\mathbb{R})$. We define, iteratively,

$$\begin{aligned}
X_t^{1,x} &\sim \mathrm{WIS}_d(x,\alpha,0,e_d^1;t),\\
X_t^{2,X_t^{1,x}} &\sim \mathrm{WIS}_d(X_t^{1,x},\alpha,0,e_d^2;t),\\
&\vdots\\
X_t^{n,\dots X_t^{1,x}} &\sim \mathrm{WIS}_d\bigl(X_t^{n-1,\dots X_t^{1,x}},\alpha,0,e_d^n;t\bigr).
\end{aligned}$$

Thus, conditionally to $X_t^{i-1,\dots X_t^{1,x}}$, $X_t^{i,\dots X_t^{1,x}}$ is sampled 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 $(\alpha,0,e_d^i)$. We have the following result.

PROPOSITION 8. *Let $X_t^{n,\dots X_t^{1,x}}$ be defined as above. Then*

$$X_t^{n,\dots X_t^{1,x}} \sim \mathrm{WIS}_d(x,\alpha,0,I_d^n;t).$$

Thanks to this proposition, we can generate a sample according to $\mathrm{WIS}_d(x,\alpha,0,I_d^n;t)$ as soon as we can simulate $\mathrm{WIS}_d(x,\alpha,0,e_d^i;t)$. These laws are the same as $\mathrm{WIS}_d(x,\alpha,0,e_d^1;t)$, up to the permutation of the first and ith coordinates. In the next subsection, it is explained how to draw such random variables.

It is really easy to give a formal proof of Proposition 8. Let $X_t^x \sim \mathrm{WIS}_d(x,\alpha,0,I_d^n;t)$ and f be a smooth function on $\mathcal{S}_d^+(\mathbb{R})$ such that the series below converge absolutely. By iterating Itô's formula, we have that $\mathbb{E}[f(X_t^x)] = \sum_{k=0}^{\infty} t^k L^k f(x)/k!$. Similarly, we also get by using the tower property of the conditional expectation that

$$\begin{aligned}
\mathbb{E}[f(X_t^{n,\dots X_t^{1,x}})] &= \mathbb{E}[\mathbb{E}[f(X_t^{n,\dots X_t^{1,x}})|X_t^{n-1,\dots X_t^{1,x}}]]\\
&= \sum_{k_n=0}^{+\infty} \frac{t^{k_n}}{k_n!}\mathbb{E}[L_{e_d^n}^{k_n} f(X_t^{n-1,\dots X_t^{1,x}})].
\end{aligned} \tag{16}$$

Simply by repeating this argument, we get that

$$\begin{aligned}\mathbb{E}[f(X_t^{n,\dots X_t^{1,x}})] &= \sum_{k_1,\dots,k_n=0}^{+\infty} \frac{t^{\sum_{i=1}^n k_i}}{k_1!\cdots k_n!} L_{e_d^1}^{k_1}\cdots L_{e_d^n}^{k_n} f(x) \\ &= \sum_{k=0}^{\infty} \frac{t^k}{k!}(L_{e_d^1}+\cdots+L_{e_d^n})^k f(x) = \mathbb{E}[f(X_t^x)].\end{aligned} \tag{17}$$

To get the second equality, we identify a Cauchy product and use that the operators $L_{e_d^1},\dots,L_{e_d^n}$ commute. To make this formal proof correct, one has to check that the series are well defined and can be switched with the expectation. This check is made in the Appendix C.1 for our framework and remains valid as soon as the operator $L_{e_d^i}$ and L are of affine type.

2.2. *Exact simulation for* $\mathrm{WIS}_d(x,\alpha,0,e_d^1;t)$. For the sake of clarity, we start with the case of $d=2$ that avoids complexities due to matrix decompositions. We deal with the general case just after.

2.2.1. *The case* $d=2$. We start by writing explicitly the infinitesimal generator $L_{e_2^1}$ of $\mathrm{WIS}_2(x,\alpha,0,e_2^1)$. From (8), we get

$$\begin{aligned} & x\in\mathcal{S}_2^+(\mathbb{R}), \\ & L_{e_2^1}f(x) = \alpha\partial_{\{1,1\}}f(x) + 2x_{\{1,1\}}\partial_{\{1,1\}}^2 f(x) \\ & \qquad + 2x_{\{1,2\}}\partial_{\{1,1\}}\partial_{\{1,2\}}f(x) + \frac{x_{\{2,2\}}}{2}\partial_{\{1,2\}}^2 f(x).\end{aligned} \tag{18}$$

We now show that this operator is in fact associated to an SDE that can be explicitly solved. We will denote by $(Z_t^1, t\geq 0)$ and $(Z_t^2, t\geq 0)$ two independent standard Brownian motions in $\mathbb{R}$.

When $x_{\{2,2\}}=0$, we also have $x_{\{1,2\}}=0$ since x is nonnegative. In that case,

$$\begin{aligned} X_0^x = x, &\qquad d(X_t^x)_{\{1,1\}} = \alpha\,dt + 2\sqrt{(X_t^x)_{\{1,1\}}}\,dZ_t^1, \\ d(X_t^x)_{\{1,2\}} = 0, &\qquad d(X_t^x)_{\{2,2\}} = 0\end{aligned} \tag{19}$$

has the infinitesimal generator (18), which is one of a CIR process (or of a squared Bessel process of dimension α to be more precise). By using an algorithm that samples exactly a noncentral chi-square distribution (see, e.g., Glasserman [11]), we can then sample $\mathrm{WIS}_2(x,\alpha,0,e_2^1;t)$ when $x_{\{2,2\}}=0$.

When $x_{\{2,2\}} > 0$, it easy to check that the SDE

$$\begin{aligned}
d(X_t^x)_{\{1,1\}} &= \alpha\, dt + 2\sqrt{(X_t^x)_{\{1,1\}} - \frac{((X_t^x)_{\{1,2\}})^2}{(X_t^x)_{\{2,2\}}}}\, dZ_t^1 \\
&\quad + 2\frac{(X_t^x)_{\{1,2\}}}{\sqrt{(X_t^x)_{\{2,2\}}}}\, dZ_t^2, \\
d(X_t^x)_{\{1,2\}} &= \sqrt{(X_t^x)_{\{2,2\}}}\, dZ_t^2, \\
d(X_t^x)_{\{2,2\}} &= 0,
\end{aligned} \tag{20}$$

starting from $X_0^x = x$, has an infinitesimal generator equal to $L_{e_2^1}$. To solve (20), we set

$$\begin{aligned}
(U_t^u)_{\{1,1\}} &= (X_t^x)_{\{1,1\}} - \frac{((X_t^x)_{\{1,2\}})^2}{(X_t^x)_{\{2,2\}}}, \\
(U_t^u)_{\{1,2\}} &= \frac{(X_t^x)_{\{1,2\}}}{\sqrt{x_{\{2,2\}}}}, \qquad (U_t^u)_{\{2,2\}} = x_{\{2,2\}}.
\end{aligned} \tag{21}$$

Here, u stands for the initial condition, that is, $u = U_0^u$. We get by using Itô calculus that

$$\begin{aligned}
d(U_t^u)_{\{1,1\}} &= (\alpha - 1)\, dt + 2\sqrt{(U_t^u)_{\{1,1\}}}\, dZ_t^1, \\
d(U_t^u)_{\{1,2\}} &= dZ_t^2 \quad \text{and} \quad d(U_t^u)_{\{2,2\}} = 0.
\end{aligned} \tag{22}$$

Therefore, $(U_t^u)_{\{1,2\}}$ and $(U_t^u)_{\{1,1\}}$ can be sampled, respectively, by independent Gaussian and noncentral chi-square variables. Then, we can get back X_t^x by inverting (21),

$$\begin{aligned}
(X_t^x)_{\{1,1\}} &= (U_t^u)_{\{1,1\}} + (U_t^u)_{\{1,2\}}^2, \\
(X_t^x)_{\{1,2\}} &= (U_t^u)_{\{1,2\}}\sqrt{(U_t^u)_{\{2,2\}}}, \qquad (X_t^x)_{\{2,2\}} = (U_t^u)_{\{2,2\}}.
\end{aligned} \tag{23}$$

This result gives an interesting way to figure out the dynamics associated to the operator $L_{e_2^1}$ by using a change of variable. It is worth noticing that the CIR process $(U_t^u)_{\{1,1\}}$ is well defined as soon as its degree $\alpha - 1$ is nonnegative, which coincides with the condition under which the Wishart process $\mathrm{WIS}_2(x, \alpha, 0, e_2^1)$ is well defined. Last, we notice that the solution of the operator $L_{e_2^1}$ involves a CIR process in the diagonal term and a Brownian motion in the nondiagonal one. A similar structure holds for larger d.

2.2.2. *The general case.* We now present a general way to sample exactly $\mathrm{WIS}_d(x,\alpha,0,e_d^1;t)$. We first write explicitly from (8) the infinitesimal generator of $\mathrm{WIS}_d(x,\alpha,0,e_d^1)$ for $x \in \mathcal{S}_d^+(\mathbb{R})$

$$\begin{aligned} L_{e_d^1} f(x) &= \alpha \partial_{\{1,1\}} f(x) + 2x_{\{1,1\}} \partial^2_{\{1,1\}} f(x) \\ &\quad + 2 \sum_{\substack{1\le m\le d\\ m\neq 1}} x_{\{1,m\}} \partial_{\{1,m\}} \partial_{\{1,1\}} f(x) \\ &\quad + \frac{1}{2} \sum_{\substack{1\le m,l\le d\\ m\neq 1, l\neq 1}} x_{\{m,l\}} \partial_{\{1,m\}} \partial_{\{1,l\}} f(x). \end{aligned} \tag{24}$$

As for $d=2$, we will construct an SDE that has the same infinitesimal generator $L_{e_d^1}$ and that can be solved explicitly. To do so however, we need to use further matrix decomposition results. In the case $d=2$, we have already noticed that we choose different SDEs whether $x_{2,2}=0$ or not. Here, the SDE will depend on the rank of the submatrix $(x_{i,j})_{2\le i,j\le d}$, and we set

$$r = \mathrm{Rk}((x_{i,j})_{2\le i,j\le d}) \in \{0,\dots,d-1\}.$$

First, we consider the case where

$$\begin{gathered} \exists c_r \in \mathcal{G}_r \text{ lower triangular}, \qquad k_r \in \mathcal{M}_{d-1-r\times r}(\mathbb{R}), \\ (x)_{2\le i,j\le d} = \begin{pmatrix} c_r & 0 \\ k_r & 0 \end{pmatrix} \begin{pmatrix} c_r^T & k_r^T \\ 0 & 0 \end{pmatrix} =: cc^T. \end{gathered} \tag{25}$$

With a slight abuse of notation, we consider that this decomposition also holds when $r=0$ with $c=0$. When $r=d-1$, $c=c_r$ is simply the usual Cholesky decomposition of $(x_{i,j})_{2\le i,j\le d}$. As it is explained in Corollary 11, we can still get such a decomposition up to a permutation of the coordinates $\{2,\dots,d\}$.

THEOREM 9. *Let us consider $x \in \mathcal{S}_d^+(\mathbb{R})$ such that (25) holds. Let $(Z_t^l)_{1\le l\le r+1}$ be a vector of independent standard Brownian motions. Then, the following SDE* [*convention $\sum_{k=1}^r(\cdots)=0$ when $r=0$*]

$$\begin{aligned} d(X_t^x)_{\{1,1\}} &= \alpha\, dt + 2\sqrt{(X_t^x)_{\{1,1\}} - \sum_{k=1}^r \left(\sum_{l=1}^r (c_r^{-1})_{k,l} (X_t^x)_{\{1,l+1\}}\right)^2}\, dZ_t^1 \\ &\quad + 2\sum_{k=1}^r \sum_{l=1}^r (c_r^{-1})_{k,l} (X_t^x)_{\{1,l+1\}}\, dZ_t^{k+1}, \\ d(X_t^x)_{\{1,i\}} &= \sum_{k=1}^r c_{i-1,k}\, dZ_t^{k+1}, \qquad i=2,\dots,d, \\ d\big((X_t^x)_{\{l,k\}}\big)_{2\le k,l\le d} &= 0 \end{aligned} \tag{26}$$

has a unique strong solution starting from x. It takes values in $\mathcal{S}_d^+(\mathbb{R})$ and has the infinitesimal generator $L_{e_d^1}$. Moreover, this solution is given explicitly by

$$X_t^x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{pmatrix}$$

$$\times \begin{pmatrix} (U_t^u)_{\{1,1\}} + \sum_{k=1}^{r} \big((U_t^u)_{\{1,k+1\}}\big)^2 & \big((U_t^u)_{\{1,l+1\}}\big)_{1\le l\le r}^T & 0 \\ \big((U_t^u)_{\{1,l+1\}}\big)_{1\le l\le r} & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{27}$$

$$\times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r^T & k_r^T \\ 0 & 0 & I_{d-r-1} \end{pmatrix},$$

where

$$\begin{aligned} d(U_t^u)_{\{1,1\}} &= (\alpha - r)\,dt + 2\sqrt{(U_t^u)_{\{1,1\}}}\,dZ_t^1, \\ u_{\{1,1\}} &= x_{\{1,1\}} - \sum_{k=1}^{r} \big(u_{\{1,k+1\}}\big)^2 \ge 0, \\ d\big((U_t^u)_{\{1,l+1\}}\big)_{1\le l\le r} &= (dZ_t^{l+1})_{1\le l\le r}, \\ \big(u_{\{1,l+1\}}\big)_{1\le l\le r} &= c_r^{-1}\big(x_{\{1,l+1\}}\big)_{1\le l\le r}. \end{aligned} \tag{28}$$

Once again, we have made a slight abuse of notation when $r = 0$, and (27) should be simply read as

$$X_t^x = \begin{pmatrix} (U_t^u)_{\{1,1\}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

in that case. In the statement above, it may seem weird that we use for u and U_t^u the same indexation as the one for symmetric matrices while we only use its first row (or column). The reason is that we can, in fact, see X_t^x as a function of U_t^u by setting

$$\begin{aligned} (U_t^u)_{\{i,j\}} &= u_{\{i,j\}} = x_{\{i,j\}} \qquad \text{for } i, j \ge 2 \quad \text{and} \\ (U_t^u)_{\{1,i\}} &= u_{\{1,i\}} = 0 \qquad \text{for } r+1 \le i \le d. \end{aligned} \tag{29}$$

Thus, (c_r, k_r, I_{d-1}) is an extended Cholesky decomposition of $((U_t^u)_{i,j})_{2\le i,j\le d}$ and can be seen as a function of U_t^u. We get from (27) that

$$X_t^x = h(U_t^u) \qquad \text{with } h(u) = \sum_{r=0}^{d-1} \mathbb{1}_{r=\mathrm{Rk}[(u_{i,j})_{2\le i,j\le d}]} h_r(u) \quad \text{and} \tag{30}$$

$$h_r(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r(u) & 0 \\ 0 & k_r(u) & I_{d-r-1} \end{pmatrix}$$
$$\times \begin{pmatrix} u_{\{1,1\}} + \sum_{k=1}^{r} (u_{\{1,k+1\}})^2 & (u_{\{1,l+1\}})_{1\le l\le r}^T & 0 \\ (u_{\{1,l+1\}})_{1\le l\le r} & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r(u)^T & k_r(u)^T \\ 0 & 0 & I_{d-r-1} \end{pmatrix},$$

where $(c_r(u), k_r(u), I_{d-1})$ is the extended Cholesky decomposition of $(u_{i,j})_{2\le i,j\le d}$ given by some algorithm (e.g., Golub and Van Loan [13], Algorithm 4.2.4). Equation (30) will later play an important role in analyzing discretization schemes.

The proof of Theorem 9 is given in Appendix C.2. It enables us to simulate exactly the distribution $\mathrm{WIS}_d(x, \alpha, 0, e_d^1; t)$ simply by sampling one noncentral chi-square distribution for $(U_t^u)_{\{1,1\}}$ (see Glasserman [11]) and r other independent Gaussian random variables. As in the $d = 2$ case, we notice that the condition which ensures that the CIR process $((U_t^u)_{\{1,1\}}, t \ge 0)$ is well defined for any $r \in \{0, \dots, d-1\}$, namely, $\alpha - (d-1) \ge 0$, is the same as the one required for the definition of $\mathrm{WIS}_d(x, \alpha, 0, e_d^1)$.

REMARK 10. From (27), we easily get by a calculation made in (47) that $\mathrm{Rk}(X_t^x) = \mathrm{Rk}((x_{i,j})_{2\le i,j\le d}) + \mathbb{1}_{(U_t^u)_{\{1,1\}} \ne 0}$, and therefore,

$$\mathrm{Rk}(X_t^x) = \mathrm{Rk}((x_{i,j})_{2\le i,j\le d}) + 1 \qquad \text{a.s.}$$

In particular, X_t^x is almost surely positive definite if $x \in \mathcal{S}_d^{+,*}(\mathbb{R})$.

Theorem 9 assumes that the initial value $x \in \mathcal{S}_d^+(\mathbb{R})$ satisfies (25). Now we explain why it is still possible, up to a permutation of the coordinates, to be in such a case. This relies on the extended Cholesky decomposition which is stated in Lemma 23.

COROLLARY 11. *Let $x \in \mathcal{S}_d^+(\mathbb{R})$ and (c_r, k_r, p) be an extended Cholesky decomposition of $(x_{i,j})_{2\le i,j\le d}$ (Lemma 23). Then, $\pi = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ is a permutation matrix, $\mathrm{WIS}_d(x, \alpha, 0, e_d^1) \underset{Law}{=} \pi^T \mathrm{WIS}_d(\pi x \pi^T, \alpha, 0, e_d^1)\pi$ and $((\pi x \pi^T)_{i,j})_{2\le i,j\le d} = \begin{pmatrix} c_r & 0 \\ k_r & 0 \end{pmatrix}\begin{pmatrix} c_r^T & k_r^T \\ 0 & 0 \end{pmatrix}$ satisfies (25).*

PROOF. The result comes directly from (13), since $\pi^T = \pi^{-1}$ and $\pi e_d^1 \pi^T = e_d^1$. $\square$

Therefore, by a combination of Corollary 11 and Theorem 9, we get a simple way to explicitly construct a process that has the infinitesimal generator $L_{e_d^1}$ for any initial condition $x \in \mathcal{S}_d^+(\mathbb{R})$. In particular, this enables us to sample exactly the Wishart distribution $\mathrm{WIS}_d(x, \alpha, 0, e_d^1; t)$. Algorithm 1 below sums up the whole procedure.

Algorithm 1: Exact simulation $\mathrm{WIS}_d(x, \alpha, 0, e_d^1; t)$

Input: $x \in \mathcal{S}_d^+(\mathbb{R})$, d, $\alpha \geq d-1$ and $t > 0$.
Output: X, sampled according to $\mathrm{WIS}_d(x, \alpha, 0, e_d^1; t)$.
Compute the extended Cholesky decomposition (p, k_r, c_r) of $(x_{i,j})_{2 \leq i,j \leq d}$ given by Lemma 23, $r \in \{0, \dots, d-1\}$ (see Golub and Van Loan [13] for an algorithm);
Set $\pi = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$, $\tilde{x} = \pi x \pi^T$, $(u_{\{1,l+1\}})_{1 \leq l \leq r} = (c_r)^{-1} (\tilde{x}_{\{1,l+1\}})_{1 \leq l \leq r}$ and $u_{\{1,1\}} = \tilde{x}_{\{1,1\}} - \sum_{k=1}^r (u_{\{1,k+1\}})^2 \geq 0$;
Sample independently r normal variables $G_2, \dots, G_{r+1} \sim \mathcal{N}(0,1)$ and $(U_t^u)_{\{1,1\}}$ as a CIR process at time t starting from $u_{\{1,1\}}$ solving $d(U_t^u)_{\{1,1\}} = (\alpha - r)\, dt + 2\sqrt{(U_t^u)_{\{1,1\}}}\, dZ_t^1$ (see Glasserman [11]).
Set $(U_t^u)_{\{1,l+1\}} = u_{\{1,l+1\}} + \sqrt{t} G_{l+1}$;
return

$$
\begin{aligned}
X = \pi^T &\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{pmatrix} \\
&\times \begin{pmatrix} (U_t^u)_{\{1,1\}} + \sum_{k=1}^r \bigl((U_t^u)_{\{1,k+1\}}\bigr)^2 & \bigl((U_t^u)_{\{1,l+1\}}\bigr)_{1 \leq l \leq r}^T & 0 \\ \bigl((U_t^u)_{\{1,l+1\}}\bigr)_{1 \leq l \leq r} & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&\times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r^T & k_r^T \\ 0 & 0 & I_{d-r-1} \end{pmatrix} \pi.
\end{aligned}
$$

Let us now discuss the complexity of Algorithm 1. The number of operations required by the extended Cholesky decomposition is of order $O(d^3)$. From a computational point of view, the permutation is handled directly and does not require any matrix multiplication so that we can consider w.l.o.g. that $\pi = I_d$. Since c_r is lower triangular, the calculation of $u_{\{1,i\}}$, $i = 1, \dots, r+1$, only requires $O(d^2)$ operations. Also, we do not perform in practice the matrix product (27), but only compute the values of $X_{\{1,i\}}$ for $i = 1, \dots, d$, which also requires $O(d^2)$ operations. Last, d samples are at most required. To sum up, it comes out that the complexity of Algorithm 1 is of order $O(d^3)$.

2.3. *Exact simulation for Wishart processes.* We have now shown all the mathematical results that enable us to give an exact simulation method for general Wishart processes. This is made in two steps.

First, we know how to sample exactly $\mathrm{WIS}_d(x, \alpha, 0, e_d^1; t)$ thanks to Theorem 9 and Corollary 11. By a simple permutation of the first and kth coordinates, we are then also able to sample according to $\mathrm{WIS}_d(x, \alpha, 0, e_d^k; t)$ for $k \in \{1, \dots, d\}$. Thus, we get by Proposition 8 an exact simulation method to sample $\mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$. It is given explicitly in Algorithm 2. Then we get an exact simulation scheme for $\mathrm{WIS}_d(x, \alpha, b, a; t)$ by using the law identity (14) (see Algorithm 3).

Algorithm 2: Exact simulation for $\mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$

Input: $x \in \mathcal{S}_d^+(\mathbb{R})$, $n \le d$, $\alpha \ge d - 1$ and $t > 0$.
Output: X, sampled according to $\mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$
$y = x$
for $k = 1$ **to** n **do**
  Set $p_{k,1} = p_{1,k} = p_{i,i} = 1$ for $i \notin \{1, k\}$ and $p_{i,j} = 0$ otherwise (permutation of the first and kth coordinates).
  $y = pYp$ where Y is sampled according to $\mathrm{WIS}_d(pyp, \alpha, 0, e_d^1; t)$ by using Algorithm 1.
end
return $X = y$.

Algorithm 3: Exact simulation for $\mathrm{WIS}_d(x, \alpha, b, a; t)$

Input: $x \in \mathcal{S}_d^+(\mathbb{R})$, $\alpha \ge d - 1$, $a, b \in \mathcal{M}_d(\mathbb{R})$ and $t > 0$.
Output: X, sampled according to $\mathrm{WIS}_d(x, \alpha, b, a; t)$.
Calculate $q_t = \int_0^t \exp(sb)a^T a \exp(sb^T)\,ds$ and (p, c_n, k_n) an extended Cholesky decomposition of q_t/t.
Set $\theta_t = p^{-1}\begin{pmatrix} c_n & 0 \\ k_n & I_{d-n} \end{pmatrix}$ and $m_t = \exp(tb)$.
return $X = \theta_t Y \theta_t^T$, where $Y \sim \mathrm{WIS}_d(\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T, \alpha, 0, I_d^n; t)$ is sampled by Algorithm 2.

Let us analyze the overall complexity of Algorithm 3. Since it basically runs n times Algorithm 1, it requires a complexity of order $O(nd^3)$ and therefore at most of order $O(d^4)$. As we have seen, the "bottleneck" of Algorithm 1 is the extended Cholesky decomposition which is in $O(d^3)$. All the other steps in Algorithm 1 require at most $O(d^2)$ operations. A natural question for Algorithm 2 is to wonder if we can reuse the Cholesky decomposition between the loops instead of calculating it from scratch. For example, if it were possible to get the Cholesky decomposition of loop $k+1$ from the one of loop k at a cost $O(d^2)$, the complexity of Algorithms 2 and 3 would then drop to $O(d^3)$. Despite our investigations, we have not been able to do so up to now.

REMARK 12. When $\alpha \geq 2d-1$, it is possible to sample $\mathrm{WIS}_d(x,\alpha,0,I_d^n;t)$ in $O(d^3)$ by another mean. If $X_t^1 \sim \mathrm{WIS}_d(x,d,0,I_d^n;t)$ and $X_t^2 \sim \mathrm{WIS}_d(0,\alpha-d,0,I_d^n;t)$ are independent, we can check that $X_t^1+X_t^2 \sim \mathrm{WIS}_d(x,\alpha,0,I_d^n;t)$. Then, X_t^1 can be sampled by using Proposition 21 and X_t^2 by using Bartlett's decomposition (31) since $X_t^2 \underset{\text{Law}}{=} t\,\mathrm{WIS}_d(0,\alpha-d,0,I_d^n;1)$ from (11).

2.4. *The Bartlett's decomposition revisited.* Now we would like to illustrate our exact simulation method on the particular case $\mathrm{WIS}_d(0,\alpha,0,I_d^n;1)$, which is known in the literature as the central Wishart distribution. In that case, we can perform explicitly the composition $X_1^{n,\dots X_1^{1,0}}$ given by Proposition 8. We will show by an induction on n that

$$X_1^{n,\dots X_1^{1,0}} = \begin{pmatrix} (L_{i,j})_{1\leq i,j\leq n} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (L_{i,j}^T)_{1\leq i,j\leq n} & 0 \\ 0 & 0 \end{pmatrix}, \tag{31}$$

where $(L_{i,j})_{1\leq j<i\leq d}$ and $L_{i,i}$ are independent random variables such that $L_{i,j} \sim \mathcal{N}(0,1)$ and $(L_{i,i})^2 \sim \chi^2(\alpha-i+1)$ and $L_{i,j}=0$ for $i<j$. This result is known as the Bartlett's decomposition and dates back to 1933 (see Kshirsagar [18] or Kabe [17]).

For $n=1$, we know from Theorem 9 that $(X_1^{1,0})_{1,1} \sim \chi^2(\alpha)$ since $d(X_t^{1,0})_{1,1} = \alpha\,dt + 2\sqrt{(X_t^{1,0})_{1,1}}\,dZ_t^1$ with $(X_0^{1,0})_{1,1}=0$, and all the other elements are equal to 0. Let us assume now that the induction hypothesis is satisfied for $n-1$. Then, we can apply once again Theorem 9 (up to the permutation of the first and nth coordinates). We have $\mathrm{Rk}(X_1^{n-1,\dots X_1^{1,0}}) = n-1$, a.s., and the Cholesky decomposition is directly given by $(L_{i,j})_{1\leq i,j\leq n-1}$. Then, we get from (27) that there are independent variables $L_{n,n}^2 \sim \chi^2(\alpha-n+1)$ and $L_{n,i} \sim \mathcal{N}(0,1)$ for $i \in \{1,\dots,n-1\}$ such that

$$\begin{aligned} X_1^{n,\dots X_1^{1,0}} &= \begin{pmatrix} (L_{i,j})_{1\leq i,j\leq n-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{d-n} \end{pmatrix} \\ &\quad\times \begin{pmatrix} I_{n-1} & (L_{n,i})_{1\leq i\leq n-1} & 0 \\ (L_{n,i})_{1\leq i\leq n-1}^T & \sum_{i=1}^n L_{n,i}^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\quad\times \begin{pmatrix} (L_{i,j})_{1\leq i,j\leq n-1}^T & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{d-n} \end{pmatrix}. \end{aligned}$$

Since

$$\begin{pmatrix} I_{n-1} & (L_{n,i})_{1\le i\le n-1} & 0 \\ (L_{n,i})^T_{1\le i\le n-1} & \sum_{i=1}^n L_{n,i}^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} I_{n-1} & 0 & 0 \\ (L_{n,i})^T_{1\le i\le n-1} & L_{n,n} & 0 \\ 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} I_{n-1} & (L_{n,i})_{1\le i\le n-1} & 0 \\ 0 & L_{n,n} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

we conclude by induction on n.

3. High-order discretization schemes for Wishart and semidefinite positive affine processes. In this section, we switch from exact sampling to approximate schemes. First, this will enable us to simulate not only Wishart processes, but also general affine processes. More importantly, the discretization schemes that we introduce are in practice faster than the exact simulation scheme, especially if one has to sample entire paths. This will be illustrated in Section 4.

When dealing with discretization schemes, splitting operators is a powerful technique to construct schemes for SDEs from other schemes obtained on simpler SDEs. This idea of splitting originates from the seminal work of Strang [27] in the field of ODEs. As pointed out by Ninomiya and Victoir [24] or Alfonsi [2], it is rather easy to analyze the weak error (i.e., the error made on marginal distributions) of schemes obtained by splitting. Indeed, this can be done simply by using the same arguments as Talay and Tubaro [28] for the Euler–Maruyama scheme. Nonetheless, when we use the splitting technique for SDEs that are defined on a given domain [$\mathcal{S}_d^+(\mathbb{R})$ in our case], one has to be careful that the discretization scheme remains in it. For example, in the case of the CIR diffusion (i.e., $d = 1$), general splitting methods such as Ninomiya and Victoir [24] fail to preserve the domain $\mathbb{R}^+$. It is, in fact, only well defined for $\alpha \ge 1$, while the CIR process exists for any $\alpha \ge 0$ (see Alfonsi [2]). Of course, the same remark holds for Wishart and affine processes. This is why we will use the ad hoc splitting (7) instead of general splitting methods, which enables us to get schemes that preserve $\mathcal{S}_d^+(\mathbb{R})$ and are defined without any restriction on the parameters.

The analysis of the strong error of our schemes is beyond the scope of this paper. In fact, behind the term "strong error" we have in mind here two different things. First, it can be the error made on pathwise expectations between the discretization scheme and the exact scheme. This kind of error is illustrated numerically in the next section (Figure 3) and seems to be of the same order as the weak error, even though we are not at all able to mathematically show this result. Second, "strong error" can also mean the pathwise error between the discretization scheme and the exact solution for a given Brownian motion $(W_t, t \ge 0)$. The rate of convergence for this kind of error has been analyzed for the CIR in Alfonsi [1] and is really low. This is mainly due to the fact that the square root is not Lipschitz near 0. Fortunately, discretization schemes are mostly used to compute expectations with a Monte Carlo algorithm. In this context, pathwise error is not so relevant.

To our knowledge, there are very few papers in the literature that deal with discretization schemes for Wishart processes. Recently, Benabid, Bensusan and Karoui [3] have proposed a Monte Carlo method to calculate expectations on Wishart processes which is based on a Girsanov change of probability. Gauthier and Possamai [10] introduce a moment-matching scheme for Wishart processes. Both methods are well defined under some restrictions on the parameters, and there is no theoretical result on their accuracy. Currently, Teichmann [29] is working on dedicated schemes for general affine processes by approximating their characteristic functions.

This section is structured as follows. First, we recall basic results on the splitting technique to get discretization schemes for SDEs. We will take the same framework as Alfonsi [2] since it is somehow designed for affine processes. Then we will explain how to get high-order schemes for $WIS_d(x, \alpha, 0, e_d^1)$ from the construction given by Theorem 9. The remarkable splitting (15) will then enable us to get high-order schemes for $WIS_d(x, \alpha, 0, I_d^n)$. From this result, we will be able to get a second-order scheme for any semidefinite positive affine processes and a third-order scheme for Wishart processes.

3.1. *Weak error analysis and splitting methods.* Let us start with some notation. We consider a time horizon $T > 0$ and the regular time grid defined by $t_i^N = iT/N, i = 0, \dots, N$. When considering a Markovian process on a domain $\mathbb{D}$, a discretization scheme is a way to sample the value at a given time step $t > 0$, starting from the current value $x \in \mathbb{D}$. It is thus described by a probability measure $\hat{p}_x(t)(dz)$ on $\mathbb{D}$, and we denote by $\hat{X}_t^x$ a random variable that follows this law. Then the full discretization on the regular time grid associated to this scheme from $x \in \mathbb{D}$ is simply a sequence $(\hat{X}^N_{t_i^N}, 0 \le i \le N)$ of random variables such that:

- $\hat{X}^N_{t_0^N} = x$,
- the law of $\hat{X}^N_{t_{i+1}^N}$ is sampled according to $\hat{p}_{\hat{X}^N_{t_i^N}}(T/N)(dz)$ independently from the previous samples, that is, $\mathbb{E}[f(\hat{X}^N_{t_{i+1}^N})|(\hat{X}^N_{t_j^N}, 0 \le j \le i)] = \int_{\mathbb{D}} f(z)\hat{p}_{\hat{X}^N_{t_i^N}}(T/N)(dz)$ for any bounded measurable function $f : \mathbb{D} \to \mathbb{R}$.

Now we focus on the analysis of the weak error $\mathbb{E}[f(X_T^x)] - \mathbb{E}[f(\hat{X}^N_{t_N^N})]$. There is a huge literature on this topic. Talay and Tubaro [28] have obtained an expansion error for Euler–Maruyama and Milstein schemes. This error has also been studied on other schemes: we cite the articles of Kusuoka [19], Lyons and Victoir [21], Ninomiya and Victoir [24], and Ninomiya and Ninomiya [23], to mention a few. However, to our knowledge, most of these papers make regularity assumptions on the SDE coefficients that are not satisfied by affine diffusions. Typically, they assume that these coefficients are $\mathcal{C}^\infty$ with bounded derivatives. This is not satisfied by general affine diffusions because of the square root diffusion term. For this reason, Alfonsi [2] introduced a framework that allows us to rigorously analyze the weak error for affine diffusions. In this paper, we will naturally work under this framework. Unfortunately, this requires us to introduce some definitions, and we present here only the main ones.

We consider a domain $\mathbb{D} \subset \mathbb{R}^\zeta$, $\zeta \in \mathbb{N}^*$, and L an operator associated to an SDE defined on $\mathbb{D}$. Mainly (but not only), we consider in this paper $\mathbb{D} = \mathcal{S}_d^+(\mathbb{R}) \subset \mathcal{S}_d(\mathbb{R}) \simeq \mathbb{R}^{d(d+1)/2}$. For $\gamma = (\gamma_1, \dots, \gamma_\zeta) \in \mathbb{N}^\zeta$, we define $\partial_\gamma = \partial_1^{\gamma_1}, \dots, \partial_\zeta^{\gamma_\zeta}$ and $|\gamma| = \sum_{i=1}^\zeta \gamma_i$ and set

$$\mathcal{C}^\infty_{\mathrm{pol}}(\mathbb{D}) = \{f \in \mathcal{C}^\infty(\mathbb{D}, \mathbb{R}), \forall \gamma \in \mathbb{N}^\zeta, \exists C_\gamma > 0, e_\gamma \in \mathbb{N}^*, \\ \forall x \in \mathbb{D}, |\partial_\gamma f(x)| \le C_\gamma(1 + \|x\|^{e_\gamma})\},$$

where $\|\cdot\|$ is a norm on $\mathbb{R}^\zeta$. We say that $(C_\gamma, e_\gamma)_{\gamma \in \mathbb{N}^\zeta}$ is a *good sequence* for $f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathbb{D})$ if one has $|\partial_\gamma f(x)| \le C_\gamma(1 + \|x\|^{e_\gamma})$. The operator L is said to satisfy the *required assumptions* if it can be written as $L = \sum_{0<|\gamma|\le 2} a_\gamma(x)\partial_\gamma$, with $a_\gamma \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathbb{D})$. This property holds for affine diffusions since any a_γ is an affine function. We will say that $\hat{X}_t^x$ is a *potential weak νth-order scheme for the operator L* if for any function $f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathbb{D})$ with a good sequence $(C_\gamma, e_\gamma)_{\gamma \in \mathbb{N}^\zeta}$, there exist positive constants C, E and η depending only on $(C_\gamma, e_\gamma)_{\gamma \in \mathbb{N}^\zeta}$ such that

$$\forall t \in (0, \eta) \\ \left|\mathbb{E}[f(\hat{X}_t^x)] - \left[f(x) + \sum_{k=1}^{\nu} \frac{1}{k!} t^k L^k f(x)\right]\right| \le C t^{\nu+1}(1 + \|x\|^E). \tag{32}$$

Roughly speaking, this is the main assumption that a discretization scheme should satisfy to get a weak error of order ν. This is precised by the following theorem given in [2] that relies on the idea developed by Talay and Tubaro [28] for the Euler–Maruyama scheme.

THEOREM 13. *Let L be an operator satisfying the required assumptions on $\mathbb{D}$. We assume that*:

(1) *$\hat{X}_t^x$ is a potential weak νth-order scheme for L, and the scheme has uniformly bounded moments, that is,*

$$\exists n_0 \in \mathbb{N}^*, \forall q \in \mathbb{N}^* \qquad \sup_{N \ge n_0, 0 \le i \le N} \mathbb{E}[\|\hat{X}^N_{t_i^N}\|^q] < \infty; \tag{33}$$

(2) *$f : \mathbb{D} \to \mathbb{R}$ is a function such that $u(t, x) = \mathbb{E}[f(X^x_{T-t})]$ is defined on $[0, T] \times \mathbb{D}$, $\mathcal{C}^\infty$, solves $\forall t \in [0, T], \forall x \in \mathbb{D}, \partial_t u(t, x) = -Lu(t, x)$ and satisfies*

$$\forall l \in \mathbb{N}, \gamma \in \mathbb{N}^\zeta, \exists C_{l,\gamma}, e_{l,\gamma} > 0, \forall x \in \mathbb{D}, t \in [0, T] \\ |\partial_t^l \partial_\gamma u(t, x)| \le C_{l,\gamma}(1 + \|x\|^{e_{l,\gamma}}). \tag{34}$$

Then, there is $K > 0$, $N_0 \in \mathbb{N}$, *such that* $|\mathbb{E}[f(\hat{X}^N_{t^N_N})] - \mathbb{E}[f(X^x_T)]| \le K/N^\nu$ *for* $N \ge N_0$.

It is really important to notice that only assumption (1) depends on the discretization scheme. Assumption (2) just depends on the underlying diffusion. Since we only have a hold over the discretization scheme, this means from a numerical point of view that we mainly have to focus on assumption (1) to construct an accurate scheme. From a mathematical point of view, the regularity of the Cauchy problem which is required by assumption (2) is a tough problem that is interesting in its own. General results have been obtained in Talay and Tubaro [28] when b and σ are $\mathcal{C}^\infty$ with bounded derivatives. In the case of Wishart processes, we are able to get (34) when $f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d(\mathbb{R}))$.

PROPOSITION 14. *Let* $(X^x_t)_{t\ge 0} \sim \mathrm{WIS}_d(x, \alpha, b, a)$ *and* L *the associated generator. Let* $f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d(\mathbb{R}))$, $x \in \mathcal{S}^+_d(\mathbb{R})$ *and* $T > 0$. *Then,* $\tilde{u}(t, x) = \mathbb{E}[f(X^x_t)]$ *is* $\mathcal{C}^\infty$ *on* $[0, T] \times \mathcal{S}^+_d(\mathbb{R})$, *solves* $\partial_t \tilde{u}(t, x) = L\tilde{u}(t, x)$ *and its derivatives satisfy*

$$\begin{aligned} &\forall l \in \mathbb{N}, \forall n \in \mathbb{N}^{d(d+1)/2}, \exists C_{l,n}, e_{l,n} > 0, \forall x \in \mathcal{S}^+_d(\mathbb{R}), \forall t \in [0, T] \\ &\left| \partial^l_t \prod_{1\le i\le j\le d} \partial^{n_{\{i,j\}}}_{\{i,j\}} \tilde{u}(t, x) \right| \le C_{l,n}(1 + \|x\|^{e_{l,n}}). \end{aligned} \tag{35}$$

The proof of this result is made in Appendix D.1. It relies on the explicit formula of the characteristic function (10) and, more exactly, on the property stated in Lemma 26. Unfortunately, we have not been able to show an analogous result for general affine processes $\mathrm{AFF}_d(x, \bar{\alpha}, B, a)$. We deem that (35) also holds in that case, but this remains an open question.

Let us now turn to assumption (1) of Theorem 13. Usually, the boundedness of moments is not a big issue and requires, in general, tedious calculations. This basically holds when the drift and the diffusion coefficients have a sublinear growth, which is the case here. Conversely, it is much more difficult to find a scheme which is a potential ν-order scheme and stays at the same time in the domain $\mathcal{S}^+_d(\mathbb{R})$. For example, the Euler–Maruyama scheme is, generally speaking, a potential first-order scheme. However, it does not stay in $\mathcal{S}^+_d(\mathbb{R})$ even for the CIR case ($d = 1$). Still, for the CIR process, higher-order schemes such as Ninomiya and Victoir [24] or Ninomiya and Ninomiya [23] stay in $\mathbb{R}^+$ only under additional restrictions on the parameters. To solve this problem and get high-order schemes that remain in $\mathcal{S}^+_d(\mathbb{R})$, we will construct ad hoc discretization schemes by taking advantage of the remarkable splitting (15). In fact, the property of being a potential νth-order schemes is really easy to handle by scheme composition, especially when $\nu = 2$. This kind of result dates back to Strang [27] in the field of ODEs. In our framework, we recall a result that is stated in [2].

PROPOSITION 15. *Let L_1, L_2 be the generators of SDEs defined on $\mathbb{D}$ that satisfy the required assumption on $\mathbb{D}$. Let $\hat{X}_t^{1,x}$ and $\hat{X}_t^{2,x}$ denote, respectively, two potential weak νth-order schemes on $\mathbb{D}$ for L_1 and L_2.*

(1) *If $L_1L_2 = L_2L_1$, $\hat{X}_t^{2,\hat{X}_t^{1,x}}$ is a potential weak νth-order discretization scheme for $L_1 + L_2$.*

(2) *Let B be an independent Bernoulli variable of parameter $1/2$. If $\nu \geq 2$,*

$$\text{(a)}\quad B\hat{X}_t^{2,\hat{X}_t^{1,x}} + (1-B)\hat{X}_t^{1,\hat{X}_t^{2,x}} \quad \text{and} \quad \text{(b)}\quad \hat{X}_{t/2}^{2,\hat{X}_t^{1,\hat{X}_{t/2}^{2,x}}}$$

are potential weak second-order schemes for $L_1 + L_2$.

Let us explain the notation above. The composition $\hat{X}_{t_2}^{2,\hat{X}_{t_1}^{1,x}}$ means that we first use the scheme 1 with time step t_1 and then, conditionally to $\hat{X}_{t_1}^{1,x}$, we sample the scheme 2 with initial value $\hat{X}_{t_1}^{1,x}$ and time step t_2. To be explicit, it has the law $\int_{\mathbb{D}} \hat{p}_y^2(t_2)(dz)\hat{p}_x^1(t_1)(dy)$, where $\hat{p}_x^i(t_i)(dz)$ denotes the law of $\hat{X}_{t_i}^{i,x}$, $i = 1, 2$.

3.2. *High-order schemes for Wishart processes.* In this paragraph, we will give a way to get weak νth-order schemes for any Wishart processes. The construction is similar to the one used for the exact scheme. First, we obtain a νth-order scheme for $\mathrm{WIS}_d(x,\alpha,0,e_d^1)$. Then, we get a νth-order scheme for $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$ from the splitting (15) and Proposition 15. Last, we use the identity in law (14) to get a weak νth-order scheme for any Wishart process.

Let us start then by introducing a potential weak νth-order scheme for $\mathrm{WIS}_d(x,\alpha,0,e_d^1)$. Roughly speaking, we obtain this scheme from the exact scheme given by Theorem 9 and Corollary 11 by replacing the Gaussian random variables with moment matching variables and the exact CIR distribution with a sample according to a potential weak νth-order scheme for the CIR.

THEOREM 16. *Let $x \in \mathcal{S}_d^+(\mathbb{R})$ and (c_r, k_r, p) be an extended Cholesky decomposition of $(x_{i,j})_{2\leq i,j\leq d}$. We set $\pi = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ and $\tilde{x} = \pi x \pi^T$, so that $(\tilde{x}_{i,j})_{2\leq i,j\leq d} = \begin{pmatrix} c_r & 0 \\ k_r & 0 \end{pmatrix}\begin{pmatrix} c_r^T & k_r^T \\ 0 & 0 \end{pmatrix}$. As in Theorem 9, we have*

$$u_{\{1,1\}} = \tilde{x}_{\{1,1\}} - \sum_{k=1}^{r} (u_{\{1,k+1\}})^2 \geq 0,$$

where

$$(u_{\{1,l+1\}})_{1\leq l\leq r} = c_r^{-1}(\tilde{x}_{\{1,l+1\}})_{1\leq l\leq r},$$

and we set $u_{\{1,i\}} = 0$ if $r + 2 \leq i \leq d$ and $u_{\{i,j\}} = \tilde{x}_{\{i,j\}}$ if $i, j \geq 2$. Let $(\hat{G}^i)_{1\leq i\leq r}$ be a sequence of independent real variables with finite moments of any order such *that*

$$\forall i \in \{1, \dots, r\}, \forall k \le 2\nu + 1 \qquad \mathbb{E}[(\hat{G}^i)^k] = \mathbb{E}[G^k] \qquad \textit{where } G \sim \mathcal{N}(0, 1).$$

Let h_r be the function defined by (30). Let $(\hat{U}_t^u)_{\{1,1\}}$ be sampled independently according to a potential weak νth-order scheme for the CIR process $d(U_t^u)_{\{1,1\}} = (\alpha - r)\,dt + 2\sqrt{(U_t^u)_{\{1,1\}}}\,dZ_t^1$ starting from $u_{\{1,1\}}$. We set

$$(\hat{U}_t^u)_{\{1,i\}} = u_{\{1,i\}} + \sqrt{t}\hat{G}^i, \qquad 2 \le i \le r + 1,$$
$$(\hat{U}_t^u)_{\{1,i\}} = 0, \qquad r + 2 \le i \le d,$$
$$(\hat{U}_t^u)_{\{i,j\}} = u_{\{i,j\}} \qquad \textit{if } i, j \ge 2.$$

Then, the scheme $\hat{X}_t^x = \pi^T h_r(\hat{U}_t^u)\pi$ is a potential νth-order scheme for $L_{e_d^1}$ and takes values in $\mathcal{S}_d^+(\mathbb{R})$.

Let us give the idea of the proof. By construction, we have $\hat{X}_t^x \in \mathcal{S}_d^+(\mathbb{R})$ since an analogous formula to (27) holds for $\hat{X}_t^x$. The tedious part is to check that it is a potential νth-order scheme. We know from Theorem 9, equation (30) and Corollary 11 that we have $X_t^x = \pi^T h_r(U_t^u)\pi$. It is easy to check that $\hat{U}_t^u$ is a potential νth-order scheme for the operator associated to the diffusion U_t^u. Let us suppose for a while that $h_r(u) \in \mathcal{C}_{\text{pol}}^\infty(\mathcal{S}_d(\mathbb{R}))$. Then, $u \mapsto f(\pi^T h_r(u)\pi)$ is also in $\mathcal{C}_{\text{pol}}^\infty(\mathcal{S}_d(\mathbb{R}))$, and for any $f \in \mathcal{C}_{\text{pol}}^\infty(\mathcal{S}_d^+(\mathbb{R}))$, there are constants $C, E, \eta > 0$ depending only on a good sequence of f such that

$$|\mathbb{E}[f(\pi^T h_r(\hat{U}_t^u)\pi)] - \mathbb{E}[f(X_t^x)]| \le Ct^{\nu+1}(1 + \|x\|^E),$$

which basically gives the desired result. Unfortunately, h_r is not in $\mathcal{C}_{\text{pol}}^\infty(\mathcal{S}_d(\mathbb{R}))$. In fact, h_r is only smooth with respect to the coefficients of the first row and the first columns. However, these coefficients are also the only ones that are changed by $\hat{U}_t^u$ [the submatrix $((\hat{U}_t^u)_{i,j})_{2\le i,j\le d} = (u_{i,j})_{2\le i,j\le d}$ is constant], and it comes out that the regularity on h_r is sufficient to get a potential νth-order scheme for $L_{e_d^1}$. This is shown rigorously in the preprint version of this paper at the cost of additional technical definitions such as the "immersion property" that we do not reproduce here.

Now we briefly comment on the practical implementation of Theorem 16. Second and third-order schemes for the CIR process satisfying can be found in Alfonsi [2]. We can therefore get second (resp., third) order schemes for $L_{e_d^1}$ by taking any variables that matches the five (resp., the seven) first moments of $\mathcal{N}(0, 1)$. This can be obtained by taking

$$\mathbb{P}(\hat{G}^i = \sqrt{3}) = \mathbb{P}(\hat{G}^i = -\sqrt{3}) = \tfrac{1}{6} \quad \text{and} \quad \mathbb{P}(\hat{G}^i = 0) = \tfrac{2}{3}, \tag{36}$$

respectively,

$$\text{(37)}\qquad \begin{aligned} &\mathbb{P}\big(\hat{G}^i = \varepsilon\sqrt{3+\sqrt{6}}\big) = \frac{\sqrt{6}-2}{4\sqrt{6}}, \\ &\mathbb{P}\big(\hat{G}^i = \varepsilon\sqrt{3-\sqrt{6}}\big) = \frac{1}{2} - \frac{\sqrt{6}-2}{4\sqrt{6}}, \qquad \varepsilon \in \{-1,1\}. \end{aligned}$$

We focus now on the construction of a potential weak νth-order scheme for $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$. Let $\hat{X}_t^{1,x}$ denote a potential weak νth-order scheme for $\mathrm{WIS}_d(x,\alpha,0,e_d^1)$. For $i \in \{2,\dots,d\}$, $\mathrm{WIS}_d(x,\alpha,0,e_d^i)$ and $\mathrm{WIS}_d(x,\alpha,0,e_d^1)$ have the same law up to the permutation of the first and ith coordinate. Let $\pi^{1\leftrightarrow i}$ denote the associated permutation matrix. Then, we easily get that

$$\hat{X}_t^{i,x} = \pi^{1\leftrightarrow i}\hat{X}_t^{1,\pi^{1\leftrightarrow i}x\pi^{1\leftrightarrow i}}\pi^{1\leftrightarrow i}$$

is a potential νth-order scheme for $\mathrm{WIS}_d(x,\alpha,0,e_d^i)$. Last, we get from Theorem 7 and the point 1 of Proposition 15 that

$$\text{(38)}\qquad \hat{X}_t^{n,\dots^{\hat{X}_t^{1,x}}} \text{ is a potential weak } \nu\text{th-order scheme for } \mathrm{WIS}_d(x,\alpha,0,I_d^n).$$

Now we are in position to construct a scheme for any Wishart process $\mathrm{WIS}_d(x,\alpha,b,a)$ thanks to the identity (14). Let $\theta_t \in \mathcal{G}_d(\mathbb{R})$ be such as in Proposition 6 and $\hat{Y}_t^y$ denote a potential weak νth-order scheme for $\mathrm{WIS}_d(y,\alpha,0,I_d^n)$. Then we consider the following scheme for $\mathrm{WIS}_d(x,\alpha,b,a)$:

$$\text{(39)}\qquad \hat{X}_t^x = \theta_t \hat{Y}_t^{\theta_t^{-1}m_t x m_t^T(\theta_t^{-1})^T}\theta_t^T.$$

Unfortunately, we need to make some technical restrictions on a and b [namely, $a \in \mathcal{G}_d(\mathbb{R})$ or $ba^Ta = a^Tab$] to show that we get like this a potential νth-order scheme. We, however, believe that this is rather due to our analysis of the error and that the scheme converges as well without this restriction. In addition, we mention that we give in the next section a second-order scheme based on Proposition 5 for which we can make our error analysis for any parameters.

PROPOSITION 17. *Let $t > 0$, $a, b \in \mathcal{M}_d(\mathbb{R})$ and $\alpha \ge d-1$. Let $m_t = \exp(tb)$, $q_t = \int_0^t \exp(sb)a^Ta\exp(sb^T)\,ds$ and $n = \mathrm{Rk}(a^Ta)$. We assume that either $a \in \mathcal{G}_d(\mathbb{R})$ or b and a^Ta commute. We define*:

- *if $n = d$, θ_t as the (usual) Cholesky decomposition of q_t/t,*
- *if $n < d$, $\theta_t = \sqrt{\frac{1}{t}\int_0^t \exp(sb)\exp(sb^T)\,ds}\,p^{-1}\begin{pmatrix} c_n & 0 \\ k_n & I_{d-n}\end{pmatrix}$ where (c_n, k_n, p) is the extended Cholesky decomposition of a^Ta otherwise.*

In both cases, $\theta_t \in \mathcal{G}_d(\mathbb{R})$ and the scheme (39) is a potential weak νth-order scheme for $\mathrm{WIS}_d(x,\alpha,b,a)$.

The proof of Proposition 17 is left in Appendix D.2. From Theorem 13, we finally get the following result by using Propositions 14, 17.

THEOREM 18. *Let* $(X_t^x)_{t\geq 0} \sim \mathrm{WIS}_d(x,\alpha,b,a)$ *such that either* $a \in \mathcal{G}_d(\mathbb{R})$ *or* $a^T ab = ba^T a$ *and* $f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d(\mathbb{R}))$. *Let* $(\hat{X}^N_{t_i^N}, 0 \leq i \leq N)$ *be sampled with the scheme defined by Proposition* 17 *and Theorem* 16 *with the third-order scheme for the CIR given in* [2]. *Then*,

$$\exists C, N_0 > 0, \forall N \geq N_0 \qquad |\mathbb{E}[f(\hat{X}^N_{t_N^N})] - \mathbb{E}[f(X_T^x)]| \leq C/N^3.$$

3.3. *Second-order schemes for affine diffusions on* $\mathcal{S}_d^+(\mathbb{R})$. In this part, we present a potential second-order scheme for $\mathrm{AFF}_d(x,\bar{\alpha},B,a)$. Thanks to Proposition 5, there is $u \in \mathcal{G}_d(\mathbb{R})$ and a diagonal matrix $\bar{\delta}$ such that $\bar{\alpha} = u^T \bar{\delta} u$, $a^T a = u^T I_d^n u$ and we have

$$\big(u^T Y_t^{(u^{-1})^T x u^{-1}} u\big)_{t\geq 0} \sim \mathrm{AFF}_d(x,\bar{\alpha},B,a)$$
$$\text{where } (Y_t^y)_{t\geq 0} \sim \mathrm{AFF}_d(y,\bar{\delta},B_u,I_d^n).$$

Using the same linear transformation, we can get a potential νth-order scheme for $\mathrm{AFF}_d(x,\bar{\alpha},B,a)$ from a potential νth-order scheme for $\mathrm{AFF}_d(y,\bar{\delta},B_u,I_d^n)$ as stated below.

LEMMA 19. *If* $\hat{Y}_t^y$ *is a potential* ν*th-order scheme for* $\mathrm{AFF}_d(y,\bar{\delta},B_u,I_d^n)$, *then* $u^T \hat{Y}_t^{(u^{-1})^T x u^{-1}} u$ *is a potential* ν*th-order scheme for* $\mathrm{AFF}_d(x,\bar{\alpha},B,a)$.

PROOF. Let $f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d^+(\mathbb{R}))$. We then have $x \mapsto f(u^T x u) \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d^+(\mathbb{R}))$. Since u is fixed, there are constants C,η,E depending only on a good sequence of f such that for $t \in (0,\eta)$, $|\mathbb{E}[f(u^T \hat{Y}_t^{(u^{-1})^T x u^{-1}} u)] - \mathbb{E}[f(X_t^x)]| = |\mathbb{E}[f(u^T \hat{Y}_t^{(u^{-1})^T x u^{-1}} u)] - \mathbb{E}[f(u^T Y_t^{(u^{-1})^T x u^{-1}} u)]| \leq C t^{\nu+1}(1 + \|(u^{-1})^T x \times u^{-1}\|^E) \leq C' t^{\nu+1}(1+\|x\|^E)$, for some constant $C' > C$. $\square$

We now focus on finding a scheme for $\mathrm{AFF}_d(y,\bar{\delta},B_u,I_d^n)$, and we will construct it from the second-order scheme for $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$ obtained in (38). Since $\bar{\delta}$ is a diagonal matrix such that $\bar{\delta} - (d-1)I_d^n \in \mathcal{S}_d^+(\mathbb{R})$, we have

$$\delta_{\min} := \min_{1\leq i\leq n} \bar{\delta}_{i,i} \geq d-1.$$

We rewrite the infinitesimal generator of Y_t^y as follows:

$$\begin{aligned} L &= \mathrm{Tr}\big([\bar{\delta} + B_u(x)]D^{\mathcal{S}}\big) + 2\,\mathrm{Tr}(x D^{\mathcal{S}} I_d^n D^{\mathcal{S}}) \\ &= \underbrace{\mathrm{Tr}\big([\bar{\delta} - \delta_{\min} I_d^n + B_u(x)]D^{\mathcal{S}}\big)}_{L_{\mathrm{ODE}}} + \underbrace{\delta_{\min}\,\mathrm{Tr}(I_d^n D^{\mathcal{S}}) + 2\,\mathrm{Tr}(x D^{\mathcal{S}} I_d^n D^{\mathcal{S}})}_{L_{\mathrm{WIS}_d(x,\delta_{\min},0,I_d^n)}}. \end{aligned} \tag{40}$$

It is the sum of the infinitesimal generator of $\mathrm{WIS}_d(x,\delta_{\min},0,I_d^n)$ and of the generator of the affine ODE

$$dX_t^{\mathrm{ODE},x} = [\bar{\delta} - \delta_{\min} I_d^n + B_u(X_t^{\mathrm{ODE},x})]\,dt, \qquad X_0^{\mathrm{ODE},x} = x \in \mathcal{S}_d^+(\mathbb{R}).$$

We know by Lemma 27 that $X_t^{\mathrm{ODE},x} \in \mathcal{S}_d^+(\mathbb{R})$ for any $t \geq 0$ since assumption (4) holds for B_u and $\bar{\delta} - \delta_{\min} I_d^n \in \mathcal{S}_d^+(\mathbb{R})$. Besides, this ODE can be solved explicitly [see formula (52)]. Let $\hat{X}_t^x$ denote the potential second-order scheme for $\mathrm{WIS}_d(x,\delta_{\min},0,I_d^n)$ obtained by (38) that uses the nested second-order scheme for the CIR given in [2]. By using Proposition 15, the schemes

$$\hat{Y}_t^x = X_{t/2}^{\mathrm{ODE},\hat{X}_t^{X_{t/2}^{\mathrm{ODE},x}}} \qquad \text{or} \qquad \hat{Y}_t^x = (1-B)\hat{X}_t^{X_t^{\mathrm{ODE},x}} + B X_t^{\mathrm{ODE},\hat{X}_t^x} \tag{41}$$

are potential second-order schemes for $\mathrm{AFF}_d(x,\bar{\delta},B_u,I_d^n)$. In the numerical experiments in Section 4, we have used $X_{t/2}^{\mathrm{ODE},\hat{X}_t^{X_{t/2}^{\mathrm{ODE},x}}}$ even though the other scheme would have worked as well; it is, in fact, a computational trade-off between solving a deterministic ODE and drawing a Bernoulli variable. Thanks to Lemma 19, Proposition 14 and Theorem 13, we finally get the following result.

THEOREM 20. *The scheme defined by Lemma* 19 *and equation* (41) *is a potential second-order scheme for* $\mathrm{AFF}_d(x,\bar{\alpha},B,a)$. *In the Wishart case* (3), *we have for* $f \in \mathcal{C}_{\mathrm{pol}}^\infty(\mathcal{S}_d(\mathbb{R}))$,

$$\exists C, N_0 > 0, \forall N \geq N_0 \qquad |\mathbb{E}[f(\hat{X}_{t_N^N}^N)] - \mathbb{E}[f(X_T^x)]| \leq C/N^2.$$

3.4. *A faster second-order scheme for* $\mathrm{AFF}_d(x,\bar{\alpha},B,a)$ *when* $\bar{\alpha} - da^T a \in \mathcal{S}_d^+(\mathbb{R})$. In this section, we focus on the complexity of the discretization schemes with respect to the dimension d. Up to now, the discretization schemes that we have considered in Theorems 18 and 20 have a complexity of $O(d^4)$. Indeed, both schemes rely on the construction (38) to sample $\mathrm{WIS}_d(x,\alpha,0,I_d^n)$, which requires n Cholesky decompositions, like the exact sampling. This requires at most $O(d^4)$ operations. Here, we present a second-order scheme whose complexity is $O(d^3)$, provided that $\bar{\alpha} - da^T a \in \mathcal{S}_d^+(\mathbb{R})$ or $\alpha \geq d$ in the Wishart case. The practical relevance of such a scheme will be illustrated in Section 4.

To do so, we use the same construction as in Section 3.3, and we remark that different splitting from (40) are possible. In fact, we could have chosen instead $L = \mathrm{Tr}([\bar{\delta} - \beta I_d^n + B_u(x)]D^{\mathcal{S}}) + \beta\,\mathrm{Tr}(I_d^n D^{\mathcal{S}}) + 2\,\mathrm{Tr}(x D^{\mathcal{S}} I_d^n D^{\mathcal{S}})$ for any $\beta \in [d-1,\delta_{\min}]$: the first part is the operator of an affine ODE which is well defined on $\mathcal{S}_d^+(\mathbb{R})$ by Lemma 27 while the second part is the generator of $\mathrm{WIS}_d(x,\beta,0,I_d^n)$. When $\delta_{\min} \geq d$, which is equivalent to $\bar{\alpha} - da^T a \in \mathcal{S}_d^+(\mathbb{R})$, the following splitting obtained with $\beta = d$

$$L = \underbrace{\mathrm{Tr}([\bar{\delta} - dI_d^n + B_u(x)]D^{\mathcal{S}})}_{\tilde{L}_{\mathrm{ODE}}} + \underbrace{d\,\mathrm{Tr}(I_d^n D^{\mathcal{S}}) + 2\,\mathrm{Tr}(xD^{\mathcal{S}} I_d^n D^{\mathcal{S}})}_{L_{\mathrm{WIS}_d(x,d,0,I_d^n)}} \tag{42}$$

is really interesting. Indeed it is known from Bru [5] that Wishart processes with $\alpha \in \mathbb{N}$ can be seen as the square of an Ornstein–Uhlenbeck process on matrices and can be simulated very efficiently. More precisely, we will use the following result that is shown in Appendix D.3.

PROPOSITION 21. *Let $x \in \mathcal{S}_d^+(\mathbb{R})$ and $c \in \mathcal{M}_d(\mathbb{R})$ be such that $c^T c = x$. We have*

$$\big((c + W_t I_d^n)^T (c + W_t I_d^n), t \ge 0\big) \underset{Law}{=} \mathrm{WIS}_d(x, d, 0, I_d^n).$$

If $\hat{G}$ denote a d-by-d matrix with independent elements sampled according to (36), $\hat{X}_t^x = (c + \sqrt{t}\hat{G} I_d^n)^T (c + \sqrt{t}\hat{G} I_d^n)$ is a potential second-order scheme for $\mathrm{WIS}_d(x, d, 0, I_d^n)$.

To compute $\hat{X}_t^x$, one has to sample d^2 random variables and to make one matrix product, which requires $O(d^3)$ operations. This is faster than the scheme obtained by (38). Then we follow the same line as in Section 3.3 and set

$$d\tilde{X}_t^{\mathrm{ODE},x} = [\bar{\delta} - \delta_{\min} I_d^n + B_u(\tilde{X}_t^{\mathrm{ODE},x})]\,dt, \qquad \tilde{X}_0^{\mathrm{ODE},x} = x \in \mathcal{S}_d^+(\mathbb{R}).$$

This ODE is well defined on $\mathcal{S}_d^+(\mathbb{R})$ and can be solved explicitly. By Proposition 15,

$$\hat{Y}_t^x = \tilde{X}_{t/2}^{\mathrm{ODE},\hat{X}_t^{\tilde{X}_{t/2}^{\mathrm{ODE},x}}} \quad \text{or} \quad \hat{Y}_t^x = (1-B)\hat{X}_t^{\tilde{X}_t^{\mathrm{ODE},x}} + B\tilde{X}_t^{\mathrm{ODE},\hat{X}_t^x} \tag{43}$$

is a potential second-order scheme for $\mathrm{AFF}_d(x, \bar{\delta}, B_u, I_d^n)$ that have still an $O(d^3)$ complexity. Thanks to Lemma 19, Proposition 14 and Theorem 13, we get a similar result to Theorem 20.

THEOREM 22. *Let us assume that $\bar{\alpha} - da^T a \in \mathcal{S}_d^+(\mathbb{R})$. The scheme defined by Lemma 19 and equation (43) is a potential second-order scheme for $\mathrm{AFF}_d(x, \bar{\alpha}, B, a)$ that requires at most $O(d^3)$ operations. In the Wishart case (3), we have for $f \in \mathcal{C}_{\mathrm{pol}}^{\infty}(\mathcal{S}_d(\mathbb{R}))$,*

$$\exists C, N_0 > 0, \forall N \ge N_0 \qquad |\mathbb{E}[f(\hat{X}_{t_N^N}^N)] - \mathbb{E}[f(X_T^x)]| \le C/N^2.$$

4. Numerical results on the simulation methods. The scope of this section is to compare the different simulation methods given in this paper. We still consider a time horizon T and the regular time-grid $t_i^N = iT/N$, for $i = 0, \dots, N$. In addition, we want to compare our schemes to a standard one, and we will consider the following corrected Euler–Maruyama scheme for $\mathrm{AFF}_d(x, \bar{\alpha}, B, a)$:

$$\hat{X}^N_{t_0^N} = x,$$

$$\hat{X}^N_{t_{i+1}^N} = \hat{X}^N_{t_i^N} + \big(\bar{\alpha} + B(\hat{X}^N_{t_i^N})\big)\frac{T}{N} + \sqrt{(\hat{X}^N_{t_i^N})^+}(W_{t_{i+1}^N} - W_{t_i^N})a$$
$$+ a^T(W_{t_{i+1}^N} - W_{t_i^N})^T\sqrt{(\hat{X}^N_{t_i^N})^+}, \qquad 0 \le i \le N-1. \tag{44}$$

Here, x^+ denotes the matrix that has the same eigenvectors as x with the same eigenvalue if it is positive and a zero eigenvalue otherwise. Namely, we set $x^+ = o\,\mathrm{diag}(\lambda_1^+, \dots, \lambda_d^+)o^T$ for $x = o\,\mathrm{diag}(\lambda_1, \dots, \lambda_d)o^T$. Thus, x^+ is by construction a positive semidefinite matrix and its square root is well defined. Without this positive part, the scheme above is not well defined for any realization of W.

First, we compare the time required by the different schemes and the exact simulation. Then, we present numerical results on the convergence of the different schemes. Last, we give an application of our scheme to the Gourieroux–Sufana model in finance.

4.1. *Time comparison between the different algorithms.* In this paragraph, we compare the time required by the different schemes given in this paper. As it has already been mentioned, the complexity of the exact scheme as well as the one of the second-order scheme (given by Theorem 20) and the third-order scheme (given by Theorem 18) is in $O(d^4)$ for one time-step. To be more precise, they require $O(d^4)$ operations that mainly correspond to d Cholesky decompositions, $O(d^2)$ generations of Gaussian (or moment-matching) variables and $O(d)$ generations of noncentral chi-square distributions (or second or third-order schemes for the CIR). The time saved by the second and third-order schemes with respect to the exact scheme only comes from the generation of random variables. For example, the generation of the moment-matching variables (36) and (37) is 2.5 faster than the generation of $\mathcal{N}(0, 1)$ on our computer. The gain between the second or third-order schemes for the CIR given in Alfonsi [2] and the exact sampling of the CIR given by Glasserman [11] is much greater, but it depends on the parameters of the CIR. When the dimension d gets larger, the absolute gain in time between the discretization schemes and the exact scheme is, of course, increased. However, the relative gain instead decreases to 1, because more and more time is devoted to matrix operations and Cholesky decompositions that are the same in both cases. Let us now quickly analyze the complexity of the other schemes. The second-order scheme given by Theorem 22 (called "second-order bis" later) has a complexity in $O(d^3)$ operations for one Cholesky decomposition and matrix multiplications, with $O(d^2)$ generations of Gaussian variables. The complexity of the corrected Euler scheme is of the same kind. At each time-step, $O(d^3)$ operations are needed for matrix multiplications and for diagonalizing the matrix in order to compute the square root of its positive part. However, diagonalizing a symmetric matrix is, in practice, much longer than computing a Cholesky decomposition even though both algorithms are in $O(d^3)$. Also, one has to sample $O(d^2)$ Gaussian variables for the Brownian increments.

In Table 1, we have calculated by a Monte Carlo method one value of the characteristic function of a Wishart process. It is also known analytically thanks to (10), and we have indicated in each case the exact value. We have considered dimensions $d = 3$ and $d = 10$. We have given in each case an example where $\alpha \geq d$ and another one where $d - 1 \leq \alpha < d$. We have used the different algorithms presented in this paper: "2nd-order bis" stands for the scheme given by Theorem 22 [with the moment-matching variables (36)], "2nd order" stands for the scheme given by Theorem 20 (with (36) and the second-order scheme for the CIR given by [2]), "3rd order" stands for the scheme given by Theorem 18 (with (37) and the third-order scheme for the CIR given by [2]) and "Corrected Euler" stands for the corrected Euler–Maruyama scheme (44). For the exact scheme, we have considered both the cases with one time-step T and N time-steps T/N. Of course, the first case is sufficient to calculate an expectation that only depends on X_T, but the second case allows us to also compute pathwise expectations. For each method, we have given the value obtained and the time needed (in seconds) on our computer (3000 MHz CPU).

First, let us mention that the exact value is in each case in the confidence interval except for the corrected Euler scheme. As one can expect, the exact method with one time-step is by far the quickest method to compute an expectation that only depends on the final value. We put aside this case and focus now on the generation of the whole path. We see from Table 1 that the second and the third-order schemes require roughly the same computation time. As expected, the second-order scheme bis is much faster when it is defined (i.e., when $\alpha \geq d$). On the contrary, the Euler scheme is much slower than the second and third-order scheme. This is due to the cost of the matrix diagonalization. Let us mention that the time required by the discretization schemes is proportional to N and do not depend on the parameters when the dimension is given. On the contrary, the time needed by the exact scheme may change according to α and can increase considerably when α is close to $d - 1$. To be more precise, the exact simulation method for the CIR given by Glasserman [11] uses a rejection sampling when the degree of freedom is lower than 1, which corresponds to the case $d - 1 \leq \alpha < d$. The rejection rate can in fact be rather high, notably when the time-step gets smaller. For $N = 30$, $d = 3$ and $\alpha = 2.2$, the exact scheme is four times slower than the second-order scheme and 2.5 slower than the exact scheme with $\alpha = 3.5$.

TABLE 1

$\mathbb{E}[\exp(-\operatorname{Tr}(iv\hat{X}^N_{t^N_N}))]$ calculated by a Monte Carlo with 10^6 samples for a Wishart process with $a = I_d$, $b = 0$, $x = 10I_d$, $v = 0.09I_d$ and $T = 1$. The starred numbers are those for which the exact value is outside the 95% confidence interval, and Δ_R (resp., Δ_I) gives the two standard deviations value on the real (resp., imaginary) part

	$N = 10$			**$N = 30$**		
Schemes	**R. value**	**Im. value**	**Time**	**R. value**	**Im. value**	**Time**
	$\alpha = 3.5$, $d = 3$, $\Delta_R = 10^{-3}$, $\Delta_{Im} = 10^{-3}$, exact value R. $= -0.527090$ and Im. $= -0.228251$					
Exact (1 step)	−0.526852	−0.227962	12			
2nd-order bis	−0.526229	−0.228663	41	−0.526486	−0.229078	125
2nd order	−0.526577	−0.228923	76	−0.526574	−0.228133	229
3rd order	−0.527021	−0.227286	82	−0.527613	−0.228376	244
Exact (N steps)	−0.526963	−0.228303	123	−0.526891	−0.227729	369
Corrected Euler	−0.525627*	−0.233863*	225	−0.525638*	−0.231449*	687
	$\alpha = 2.2$, $d = 3$, $\Delta_R = 0.9 \times 10^{-3}$, $\Delta_{Im} = 1.3 \times 10^{-3}$, exact value R. $= -0.591411$ and Im. $= -0.036346$					
Exact (1 step)	−0.591579	−0.037651	12			
2nd order	−0.590444	−0.037024	77	−0.590808	−0.036487	229
3rd order	−0.591234	−0.034847	82	−0.590818	−0.036210	246
Exact (N steps)	−0.591169	−0.036618	174	−0.592145	−0.037411	920
Corrected Euler	−0.589735*	−0.042002*	223	−0.590079*	−0.039937*	680
	$\alpha = 10.5$, $d = 10$, $\Delta_R = 1.4 \times 10^{-3}$, $\Delta_{Im} = 1.3 \times 10^{-3}$, exact value R. $= 0.063960$ and Im. $= -0.063544$					
Exact (1 step)	0.062712	−0.063757	181			
2nd-order bis	0.064237	−0.063825	921	0.064573	−0.062747	2762
2nd order	0.064922	−0.064103	1431	0.063534	−0.063280	4283
3rd order	0.064620	−0.064543	1446	0.064120	−0.063122	4343
Exact (N steps)	0.063418	−0.064636	1806	0.063469	−0.064380	5408
Corrected Euler	0.068298*	−0.058491*	2312	0.061732*	−0.056882*	7113
	$\alpha = 9.2$, $d = 10$, $\Delta_R = 1.4 \times 10^{-3}$, $\Delta_{Im} = 1.4 \times 10^{-3}$, exact value R. $= -0.036064$ and Im. $= -0.093275$					
Exact (1 step)	−0.036869	−0.094156	177			
2nd order	−0.036246	−0.094196	1430	−0.035944	−0.092770	4285
3rd order	−0.035408	−0.093479	1441	−0.036277	−0.093178	4327
Exact (N steps)	−0.036478	−0.092860	1866	−0.036145	−0.093003	6385
Corrected Euler	−0.028685*	−0.094281*	2321	−0.030118*	−0.088988*	7144

Let us draw a conclusion from this time comparison between the different schemes. Obviously, we recommend the use of the exact scheme when calculating expectations that depend on one or few dates. Instead, when calculating pathwise expectations of affine processes by Monte Carlo, we would recommend the use of, in general, the second-order bis scheme when $\alpha \geq d$ and the second order (or third order for Wishart processes) when $d-1 \leq \alpha < d$.

4.2. *Numerical results on the convergence.* Now we want to illustrate the theoretical results of convergence obtained in this paper for the different schemes. To do so, we have plotted for each scheme $\mathbb{E}[\exp(-\mathrm{Tr}(iv\hat{X}^N_{t_N^N}))]$ in function of the time step T/N. This expectation is calculated by a Monte Carlo method. As for the time comparison, we illustrate the convergence for $d=3$ in Figure 1 and $d=10$ in Figure 2. Each time, we consider a case where $\alpha \geq d$ and a case where $d-1 \leq \alpha < d$, which is in general tougher. In these figures:

- scheme 1 denotes the value obtained by the exact scheme with one time-step,
- scheme 2 stands for the second-order scheme given by Theorem 20,
- scheme 3 denotes the third-order scheme given by Theorem 18,
- scheme 4 is the corrected Euler scheme (44).

Here, we have not plotted the convergence of the second-order (bis) scheme given by Theorem 22 because it would have given almost the same convergence as the other second-order scheme.

As expected, we observe in both Figures 1 and 2 convergences that fit our theoretical results. Namely, scheme 2 converges in $O(1/N^2)$ and scheme 3 converges faster in $O(1/N^3)$. In some cases, such as Figure 2, scheme 3 already matches the exact value from $N=2$. Even though it seems to converge at an $O(1/N)$ speed, the corrected Euler scheme is clearly not competitive with respect to the

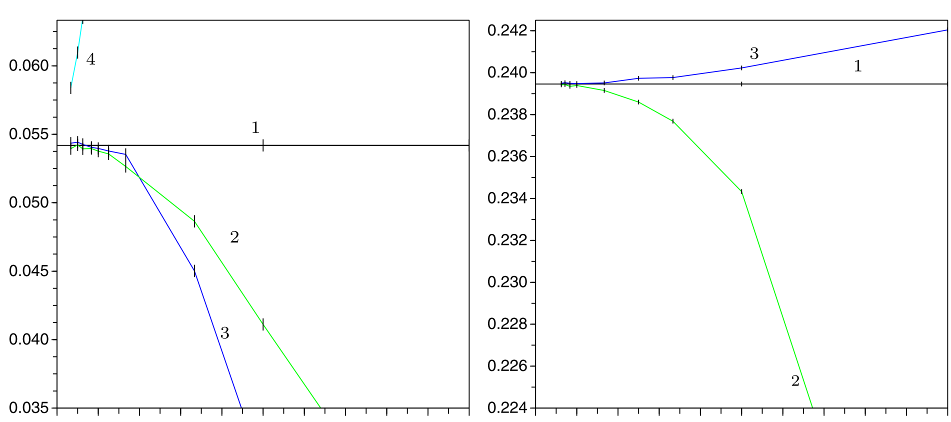

FIG. 1. $d=3$, 10^7 *Monte Carlo samples*, $T=10$. *The real value of* $\mathbb{E}[\exp(-\mathrm{Tr}(iv\hat{X}^N_{t_N^N}))]$ *in function of the time-step* T/N. *Left*: $v=0.05I_d$ *and Wishart parameters* $x=0.4I_d$, $\alpha=4.5$, $a=I_d$ *and* $b=0$. *Exact value*: 0.054277. *Right*: $v=0.2I_d+0.04q$ *and Wishart parameters* $x=0.4I_d+0.2q$, $\alpha=2.22$, $a=I_d$ *and* $b=-0.5I_d$. *Exact value*: 0.239836. *Here*, q *is the matrix defined by*: $q_{i,j}=\mathbb{1}_{i\neq j}$. *The width of each point represents the* 95% *confidence interval.*

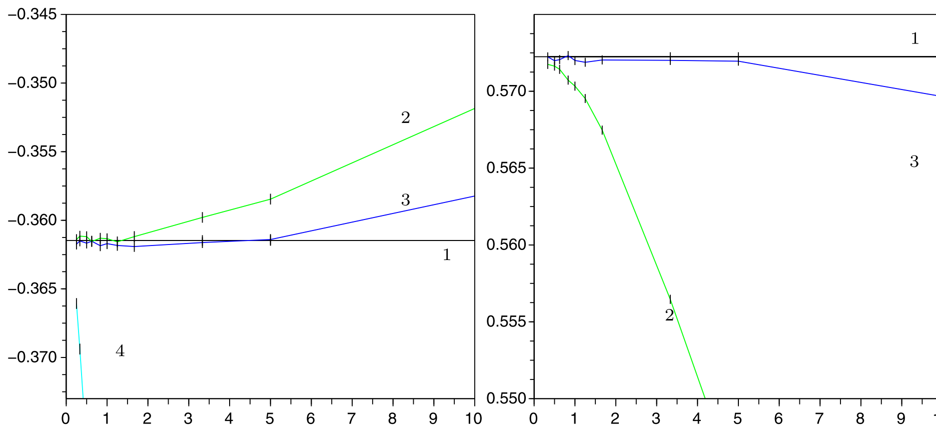

FIG. 2. $d = 10$, 10^7 *Monte Carlo samples*, $T = 10$. *Left*: *imaginary value of* $\mathbb{E}[\exp(-\mathrm{Tr}(iv\hat{X}^N_{t^N_N}))]$ *with* $v = 0.009I_d$ *in function of the time-step* T/N. *Wishart parameters*: $x = 0.4I_d$, $\alpha = 12.5$, $b = 0$ *and* $a = I_d$. *Exact value*: -0.361586. *Right*: *real value of* $\mathbb{E}[\exp(-\mathrm{Tr}(iv\hat{X}^N_{t^N_N}))]$ *with* $v = 0.009I_d$ *in function of* T/N. *Wishart parameters*: $x = 0.4I_d$, $\alpha = 9.2$, $b = -0.5I_d$ *and* $a = I_d$. *Exact value* 0.572241. *The width of each point represents the* 95% *confidence interval.*

other schemes. In the tough case $d - 1 \le \alpha \le d$, the values obtained by the Euler scheme are in fact outside the figures, and we have put the corresponding values in Table 2.

We want to conclude this section by numerically testing the convergence of our schemes when we calculate pathwise expectations. Of course, our theoretical results only bring on the weak error, but we may hope that our schemes converge also quickly when considering more intricate expectations. In Figure 3, we approximate $\mathbb{E}[\max_{0\le t\le T} \mathrm{Tr}(X^x_t)]$ with the different schemes by computing the maximum on the time-grid. The convergence seems to be roughly in $O(1/\sqrt{N})$ for all the schemes (see Figure 3, left), including the exact scheme. However, the main error seems to come from the approximation of $\max_{0\le t\le T} \mathrm{Tr}(X^x_t)$ by $\max_{0\le k\le N} \mathrm{Tr}(X^x_{t^N_k})$. In fact, we have plotted in Figure 3 (right) the difference between $\mathbb{E}[\max_{0\le k\le N} \mathrm{Tr}(\hat{X}^N_{t^N_k})]$ and $\mathbb{E}[\max_{0\le k\le N} \mathrm{Tr}(X^x_{t^N_k})]$. Then, we find convergences that are very similar to those obtained for the weak error: schemes 2

TABLE 2
Values obtained by the Euler scheme in the numerical experiments of Figures 1 and 2

N	2	4	8	10	16	30
Figure 1, right	−0.000698	0.000394	0.033193	0.111991	0.185128	0.210201
Figure 2, right	0.494752	−0.464121	0.657041	0.643042	0.637585	0.619553

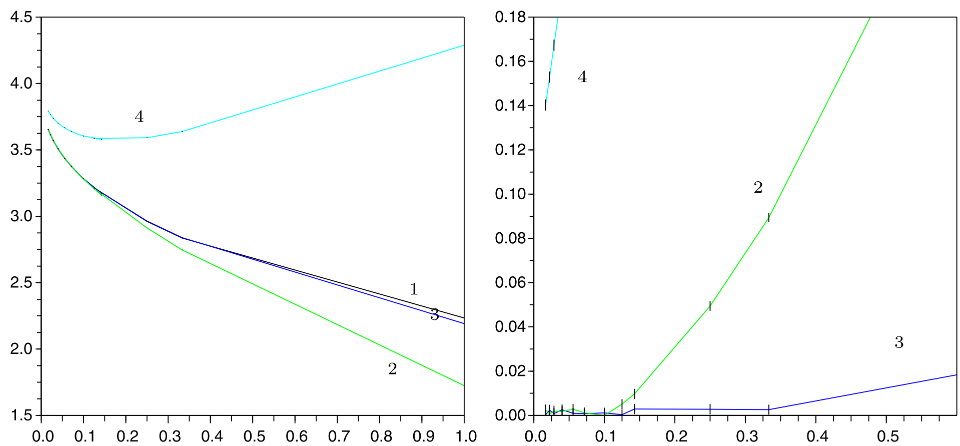

FIG. 3. $d=3$, 10^7 *Monte Carlo samples*, $T=1$. *Wishart parameters* $x=0.4I_d+0.2q$ *with* $q_{i,j}=\mathbb{1}_{i\neq j}$, $\alpha=2.2$, $b=0$ *and* $a=I_d$. *Left*: $\mathbb{E}[\max_{0\le k\le N}\operatorname{Tr}(\hat{X}^N_{t^N_k})]$. *Right*: $\mathbb{E}[\max_{0\le k\le N}\operatorname{Tr}(\hat{X}^N_{t^N_k})]-\mathbb{E}[\max_{0\le k\le N}\operatorname{Tr}(X^x_{t^N_k})]$ *in function of* T/N. *The width of each point gives the precision up to two standard deviations.*

and 3 converge at a speed which is, respectively, compatible with $O(1/N^2)$ and $O(1/N^3)$. Scheme 4 seems also to give an $O(1/N)$ convergence. It would be hasty to draw a global conclusion from this simple example. Nonetheless, the convergence of schemes 2 and 3 is really encouraging on pathwise expectations, if we put aside the problem of approximating a function of $(X^x_t, 0\le t\le T)$ by a function of $(X^x_{t^N_k}, 0\le k\le N)$.

4.3. *An application in finance to the Gourieroux and Sufana model.* In this paragraph, we want to give a possible application of our schemes in finance. More precisely, we will consider the model introduced by Gourieroux and Sufana [14]. This is a model for d risky assets $S^1_t,\dots,S^d_t$. Let $(B_t, t\ge 0)$ denote a standard Brownian motion on $\mathbb{R}^d$ that is independent from $(W_t, t\ge 0)$. Then, we consider the following dynamics for the assets:

$$(45)\qquad t\ge 0, 1\le l\le d,\qquad S^l_t=S^l_0+r\int_0^t S^l_u\,du+\int_0^t S^l_u(\sqrt{X_u}\,dB_u)_l,$$

where $X_t=X_0+\int_0^t(\alpha a^Ta+bX_u+X_ub^T)\,du+\int_0^t(\sqrt{X_u}\,dW_ua+a^T\,dW_u^T\sqrt{X_u})$ is a Wishart process. Here, $(\sqrt{X_u}\,dB_u)_l$ is simply the lth coordinates of the vector $\sqrt{X_u}\,dB_u$. We can easily check that the instantaneous quadratic covariation matrix between the log-prices of the assets is X_t. Last, r denotes the instantaneous interest rate.

To simulate both assets and the Wishart matrix, we proceed as follows. We observe that the generator of (S_t, X_t) can be written as

$$L = L^S + L^X \qquad \text{where } L^S = \sum_{i=1}^{d} r s_i \partial_{s_i} + \frac{1}{2} \sum_{i,j=1}^{d} s_i s_j x_{i,j} \partial_{s_i} \partial_{s_j},$$

and L^X is the generator of the Wishart process $\mathrm{WIS}_d(x, \alpha, b, a)$. The operator L^S is associated to the SDE $dS_t^l = rS_t^l + S_t^l(\sqrt{x}\, dB_t)_l$ that can be solved explicitly. We have indeed $S_t^l = S_0^l \exp[(r - x_{l,l}/2)t + (\sqrt{x} B_t)_l]$. Let us also remark that $\sqrt{x} B_t \underset{\text{Law}}{=} cB_t$ if we have $cc^T = x$; both are centered Gaussian vectors with the same covariance matrix. In practice, it is more efficient to use $S_t^l = S_0^l \exp[(r - x_{l,l}/2)t + (cB_t)_l]$ where c is computed with an extended Cholesky decomposition of x rather than calculating $\sqrt{x}$, which requires a diagonalization. Then we consider the scheme given by 2(a) in Proposition 15, where we take the second-order scheme for $\mathrm{WIS}_d(x, \alpha, b, a)$ and the exact scheme for L^S. This construction is known to preserve the second-order convergence. To be consistent with Section 4.2, this scheme will be denoted by scheme 2 in this paragraph. To compare this scheme with a more basic one, we consider the Euler–Maruyama scheme defined by (44) and

$$\hat{S}^{l,N}_{t_0^N} = S_0^l,$$

$$\hat{S}^{l,N}_{t_{i+1}^N} = \hat{S}^{l,N}_{t_i^N}\big(1 + rT/N + \big(\sqrt{(\hat{X}^N_{t_i^N})^+}(B_{t_{i+1}^N} - B_{t_i^N})\big)_l\big), \qquad 0 \le i \le N-1.$$

It is denoted by scheme 4 as in Section 4.2.

We have plotted in Figure 4 the price of a put option on the maximum of two risky assets ($d = 2$). The Gourieroux and Sufana model is an affine model, and the characteristic function of S_t is explicitly known (see [14]). Thus, it is possible to adapt the method proposed by Carr and Madan [6] and to calculate by numerical integration (which is possible for small dimensions) the value of this put option. We have given in Figure 4 the exact value obtained by this method. As one might have guessed, we observe a quadratic convergence for scheme 2 and a linear convergence for scheme 4. The benefit of using scheme 2 is clear since it already fits with the exact value from $N = 5$ in both cases; its convergence is really satisfactory.

5. Conclusion and prospects. Let us draw a brief summary of this paper. Thanks to a remarkable splitting of the infinitesimal generator of Wishart processes, we have been able to sample exactly any Wishart distribution. We have also proposed a third-order scheme for Wishart processes and a second-order scheme for general affine diffusions. We have confirmed these rates of convergence with numerical tests and analyzed the time complexity of each method. It comes out

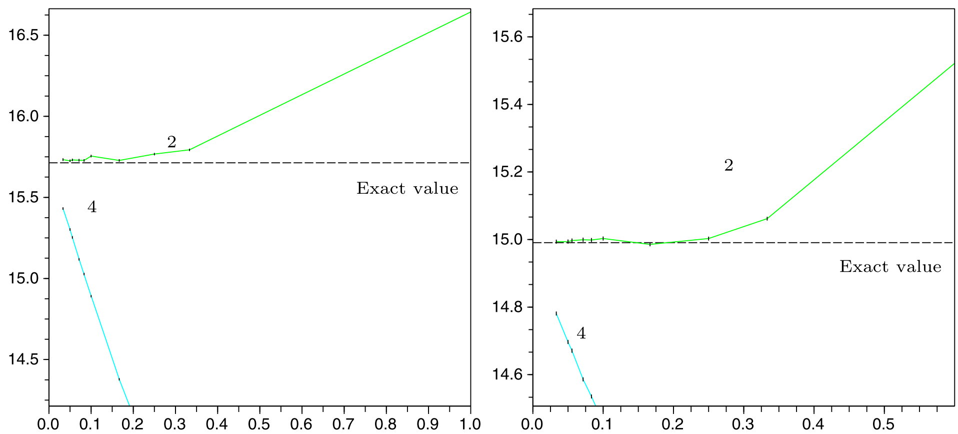

FIG. 4. $\mathbb{E}[e^{-rT}(K - \max(\hat{S}^{1,N}_{t^N_N}, \hat{S}^{2,N}_{t^N_N}))^+]$ *in function of* T/N. $d = 2$, $T = 1$, $K = 120$, $S^1_0 = S^2_0 = 100$ *and* $r = 0.02$. *Wishart parameters*: $x = 0.04I_d + 0.02q$ *with* $q_{i,j} = \mathbb{1}_{i \neq j}$, $a = 0.2I_d$, $b = 0.5I_d$ *and* $\alpha = 4.5$ (*left*), $\alpha = 1.05$ (*right*). *The width of each point gives the precision up to two standard deviations* (10^6 *Monte Carlo samples*).

that we recommend to use the exact scheme to compute expectations that depend on one (or few) times. To calculate pathwise expectations, we instead recommend generally to use discretization schemes. More precisely, the second-order scheme given by Theorem 22 has to be preferred when $\alpha \geq d$. Otherwise, we recommend to use the third-order scheme given by Theorem 18 for Wishart processes or the second-order scheme given by Theorem 20 for general affine diffusions.

Let us give now some prospects of this work. As a possible continuation of this paper, it is natural to study how it is possible to extend our schemes to affine diffusions on positive semidefinite matrices that include jumps (see Cuchiero et al. [7]). From a modeling point of view, we believe that Wishart processes could be used in a wide range of applications. In fact, they can be used as soon as one has to model dependence dynamics. Thus, we hope that the possibility of sampling such processes will stimulate different kinds of dependence models.

APPENDIX A: THE EXTENDED CHOLESKY DECOMPOSITION

LEMMA 23. *Let* $q \in \mathcal{S}^+_d(\mathbb{R})$ *be a matrix with rank* r. *Then there is a permutation matrix* p, *an invertible lower triangular matrix* $c_r \in \mathcal{G}_r(\mathbb{R})$ *and* $k_r \in \mathcal{M}_{d-r\times r}(\mathbb{R})$ *such that*

$$pqp^T = cc^T, \qquad c = \begin{pmatrix} c_r & 0 \\ k_r & 0 \end{pmatrix}.$$

The triplet (c_r, k_r, p) *is called an extended Cholesky decomposition of* q. *Besides*, $\tilde{c} = \begin{pmatrix} c_r & 0 \\ k_r & I_{d-r} \end{pmatrix} \in \mathcal{G}_d(\mathbb{R})$, *and we have*

$$q = (\tilde{c}^T p)^T I^r_d \tilde{c}^T p.$$

The proof and a numerical procedure to get such a decomposition can be found in Golub and Van Loan ([13], Algorithm 4.2.4). When $r = d$, we can take $p = I_d$, and c_r is the usual Cholesky decomposition.

APPENDIX B: PROOFS OF SECTION 1

B.1. Proof of Proposition 4. We will need in the proof the following basic lemma.

LEMMA 24. *Let $b, c \in \mathcal{S}_d(\mathbb{R})$. If either $b \in \mathcal{S}_d^+(\mathbb{R})$ or $c \in \mathcal{S}_d^+(\mathbb{R})$, then $I_d + ibc$ is invertible. In particular, if $b \in \mathcal{S}_d^{+,*}(\mathbb{R})$, $b + ic$ is invertible.*

PROOF. Let $v \in \mathcal{S}_d(\mathbb{R})$ such that $\forall s \in [0, t]$, $I_d - 2q_s v \in \mathcal{G}_d(\mathbb{R})$. As it is usual for affine diffusions, the Laplace transform can be formulated with ODE solutions. Namely, we will show that $\mathbb{E}[\exp(\mathrm{Tr}(vX_t^x))] = \exp[\phi(t, v) + \mathrm{Tr}(\psi(t, v)x)]$, where ψ and ϕ solve the following ODEs (see, e.g., Cuchiero et al. [7]):

$$\partial_t \psi(t, v) = \psi(t, v)b + b^T \psi(t, v) + 2\psi(t, v)a^T a\psi(t, v); \qquad \psi(0, v) = v,$$
$$\partial_t \phi(t, v) = \alpha \,\mathrm{Tr}(\psi(t, v)); \qquad \phi(0, v) = 0.$$

The function ψ solves an usual matrix Riccati ODE. As shown by Levin [20], ψ can be obtained explicitly by the mean of an exponential matrix, and we get

$$\psi(t, v) = \exp(tb^T)(I_d - 2vq_t)^{-1} v \exp(tb),$$

provided that $I_d - 2q_s v$ is invertible for $s \in [0, t]$, which holds by assumption. Therefore we get, for $x \in \mathcal{S}_d(\mathbb{R})$,

$$\begin{aligned}\mathrm{Tr}(\psi(t, v)x) &= \mathrm{Tr}\big((I_d - 2vq_t)^{-1} v \exp(tb)x \exp(tb^T)\big)\\ &= \mathrm{Tr}\big(v(I_d - 2q_t v)^{-1} \exp(tb)x \exp(tb^T)\big),\end{aligned}$$

since $v(I_d - 2q_t v)^{-1} = (I_d - 2vq_t)^{-1} v$. As explained by Grasselli and Tebaldi ([15], Section 4.2), ϕ can also be calculated explicitly by the mean of the exponential matrix above, and we get

$$\phi(t, v) = -\frac{\alpha}{2} \mathrm{Tr}\big(\log[(I_d - 2vq_t)\exp(tb^T)] - t\,\mathrm{Tr}(b)\big).$$

By using that $\exp(\mathrm{Tr}(\log(A))) = \det(A)$ for $A \in \mathcal{G}_d(\mathbb{R})$, we deduce then that

$$\begin{aligned}\exp(\phi(t, v)) &= \exp\left(\frac{\alpha}{2} t\,\mathrm{Tr}(b)\right)\big(\det\{(I_d - 2vq_t)\}\det\{\exp(tb^T)\}\big)^{-\alpha/2}\\ &= \frac{1}{\det(I_d - 2q_t v)^{\alpha/2}}.\end{aligned}$$

Now it remains to show that (10) indeed holds. By Itô calculus, we get that for $s \in (0, t)$,

$$\begin{aligned} &d \exp[\phi(t-s,v) + \mathrm{Tr}(\psi(t-s,v)X_s^x)] \\ &\quad = \exp[\phi(t-s,v) + \mathrm{Tr}(\psi(t-s,v)X_s^x)] \\ &\qquad \times \mathrm{Tr}[\psi(t-s,v)(\sqrt{X_s^x}\, dW_s a + a^T\, dW_s^T \sqrt{X_s^x})]. \end{aligned} \tag{46}$$

Thus, $\exp[\phi(t-s,v) + \mathrm{Tr}(\psi(t-s,v)X_s^x)]$ is a positive local martingale and therefore a supermartingale, which gives that $\mathbb{E}[\exp(\mathrm{Tr}(vX_t^x))] \le \exp[\phi(t,v) + \mathrm{Tr}(\psi(t,v)x)] < \infty$, that is, $\mathcal{D}_{b,a;t} \subset \tilde{\mathcal{D}}_{x,\alpha,b,a;t}$, where

$$\mathcal{D}_{b,a;t} := \{v \in \mathcal{S}_d(\mathbb{R}), \forall s \in [0,t], I_d - 2q_s v \in \mathcal{G}_d(\mathbb{R})\}$$

and

$$\tilde{\mathcal{D}}_{x,\alpha,b,a;t} := \{v \in \mathcal{S}_d(\mathbb{R}), \mathbb{E}[\exp(\mathrm{Tr}(vX_t^x))] < \infty\}.$$

On the other hand, when $-v \in \mathcal{S}_d^{+,*}(\mathbb{R})$, we can check that $\exp[\phi(t-s,v) + \mathrm{Tr}(\psi(t-s,v)X_s^x)] \le 1$ by observing that $\det(I_d - 2q_t v) = \det(I_d + 2\sqrt{-v} \times q_t \sqrt{-v}) \ge 1$ and $\mathrm{Tr}(v(I_d - 2q_t v)^{-1} \exp(tb)x \exp(tb^T)) = -\mathrm{Tr}(\sqrt{-v}(I_d + 2 \times \sqrt{-v} q_t \sqrt{-v})^{-1} \sqrt{-v} \exp(tb)x \exp(tb^T)) \le 0$. In that case, $\exp[\phi(t-s,v) + \mathrm{Tr}(\psi(t-s,v)X_s^x)]$ is a martingale from (46), and (10) holds.

Let us now observe that $\mathcal{D}_{b,a;t}$ is convex. In fact, we have $\det(I_d - 2q_s v) = \det(I_d - 2\sqrt{q_s} v \sqrt{q_s})$, and therefore, $\mathcal{D}_{b,a;t} = \{v \in \mathcal{S}_d(\mathbb{R}), \forall s \in [0,t], I_d - 2\sqrt{q_s} v \sqrt{q_s} \in \mathcal{S}_d^{+,*}(\mathbb{R})\}$ which is obviously convex. The Laplace transform $v \mapsto \mathbb{E}[\exp(\mathrm{Tr}(vX_t^x))]$ is an analytic function on $\mathcal{D}_{b,a;t}$ (see, e.g., [9], Lemma 10.8). The right-hand side of (10) is also analytic on $\mathcal{D}_{b,a;t}$ and coincides with the Laplace transform when $-v \in \mathcal{S}_d^{+,*}(\mathbb{R})$. Therefore, (10) holds for $v \in \mathcal{D}_{b,a;t}$ since $\mathcal{D}_{b,a;t}$ is convex. Now, we can extend to complex values of v. Indeed, the right-hand side of (10) is well defined for $v = v_R + iv_I$ with $v_R \in \mathcal{D}_{b,a;t}$, thanks to Lemma 24. Since both-hand sides are analytic functions of v, (10) holds for $v = v_R + iv_I$.

Last, we want to show that $\mathcal{D}_{b,a;t} = \tilde{\mathcal{D}}_{x,\alpha,b,a;t}$. We first consider the case $b = 0$ and assume by a way of contradiction that there is $v \in \tilde{\mathcal{D}}_{x,\alpha,0,a;t} \setminus \mathcal{D}_{0,a;t}$ for some x, α, a and $t > 0$. Let $\tilde{t} = \min\{s \in [0,t], I_d - 2q_s v \notin \mathcal{G}_d(\mathbb{R})\} \in (0,t]$. On the one hand, we have $v \notin \mathcal{D}_{0,a;\tilde{t}}$ and $v \in \mathcal{D}_{0,a;s}$ for $s \in [0,\tilde{t})$. On the other hand, we have, by Jensen's inequality

$$s \in [0,t], \qquad \exp(\alpha(t-s)\mathrm{Tr}(va^T a)) \exp(\mathrm{Tr}(vX_s^x)) \le \mathbb{E}[\exp(\mathrm{Tr}(vX_t^x)) | \mathcal{F}_s],$$

which gives $s \in [0,t] \mapsto \exp(-\alpha s \mathrm{Tr}(va^T a)) \mathbb{E}[\exp(\mathrm{Tr}(vX_s^x))]$ is nondecreasing and finite. Since (10) holds for $s < \tilde{t}$, we get that $\mathbb{E}[\exp(\mathrm{Tr}(vX_{\tilde{t}}^x))] = +\infty$, which leads to a contradiction. Let us now consider the case $b \neq 0$. From Proposition 6 (which is a consequence of the characteristic function obtained above), we have

$$\begin{aligned} v \in \tilde{\mathcal{D}}_{x,\alpha,b,a;t} \quad &\iff \quad \theta_t^T v \theta_t \in \mathcal{D}_{0,I_d^n;t} \\ &\iff \quad \forall s \in [0,t] \qquad \det(I_d - 2(s/t)q_t v) \neq 0. \end{aligned}$$

In particular, $\tilde{\mathcal{D}}_{x,\alpha,b,a;t}$ is an open set. For $v \in \mathcal{G}_d(\mathbb{R})$, we have $\det(I_d - 2(s/t)q_t v) \neq 0 \iff \det(v^{-1} - 2(s/t)q_t) \neq 0$ [resp., $\det(I_d - 2q_s v) \neq 0 \iff \det(v^{-1} - 2q_s) \neq 0$]. Since $sq_t \leq s'q_t$ (resp., $q_s \leq q_{s'}$) for $s \leq s'$, we know from Theorem 8.1.5 in [13] that the (real) eigenvalues of $v^{-1} - 2(s/t)q_t$ (resp., $v^{-1} - 2q_s$) are nonincreasing w.r.t. s. Since they are also continuous, and $v^{-1} - 2(s/t)q_t = v^{-1} - 2q_s$ for $s \in \{0, t\}$, we get that $\forall s \in [0,t], \det(v^{-1} - 2(s/t)q_t) \neq 0 \iff \forall s \in [0,t], \det(v^{-1} - 2q_s) \neq 0$ and thus $\tilde{\mathcal{D}}_{x,\alpha,b,a;t} \cap \mathcal{G}_d(\mathbb{R}) = \mathcal{D}_{b,a;t} \cap \mathcal{G}_d(\mathbb{R})$. Let $v \in \tilde{\mathcal{D}}_{x,\alpha,b,a;t}$. Since $\tilde{\mathcal{D}}_{x,\alpha,b,a;t}$ is an open set, there is $\varepsilon > 0$ such that $v \pm \varepsilon I_d \in \tilde{\mathcal{D}}_{x,\alpha,b,a;t} \cap \mathcal{G}_d(\mathbb{R})$. Since $\mathcal{D}_{b,a;t}$ is convex, $v = (v + \varepsilon I_d + v - \varepsilon I_d)/2 \in \mathcal{D}_{b,a;t}$. $\square$

B.2. Proof of Proposition 5. Once u is given, the identity in law comes directly from (13). We now give a constructive proof of the existence of u, which takes back the arguments given by Golub and Van Loan ([13], Theorem 8.7.1). Nonetheless, we explain it entirely since it gives a practical way to get u.

Let us consider $\bar{\alpha} + a^T a \in \mathcal{S}_d^+(\mathbb{R})$. From the extended Cholesky decomposition given in Lemma 23 there is a matrix $v \in \mathcal{G}_d(\mathbb{R})$ such that $v^T \bar{\alpha} v + v^T a^T a v = I_d^r$, where $r = \mathrm{Rk}(\bar{\alpha} + a^T a)$. Since $v^T \bar{\alpha} v \in \mathcal{S}_d^+(\mathbb{R})$, $v^T a^T a v \in \mathcal{S}_d^+(\mathbb{R})$ and $z^T I_d^r z = 0$ for $z \in \mathbb{R}^d$ such that $z_1 = \cdots = z_r = 0$, there are $s_1, s_2 \in \mathcal{S}_n^+(\mathbb{R})$ such that

$$v^T \bar{\alpha} v = \begin{pmatrix} s_1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad v^T a^T a v = \begin{pmatrix} s_2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Let o_2 be an orthogonal matrix such that $o_2^T s_2 o_2$ is a diagonal matrix. We assume without loss of generality that only the first n elements of this diagonal are positive: $o_2^T s_2 o_2 = \mathrm{diag}(\eta_1, \dots, \eta_n, 0, \dots, 0)$. We set $o = \begin{pmatrix} o_2 & 0 \\ 0 & I_{d-r} \end{pmatrix}$ and get $I_d^r = o^T v^T \bar{\alpha} v o + o^T v^T a^T a v o$, which gives that $o^T v^T \bar{\alpha} v o$ is a diagonal matrix. Thus, we get the desired result by taking $u = \mathrm{diag}(\sqrt{\eta_1}, \dots, \sqrt{\eta_n}, 1, \dots, 1) o^{-1} v^{-1}$.

APPENDIX C: PROOFS OF SECTION 2

C.1. Proof of Proposition 8. Let $X_t^x \sim \mathrm{WIS}_d(x, \alpha, 0, I_d^n; t)$. We will check that for any polynomial function f of the matrix elements, we have $\mathbb{E}[f(X_t^x)] = \mathbb{E}[f(X_t^{n,\dots X_t^{1,x}})]$. Let us consider a polynomial function f of degree m,

$$x \in \mathcal{S}_d(\mathbb{R}), \qquad f(x) = \sum_{\gamma \in \mathbb{N}^{d(d+1)/2}, |\gamma| \leq m} a_\gamma \bar{x}^\gamma,$$

where $|\gamma| = \sum_{1 \leq i \leq j \leq d} |\gamma_{\{i,j\}}|$ and $\bar{x}^\gamma = \prod_{1 \leq i \leq j \leq d} x_{\{i,j\}}^{\gamma_{\{i,j\}}}$. Since the operators are affine, it is easy to check that $Lf(x)$ and $L_{e_d^i} f(x)$ are also polynomial functions of degree m. We set

$$\|f\|_{\mathbb{P}} = \sum_{\gamma \in \mathbb{N}^{d(d+1)/2}, |\gamma| \le m} |a_\gamma| \quad \text{and} \quad |L| = \max_{\gamma \in \mathbb{N}^{d(d+1)/2}, |\gamma| \le m} \|L\bar{x}^\gamma\|_{\mathbb{P}},$$

so that $\|L^k f\|_{\mathbb{P}} \le |L|^k \|f\|_{\mathbb{P}}$ for any $k \in \mathbb{N}$. Therefore, the series $\sum_{k=0}^{\infty} t^k L^k f(x)/k!$ converges absolutely. By using $l+1$ times Itô's formula, we get

$$\mathbb{E}[f(X_t^x)] = \sum_{k=0}^{l} \frac{t^k}{k!} L^k f(x) + \int_0^t \frac{(t-s)^l}{l!} \mathbb{E}[L^{l+1} f(X_s^x)]\, ds.$$

Wishart processes have bounded moments since the drift and diffusion coefficients have a sublinear growth. Thus, $C = \max_{\gamma \in \mathbb{N}^{d(d+1)/2}, |\gamma| \le m} \sup_{s \in [0,t]} \mathbb{E}[|\bar{X}_s^{x^\gamma}|] < \infty$ and we obtain that $|\int_0^t \frac{(t-s)^l}{l!} \mathbb{E}[L^{l+1} f(X_s^x)]\, ds| \le C \|f\|_{\mathbb{P}} (t|L|)^{l+1}/(l+1)! \underset{l \to +\infty}{\to} 0$. Thus, we have $\mathbb{E}[f(X_t^x)] = \sum_{k=0}^{\infty} t^k L^k f(x)/k!$ and similarly we get that

$$\mathbb{E}\big[f\big(X_t^{n,\dots X_t^{1,x}}\big) | X_t^{n-1,\dots X_t^{1,x}}\big] = \sum_{k_n=0}^{+\infty} \frac{t^{k_n}}{k_n!} L_{e_d^n}^{k_n} f\big(X_t^{n-1,\dots X_t^{1,x}}\big).$$

Now, we remark that $\tilde{C} = \max_{\gamma \in \mathbb{N}^{d(d+1)/2}, |\gamma| \le m} \sup_{s \in [0,t]} \max(\mathbb{E}[|\bar{X}_t^{1,x^\gamma}|], \dots, \mathbb{E}[|\bar{X}_t^{n,\dots X_t^{1,x^\gamma}}|]) < \infty$ by using once again that Wishart processes have bounded moments. Since $\mathbb{E}[|L_{e_d^n}^{k_n} f(X_t^{n-1,\dots X_t^{1,x}})|] \le \tilde{C} \|f\|_{\mathbb{P}} |L_{e_d^n}|^{k_n}$, we can switch the expectation with the series and get (16). Then, since $L_{e_d^n}^{k_n} f(x)$ are polynomial function of degree m, we can iterate this argument and finally get (17), which gives the result.

C.2. Proof of Theorem 9. The proof is divided into two parts. First, we prove that the SDE (26) has a unique strong solution which is given by (27) and is well defined on $\mathcal{S}_d^+(\mathbb{R})$. Second, we show that its infinitesimal generator is equal to the operator $L_{e_d^1}$ defined in (18).

First step. Let us assume that $(X_t^x)_{t \ge 0}$ is a solution to (26). We use the matrix decomposition of $(x_{i,j})_{2 \le i,j \le d}$ given by (25) and set

$$(U_t)_{\{1,l+1\}} = \sum_{i=1}^{r} (c_r^{-1})_{l,i} (X_t^x)_{\{1,i+1\}}, \qquad l \in \{l, \dots, r\},$$

$$(U_t)_{\{1,1\}} = (X_t^x)_{\{1,1\}} - \sum_{l=1}^{r} \left(\sum_{i=1}^{r} (c_r^{-1})_{l,i} (X_t^x)_{\{1,i+1\}} \right)^2$$

$$= (X_t^x)_{\{1,1\}} - \sum_{l=1}^{r} \big((U_t)_{\{1,l+1\}}\big)^2.$$

We get by using Lemma 25 that

$$\begin{aligned}
&\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{pmatrix} \\
&\quad\times \begin{pmatrix} (U_t)_{\{1,1\}} + \sum_{k=1}^{r} \big((U_t)_{\{1,k+1\}}\big)^2 & \big((U_t)_{\{1,l+1\}}\big)_{1\le l\le r}^T & 0 \\ \big((U_t)_{\{1,l+1\}}\big)_{1\le l\le r} & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&\quad\times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r^T & k_r^T \\ 0 & 0 & I_{d-r-1} \end{pmatrix} \\
&\quad= \begin{pmatrix} (U_t)_{\{1,1\}} + \sum_{k=1}^{r} \big((U_t)_{\{1,k+1\}}\big)^2 & \big((U_t)_{\{1,l+1\}}\big)_{1\le l\le r}^T c_r^T & \big((U_t)_{\{1,l+1\}}\big)_{1\le l\le r}^T k_r^T \\ c_r\big((U_t)_{\{1,l+1\}}\big)_{1\le l\le r} & c_r c_r^T & c_r k_r^T \\ k_r\big((U_t)_{\{1,l+1\}}\big)_{1\le l\le r} & k_r c_r^T & 0 \end{pmatrix} \\
&\quad= X_t^x.
\end{aligned}$$

Since

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{pmatrix}$$

is invertible, $X_t^x \in \mathcal{S}_d^+(\mathbb{R})$ if, and only if

$$\begin{aligned}
&\forall z \in \mathbb{R}^d \\
&\qquad z^T \begin{pmatrix} (U_t)_{\{1,1\}} + \sum_{i=1}^{r} \big((U_t)_{\{1,i+1\}}\big)^2 & \big((U_t)_{\{1,l\}}\big)_{2\le l\le r+1} & 0 \\ \big((U_t)_{\{l,1\}}\big)_{2\le l\le r+1} & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} z \\
&\qquad = z_1^2 (U_t)_{\{1,1\}} + \sum_{i=1}^{r} \big(z_{i+1} + (U_t)_{\{1,i+1\}} z_1\big)^2 \\
&\qquad \ge 0 \quad \Longleftrightarrow \quad (U_t)_{\{1,1\}} \ge 0.
\end{aligned} \tag{47}$$

In particular, we get that $(U_0)_{\{1,1\}} = u_{\{1,1\}} \ge 0$ since $x \in \mathcal{S}_d^+(\mathbb{R})$. Now, by Itô calculus, we get from (26) that

$$d(U_t)_{\{1,l+1\}} = \sum_{i=1}^{r} \sum_{k=1}^{r} (c_r^{-1})_{l,i} (c_r)_{i,k}\, dZ_t^{k+1} = dZ_t^{l+1}$$

and

$$
\begin{aligned}
d(U_t)_{\{1,1\}} &= (\alpha - r)\,dt + 2\sqrt{(U_t)_{\{1,1\}}}\,dW_t^1 \\
&\qquad + 2\sum_{l=1}^{r}\sum_{k=1}^{r}(c_r^{-1})_{l,k}(X_t)_{\{1,k+1\}}\,dW_t^{l+1} \\
&\qquad - \sum_{l=1}^{r} 2\big((U_t)_{\{1,l+1\}}\big)\,dW_t^{l+1} \\
&= (\alpha - r)\,dt + 2\sqrt{(U_t)_{\{1,1\}}}\,dW_t^1.
\end{aligned}
$$

Thus, the solution $(X_t^x)_{t\geq 0}$ is necessarily the one given by (27) [pathwise uniqueness holds for $((U_t^u)_{\{1,l\}})_{1\leq l\leq r+1}$, and especially for the CIR diffusion $(U_t^u)_{\{1,1\}}$ since $\alpha \geq d-1 \geq r$]. Reciprocally, it is easy to check by Itô calculus that (27) solves (26).

Second step. Now we want to show that $L_{e_d^1}$ is the infinitesimal operator associated to the process $(X_t^x)_{t\geq 0}$. It is sufficient to compare the drift and the quadratic covariation of the process X_t^x with $L_{e_d^1}$. Since the drift part of $(X_t^x)_{t\geq 0}$ clearly corresponds to the first order of $L_{e_d^1}$, we study directly the quadratic part. From (26), we have for $i, j \in \{2, \dots, d\}^2$,

$$
\begin{aligned}
&d\big\langle (X_t^x)_{\{1,1\}}, (X_t^x)_{\{1,1\}}\big\rangle \\
&\quad = 4\Bigg((X_t^x)_{\{1,1\}} - \sum_{k=1}^{r}\Bigg[\sum_{l=1}^{r}(c_r^{-1})_{k,l}(X_t^x)_{\{1,l+1\}}\Bigg]^2 \\
&\qquad\qquad + \sum_{k=1}^{r}\Bigg[\sum_{l=1}^{r}(c_r^{-1})_{k,l}(X_t^x)_{\{1,l+1\}}\Bigg]^2\Bigg) \\
&\quad = 4(X_t^x)_{\{1,1\}}\,dt, \\
&d\big\langle (X_t^x)_{\{1,i\}}, (X_t^x)_{\{1,j\}}\big\rangle \\
&\quad = \sum_{k=1}^{r}(c_r)_{i-1,k}(c_r)_{j-1,k}\,dt = (cc^T)_{i-1,j-1}\,dt \\
&\quad = (X_t^x)_{\{i,j\}}\,dt, \\
&d\big\langle (X_t^x)_{\{1,1\}}, (X_t^x)_{\{1,i\}}\big\rangle \\
&\quad = 2\sum_{k=1}^{r}\sum_{l=1}^{r}(c_r)_{i-1,k}(c_r^{-1})_{k,l}(X_t^x)_{\{1,l+1\}}\,dt \\
&\quad = 2(X_t^x)_{\{1,i\}}\,dt \qquad \text{if } i \leq r+1,
\end{aligned}
$$

$$
\begin{aligned}
d\langle (X_t^x)_{\{1,1\}}, (X_t^x)_{\{1,i\}}\rangle
&= 2\sum_{k=1}^{r}\sum_{l=1}^{r}(k_r)_{i-1-r,k}(c_r^{-1})_{k,l}(X_t^x)_{\{1,l+1\}}\,dt \\
&= 2\sum_{l=1}^{r}(k_r c_r^{-1})_{i-1-r,l}(X_t^x)_{\{1,l+1\}}\,dt \\
&= 2(X_t^x)_{\{1,i\}}\,dt \qquad \text{if } i > r+1 \qquad \text{by Lemma 25.}
\end{aligned}
$$

Thus, we deduce that $L_{e_d^1}$ is the infinitesimal generator of $(X_t^x)_{t\geq 0}$.

LEMMA 25. *Let $y \in \mathcal{S}_d^+(\mathbb{R})$. We set $r = \mathrm{Rk}((y_{i,j})_{2\leq i,j\leq d})$, $y_1^r = (y_{1,i+1})_{1\leq i\leq r}$ and $y_1^{r,d} = (y_{1,i+1})_{r+1\leq i\leq d}$. We assume that there are an invertible matrix c_r and a matrix k_r defined on $\mathcal{M}_{d-r-1\times r}(\mathbb{R})$, such that*

$$(y_{i,j})_{2\leq i,j\leq d} = \begin{pmatrix} c_r & 0 \\ k_r & 0 \end{pmatrix}\begin{pmatrix} c_r^T & k_r^T \\ 0 & 0 \end{pmatrix}.$$

Then, we have $y_1^{r,d} = k_r c_r^{-1} y_1^r$.

PROOF. We set

$$
p = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{array}\right) \quad \text{and have} \quad p^{-1} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & c_r^{-1} & 0 \\ 0 & -k_r c_r^{-1} & I_{d-r-1} \end{array}\right).
$$

Since the matrix

$$
p^{-1} y (p^{-1})^T = \left(\begin{array}{c|cc} y_{1,1} & (c_r^{-1} y_1^r)^T & (y_1^{r,d} - k_r c_r^{-1} y_1^r)^T \\ \hline c_r^{-1} y_1^r & I_r & 0 \\ y_1^{r,d} - k_r c_r^{-1} y_1^r & 0 & 0 \end{array}\right)
$$

is positive semidefinite, we necessarily have $y_1^{r,d} - k_r c_r^{-1} y_1^r = 0$. □

APPENDIX D: PROOFS OF SECTION 3

D.1. Proof of Proposition 14.

LEMMA 26. *Let $(X_t^x)_{t\geq 0} \underset{Law}{\sim} \mathrm{WIS}_d(x, \alpha, b, a)$ and $v = v_R + i v_I$ such that $v_R \in \mathcal{D}_{b,a;t}$ and $v_I \in \mathcal{S}_d(\mathbb{R})$. We denote by $\phi(t, \alpha, x, v)$ the Laplace transform of X_t^x given by (10), the other parameters a, b being fixed. Then, the derivative w.r.t. $x_{\{k,l\}}$ satisfies the equality*

$$\partial_{\{k,l\}}\phi(t, \alpha, x, v) = \phi(t, \alpha + 2, x, v) p_t^{\{k,l\}}(v),$$

where $p_t^{\{k,l\}}$ *is a polynomial function of the matrix elements of degree* d *defined by*

$$p_t^{\{k,l\}}(v) = \mathrm{Tr}[v \,\mathrm{adj}(I_d - 2q_t v) m_t (e_d^{k,l} + \mathbb{1}_{k \neq l} e_d^{l,k}) m_t^T] \\ =: \sum_{\gamma \in \mathbb{N}^{d(d+1)/2}, |\gamma| \leq d} a_t^{\gamma,\{k,l\}} \bar{v}^{\gamma},$$

where

$$\bar{v}^{\gamma} = \prod_{\{i,j\}} v_{\{i,j\}}^{\gamma_{\{i,j\}}}.$$

Moreover, its coefficients are bounded uniformly in time,

$$\exists K_t > 0, \forall s \in [0,t] \qquad \max_{\gamma \in \mathbb{N}^{d(d+1)/2}, |\gamma| \leq d} (|a_s^{\gamma,\{k,l\}}|) \leq K_t.$$

PROOF. We get from (10)

$$\partial_{\{k,l\}} \phi(t,\alpha,x,v) = \frac{\mathrm{Tr}[v \,\mathrm{adj}(I_d - 2q_t v) m_t (e_d^{k,l} + \mathbb{1}_{k \neq l} e_d^{l,k}) m_t^T]}{\det(I_d - 2q_t v)} \\ \times \frac{\exp(\mathrm{Tr}[v(I_d - 2q_t v)^{-1} m_t x m_t^T])}{\det(I_d - 2q_t v)^{\alpha/2}} \\ = \phi(t, \alpha+2, x, v)\, \mathrm{Tr}[v \,\mathrm{adj}(I_d - 2q_t v) m_t (e_d^{k,l} + \mathbb{1}_{k \neq l} e_d^{l,k}) m_t^T].$$

Since $s \mapsto \|m_s\|$ and $s \mapsto \|q_s\|$ are continuous functions on $[0,t]$, we obtain the bounds on the polynomial coefficients. $\square$

PROOF OF PROPOSITION 14. Let $f \in \mathcal{C}^{\infty}_{\mathrm{pol}}(\mathcal{S}_d(\mathbb{R}))$. First, let us observe that (35) is obvious when $l = |n| = 0$. Since we have $\forall l \in \mathbb{N}, L^l f \in \mathcal{C}^{\infty}_{\mathrm{pol}}(\mathcal{S}_d(\mathbb{R}))$, and $\partial_t^l \tilde{u}(t,x) = \mathbb{E}(L^l f(X_t^x))$, it is sufficient to prove (35) only for the derivatives w.r.t. x.

We first focus on the case $|n| = 1$ and want to show that $\partial_{\{k,l\}} \tilde{u}(t,x)$ satisfies (35). The sketch of this proof is to write f as the inverse Fourier transform of its Fourier transform and then use Lemma 26. Unfortunately, f has not a priori the required integrability to do that, and we have to introduce an auxiliary function f_ρ.

Definition of the new function f_ρ. Since $\mathcal{D}_{b,a;T}$ given by (9) is an open set and $0 \in \mathcal{D}_{b,a;T}$, there is $\rho > 0$ such that $\rho I_d \in \mathcal{D}_{b,a;T}$. Let $\mu : \mathbb{R} \to \mathbb{R}$ be the function such that $\mu(x) = 0$ if $x \leq -1$ or $x \geq 0$, $\mu(x) = \exp(\frac{1}{x(x+1)})$ if $-1 < x < 0$. We have $\mu \in \mathcal{C}^{\infty}(\mathbb{R})$.

Then we consider he cutoff function $\zeta : \mathbb{R} \to \mathbb{R} \in \mathcal{C}^{\infty}(\mathbb{R})$ defined as $\forall x \in \mathbb{R}$, $\zeta(x) = \frac{\int_{-\infty}^x \mu(y)\,dy}{\int_{\mathbb{R}} \mu(y)\,dy}$. It is nondecreasing, such that $0 \leq \zeta(x) \leq 1$, $\zeta(x) = 0$ if $x \leq -1$

and $\zeta(x) = 1$ if $x \geq 0$. Besides, we have $\zeta \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{R})$ since all its derivatives have a compact support. Now, we define a $\vartheta \in \mathcal{C}^\infty_{\text{pol}}(\mathcal{S}_d(\mathbb{R}))$ as

$$\vartheta : \mathcal{S}_d(\mathbb{R}) \to \mathbb{R}, \qquad x \mapsto \prod_{i=1}^d \zeta\big(x_{\{i,i\}}\big) \prod_{i \neq j} \zeta\big(x_{\{j,j\}} x_{\{i,i\}} - x^2_{\{i,j\}}\big).$$

It is important to notice that $0 \leq \vartheta \leq 1$, $\vartheta(x) = 1$ if $x \in \mathcal{S}^+_d(\mathbb{R})$ and $\vartheta(x) = 0$ if there is $i \in \{1, \dots, d\}$ such that $x_{\{i,i\}} < -1$ or $i < j \in \{1, \dots, d\}$ such that $x^2_{\{i,j\}} > 1 + x_{\{i,i\}} x_{\{i,i\}}$. Let $\gamma \in \mathbb{N}^{d(d-1)/2}$. Since $f \in \mathcal{C}^\infty_{\text{pol}}(\mathcal{S}_d(\mathbb{R}))$, there are constants $K, E > 0$ and $K', E' > 0$ such that, $\forall x \in \mathcal{S}_d(\mathbb{R})$

$$\begin{aligned}
|\partial^\gamma(\vartheta f)(x)| &\leq K(1 + \|x\|^E) \prod_{i=1}^d \big(1_{\{|x_{\{i,i\}}| > -1\}}\big) \prod_{1 \leq i < j \leq d} \big(\mathbb{1}_{\{x^2_{\{i,j\}} \leq 1 + x_{\{i,i\}} x_{\{j,j\}}\}}\big) \\
&\leq K'\big(1 + \big\|\big(x_{\{i,i\}}\big)_{1 \leq i \leq d}\big\|^{E_1}\big) \\
&\quad \times \prod_{i=1}^d \big(1_{\{|x_{\{i,i\}}| > -1\}}\big) \prod_{1 \leq i < j \leq d} \big(\mathbb{1}_{\{x^2_{\{i,j\}} \leq 1 + x_{\{i,i\}} x_{\{j,j\}}\}}\big).
\end{aligned}$$

Here, the upper bound only involves the diagonal coefficients. We define

$$x \in \mathcal{S}_d(\mathbb{R}), \qquad f_\rho(x) := \vartheta(x) f(x) \exp(-\operatorname{Tr}(\rho x))$$

and obtain from the last inequality that f_ρ belongs to the Schwartz space of rapidly decreasing functions since $\rho > 0$. Thus, its Fourier transform also belongs to the Schwartz space and we have

$$f_\rho(x) = \frac{1}{(2\pi)^{d(d+1)/2}} \int_{\mathbb{R}^{d(d+1)/2}} \exp(-\operatorname{Tr}(ivx)) \mathcal{F}(f_\rho)(v)\, dv,$$

where

$$\mathcal{F}(f_\rho)(v) = \int_{\mathbb{R}^{d(d+1)/2}} \exp(\operatorname{Tr}(ivx)) f_\rho(x)\, dx$$

and, in particular, $f_\rho, \mathcal{F}(f_\rho) \in L^1(\mathcal{S}_d(\mathbb{R})) \cap L^\infty(\mathcal{S}_d(\mathbb{R}))$.

A new representation of $\tilde{u}(t,x)$. We have $f(x) = \exp(\rho \operatorname{Tr}(x)) f_\rho(x)$ for $x \in \mathcal{S}^+_d(\mathbb{R})$, and therefore

$$\begin{aligned}
\tilde{u}(t,x) &= \mathbb{E}[\exp(\operatorname{Tr}(\rho X^x_t)) f_\rho(X^x_t)] \\
&= \frac{1}{(2\pi)^{d(d+1)/2}} \mathbb{E}\left[\int_{\mathbb{R}^{d(d+1)/2}} \exp\big(\operatorname{Tr}[(-iv + \rho I_d) X^x_t]\big) \mathcal{F}(f_\rho)(v)\, dv\right] \\
&= \frac{1}{(2\pi)^{d(d+1)/2}} \int_{\mathbb{R}^{d(d+1)/2}} \mathbb{E}\big[\exp\big(\operatorname{Tr}[(-iv + \rho I_d) X^x_t]\big)\big] \mathcal{F}(f_\rho)(v)\, dv.
\end{aligned}$$

The last equality holds since

$$\int_{\mathbb{R}^{d(d+1)/2}} \big|\mathbb{E}\big[\exp\big(\mathrm{Tr}[(-iv+\rho I_d)X_t^x]\big)\big]\big| |\mathcal{F}(f_\rho)(v)|\,dv \\ \le \phi(t,\alpha,x,\rho I_d)\|\mathcal{F}(f_\rho)\|_1 < \infty.$$

Here we have used that $\rho I_d \in \mathcal{D}_{b,a;T}$ to get $\phi(t,\alpha,x,\rho I_d) < \infty$.

Derivation with respect to $x_{\{k,l\}}$, $k,l \in \{1,\dots,d\}$. From Lemma 26, we have by Lebesgue's theorem

$$\begin{aligned}\partial_{\{k,l\}}\tilde{u}(t,x) &= \frac{1}{(2\pi)^{d(d+1)/2}} \int_{\mathbb{R}^{d(d+1)/2}} \phi(t,\alpha+2,x,-iv+\rho I_d) \\ &\quad \times p_t^{\{k,l\}}(\rho I_d - iv)\mathcal{F}(f_\rho)(v)\,dv\end{aligned} \tag{48}$$

since $|\partial^x_{\{k,l\}}\phi(t,\alpha,x,-iv+\rho I_d)\mathcal{F}(f_\rho)(v)| \le |\phi(t,\alpha+2,x,\rho I_d)||p_t^{\{k,l\}}(\rho I_d - iv)\mathcal{F}(f_\rho)(v)|$ and $p_t^{\{k,l\}}(\rho I_d - iv)\mathcal{F}(f_\rho)(v)$ is a rapidly decreasing function.

Let $1 \le k', l' \le d$. An integration by part gives $\int_{\mathbb{R}} (\rho I_d - iv)_{\{k',l'\}} \exp(\mathrm{Tr}[x(iv - \rho I_d)])\vartheta(x) f(x)\,dx_{\{k',l'\}} = (\frac{\mathbb{1}_{k'\ne l'}}{2} + \mathbb{1}_{k'=l'}) \int_{\mathbb{R}} \exp(\mathrm{Tr}[x(iv-\rho I_d)])\partial_{\{k',l'\}}(\vartheta(x) \times f(x))\,dx_{\{k',l'\}}$, and thus

$$(\rho I_d - iv)_{\{k',l'\}}\mathcal{F}(\exp[-\rho\,\mathrm{Tr}(x)]\vartheta(x) f(x))(v) \\ = \left(\frac{\mathbb{1}_{k'\ne l'}}{2} + \mathbb{1}_{k'=l'}\right)\mathcal{F}\big(\exp[-\rho\,\mathrm{Tr}(x)]\partial_{\{k',l'\}}[\vartheta(x) f(x)]\big)(v).$$

We set $\varphi(\gamma) = \prod_{1\le k'\le l'\le d}(\frac{\mathbb{1}_{k'\ne l'}}{2} + \mathbb{1}_{k'=l'})^{\gamma_{\{k',l'\}}}$ for $\gamma \in \mathbb{N}^{d(d+1)/2}$ and get by iterating the argument that

$$\begin{aligned}&\prod_{1\le k'\le l'\le d} (\rho I_d - iv)_{\{k',l'\}}^{\gamma_{\{k',l'\}}}\mathcal{F}(f_\rho)(v) \\ &\quad = \varphi(\gamma)\mathcal{F}\big(\exp[-\rho\,\mathrm{Tr}(x)]\partial_\gamma(\vartheta \times f)(x)\big)(v).\end{aligned} \tag{49}$$

Since $p_t^{\{k,l\}}(\rho I_d - iv) = \sum_{\gamma\in\mathbb{N}^{d(d+1)/2}, |\gamma|\le d} a_t^{\gamma,\{k,l\}} \prod_{1\le k'\le l'\le d}(\rho I_d - iv)_{\{k',l'\}}^{\gamma_{\{k',l'\}}}$, we get from (48) and (49)

$$\begin{aligned}\partial_{\{k,l\}}u(t,x) &= \sum_{|\gamma|\le d} a_t^{\gamma,\{k,l\}}\varphi(\gamma)\mathbb{E}\big(\partial_\gamma(f \times \vartheta)(Y_t^x)\big) \\ &= \sum_{|\gamma|\le d} a_t^{\gamma,\{k,l\}}\varphi(\gamma)\mathbb{E}(\partial_\gamma f(Y_t^x)),\end{aligned} \tag{50}$$

where $(Y_t^x)_{t\ge 0} \underset{\text{Law}}{\sim} \mathrm{WIS}_d(x,\alpha+2,b,a)$. Here we have used that $\partial_\gamma(\vartheta \times f)(y) = \partial_\gamma f(y)$ for $y \in \mathcal{S}_d^+(\mathbb{R})$. From Lemma 26 $(a_t^{\gamma,\{k,l\}})_{\gamma\in\mathbb{N}^{d(d+1)/2}, |\gamma|\le d}$ is bounded for $t \in [0,T]$, and we get (35) when $|n| = 1$ since $\partial_\gamma f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d(\mathbb{R}))$. Thanks to (50), a derivative of order $|n|$, can be seen as a (bounded) linear combination of derivatives of order $|n|-1$, and we easily get (35) by an induction on $|n|$.

It remains to check that we have indeed $\partial_t \tilde{u}(t,x) = Lu(t,x)$. Let $t,h>0$. By the Markov property, we have $\tilde{u}(t+h,x) = \mathbb{E}[\tilde{u}(t,X_h^x)]$. From (35) and Itô's formula, we get $[\tilde{u}(t+h,x)-u(t,x)]/h \underset{h\to 0^+}{\to} Lu(t,x)$. $\square$

LEMMA 27. *Let $\alpha, x \in \mathcal{S}_d^+(\mathbb{R})$, $B \in \mathcal{L}(\mathcal{S}_d^+(\mathbb{R}))$ that satisfies (4), and $x(t)$ be the solution of the ODE*

$$x(t) = x + \int_0^t \big(\alpha + B(x(s))\big)\,ds. \tag{51}$$

Then we have $x(t) \in \mathcal{S}_d^+(\mathbb{R})$ for $t \geq 0$.

PROOF. The ODE (51) is affine and has unique solution on $\mathcal{S}_d^+(\mathbb{R})$ which is given by

$$t \geq 0, \qquad x(t) = \exp(tB)(x) + \int_0^t \exp(sB)(\alpha)\,ds, \tag{52}$$

where $\forall t \in \mathbb{R}^+, \forall x \in \mathcal{S}_d(\mathbb{R}), \exp(tB)(x) = \sum_{k=0}^{\infty} \frac{t^k B^k(x)}{k!}$, $B^k(x) = \underbrace{B \circ \cdots \circ B}_{k \text{ times}}(x)$ such that $B^0(x) = x$.

We first assume that $\alpha, x \in \mathcal{S}_d^{+,*}(\mathbb{R})$ and consider $\tau = \inf\{t \geq 0, x(t) \notin \mathcal{S}_d^+(\mathbb{R})\}$, with the convention $\inf \varnothing = +\infty$. We have $\tau > 0$. Let us assume by a way of contradiction that $\tau < \infty$. Then $x(\tau)$ cannot be invertible and there is $y \in \mathcal{S}_d^+(\mathbb{R})$ such that $y \neq 0$ and $\mathrm{Tr}(yx(\tau)) = 0$. From (52) and (4), we get

$$\mathrm{Tr}(x'(\tau)y) = \mathrm{Tr}\big([B(x(\tau)) + \alpha]y\big) > 0,$$

since α is positive definite. Therefore, there is $\epsilon \in (0,\tau)$ such that $\mathrm{Tr}(yx(\tau-\epsilon)) < 0$. Let us now recall that $z \in \mathcal{S}_d^+(\mathbb{R}) \iff \forall y \in \mathcal{S}_d^+(\mathbb{R}), \mathrm{Tr}(yz) \geq 0$. Thus, $x(\tau-\epsilon) \notin \mathcal{S}_d^+(\mathbb{R})$, which contradicts the definition of τ.

In the general case $\alpha, x \in \mathcal{S}_d^+(\mathbb{R})$, we observe that the solution (52) is continuous w.r.t. x and α, and thus $\forall t \geq 0, x(t) \in \mathcal{S}_d^+(\mathbb{R})$ since $\mathcal{S}_d^+(\mathbb{R})$ is a closed set. $\square$

D.2. Proof of Proposition 17. First, let us check that $\theta_t \in \mathcal{G}_d(\mathbb{R})$ is well defined, such that $q_t/t = \theta_t I_d^n \theta_t^T$ and satisfies

$$\exists K, \eta > 0, \forall t \in (0,\eta) \qquad \max(\|\theta_t\|, \|\theta_t\|^{-1}) \leq K. \tag{53}$$

When $n = d$, q_t/t is definite positive as a convex combination of definite positive matrices and the usual Cholesky decomposition is well defined. Moreover, (53) holds since q_t/t goes to $a^T a$ which is invertible when $t \to 0^+$. When $n < d$, we have assumed, in addition, that b and $a^T a$ commute. Therefore,

$q_t = a^T a(\int_0^t \exp(sb)\exp(sb^T)\,ds/t)$. Since $a^T a$ and $(\int_0^t \exp(sb)\exp(sb^T)\,ds/t)$ are positive semidefinite matrices that commute, we have

$$q_t = \sqrt{\frac{1}{t}\int_0^t \exp(sb)\exp(sb^T)\,ds}\, a^T a \sqrt{\frac{1}{t}\int_0^t \exp(sb)\exp(sb^T)\,ds}.$$

Once again, $\frac{1}{t}\int_0^t \exp(sb)\exp(sb^T)\,ds$ is definite positive as a convex combination of definite positive matrices and we get that $\theta_t = \sqrt{\frac{1}{t}\int_0^t \exp(sb)\exp(sb^T)\,ds} \times p^{-1}\begin{pmatrix} c_n & 0 \\ k_n & I_{d-n}\end{pmatrix} \in \mathcal{G}_d(\mathbb{R})$ satisfies $q_t/t = \theta_t I_d^n \theta_t^T$ by Lemma 23. Similarly, (53) holds since $p^{-1}\begin{pmatrix} c_n & 0 \\ k_n & I_{d-n}\end{pmatrix}$ does not depend on t and $\sqrt{\frac{1}{t}\int_0^t \exp(sb)\exp(sb^T)\,ds}$ goes to I_d when $t \to 0^+$.

Let $f \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d^+(\mathbb{R}))$. Let $X_t^x \sim \mathrm{WIS}_d(x, \alpha, b, a; t)$. Since the exact scheme is a potential νth-order scheme, there are constants $C, E, \eta > 0$ depending only on a good sequence of f such that

$$\forall t \in (0,\eta) \qquad \left|\mathbb{E}[f(X_t^x)] - \sum_{k=0}^{\nu}\frac{t^k}{k!}L^k f(x)\right| \le Ct^{\nu+1}(1+\|x\|^E). \tag{54}$$

On the other hand, we have from Proposition 6,

$$\begin{aligned} &\mathbb{E}[f(\hat{X}_t^x)] - \mathbb{E}[f(X_t^x)] \\ &\quad = \mathbb{E}\big[f\big(\theta_t \hat{Y}_t^{\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T}\theta_t^T\big)\big] - \mathbb{E}\big[f\big(\theta_t Y_t^{\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T}\theta_t^T\big)\big]. \end{aligned} \tag{55}$$

Let us introduce $f_{\theta_t}(y) := f(\theta_t y \theta_t^T) \in \mathcal{C}^\infty_{\mathrm{pol}}(\mathcal{S}_d^+(\mathbb{R}))$. By the chain rule, we have $\partial_{\{i,j\}} f_{\theta_t}(y) = \mathrm{Tr}[\theta_t(e_d^{i,j} + \mathbb{1}_{i\neq j} e_d^{j,i})\theta_t^T \partial f(\theta_t y \theta_t^T)]$, where $(\partial f(x))_{k,l} = (\mathbb{1}_{k=l} + \frac{1}{2}\mathbb{1}_{k\neq l})\partial_{\{k,l\}} f(x)$ and $e_d^{i,j} = (\mathbb{1}_{k=i,l=j})_{1\le k,l\le d}$. From (53), we see that there is a good sequence $(C_\gamma, e_\gamma)_{\gamma\in\mathbb{N}^{d(d+1)/2}}$ that can be obtained from a good sequence of f such that

$$\forall t \in (0,\eta), \forall y \in \mathcal{S}_d^+(\mathbb{R}) \qquad |\partial_\gamma f_{\theta_t}(y)| \le C_\gamma(1+\|y\|^{e_\gamma}).$$

Therefore, we get that there are constants still denoted by $C, E, \eta > 0$ such that

$$\begin{aligned} &\forall t \in (0,\eta) \\ &\qquad \big|\mathbb{E}\big[f\big(\theta_t \hat{Y}_t^{\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T}\theta_t^T\big)\big] - \mathbb{E}\big[f\big(\theta_t Y_t^{\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T}\theta_t^T\big)\big]\big| \\ &\qquad\quad \le Ct^{\nu+1}\big(1 + \|\theta_t^{-1} m_t x m_t^T (\theta_t^{-1})^T\|^E\big). \end{aligned} \tag{56}$$

From (53), we get that there is a constant $K' > 0$ such that $\|\theta_t^{-1} m_t x \times m_t^T (\theta_t^{-1})^T\|^E \le K'\|x\|^E$ for $t \in (0,\eta)$. Thus, we get the result by gathering (54), (55) and (56).

D.3. Proof of Proposition 21. We have, by using Itô calculus, $dX_t^x = (c + W_t I_d^n)^T dW_t I_d^n + I_d^n dW_t^T (c + W_t I_d^n) + d I_d^n dt$. By using Lemma 2, the quadratic covariation of $(X_t^x)_{i,j}$ and $(X_t^x)_{m,n}$ is given by $d\langle (X_t^x)_{i,j}, (X_t^x)_{m,n}\rangle = (X_t^x)_{i,m}(I_d^n)_{j,n} + (X_t^x)_{i,n}(I_d^n)_{j,m} + (X_t^x)_{j,m}(I_d^n)_{i,n} + (X_t^x)_{j,n}(I_d^n)_{i,m}$. Therefore, $(X_t^x)_{t\geq 0}$ solves the same martingale problem as $\mathrm{WIS}_d(x, d, 0, I_d^n)$, which is known to have a unique solution from Cuchiero et al. [7].

Let us now show that $\hat{X}_t^x$ is a potential second-order scheme. We can see $c + \sqrt{t}\hat{G} I_d^n$ as the Ninomiya–Victoir scheme with moment-matching variables (see [2], Theorem 1.18) associated to $\frac{1}{2}\sum_{i=1}^d \sum_{j=1}^n \partial_{i,j}^2$ on $\mathcal{M}_d(\mathbb{R})$. Let $f \in \mathcal{C}_{\mathrm{pol}}^{\infty}(\mathcal{S}_d^+(\mathbb{R}))$. Then, $x \in \mathcal{M}_d(\mathbb{R}) \mapsto f(x^T x) \in \mathcal{C}_{\mathrm{pol}}^{\infty}(\mathcal{M}_d(\mathbb{R}))$ and there are constants $C, E, \eta > 0$ depending only on a good sequence of f such that

$$\forall t \in (0, \eta)$$
$$\bigl|\mathbb{E}[f((c+\sqrt{t}\hat{G}I_d^n)^T(c+\sqrt{t}\hat{G}I_d^n))] - \mathbb{E}[f((c+W_t I_d^n)^T(c+W_t I_d^n))]\bigr| \leq C t^{\nu+1}(1+\|c\|^E).$$

Let us now observe that the Frobenius norm of c is $\sqrt{\mathrm{Tr}(c^T c)} = \sqrt{\mathrm{Tr}(x)} \leq \sqrt{d + \mathrm{Tr}(x^2)} \leq \sqrt{d} + \sqrt{\mathrm{Tr}(x^2)}$. Therefore, for any norm, there is a constant $K > 0$ such that $\|c\| \leq K(1 + \|x\|)$, which gives the result.

REFERENCES

[1] ALFONSI, A. (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl*. **11** 355–384. MR2186814

[2] ALFONSI, A. (2010). High order discretization schemes for the CIR process: Application to affine term structure and Heston models. *Math*. *Comp*. **79** 209–237. MR2552224

[3] BENABID, A., BENSUSAN, H. and EL KAROUI, N. (2010). Wishart stochastic volatility: Asymptotic smile and numerical framework. Preprint.

[4] BRU, M. F. (1987). Thèse 3ème cycle. Résistence d'Escherichie coli aux antibiotiques. Sensibilités des analyses en composantes principales aux perturbations Browniennes et simulation. Ph.D. thesis, Univ. Paris Nord.

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

[6] CARR, P. and MADAN, A. (1999). Option pricing and the fast Fourier transform. *J*. *Comput*. *Finance* **2** 61–73.

[7] CUCHIERO, C., FILIPOVIĆ, D., MAYERHOFER, E. and TEICHMANN, J. (2011). Affine processes on positive semidefinite matrices. *Ann*. *Appl*. *Probab*. **21** 397–463. MR2807963

[8] DA FONSECA, J., GRASSELLI, M. and TEBALDI, C. (2008). Option pricing when correlations are stochastic: An analytical framework. *Review of Derivatives Research* **10** 151–180.

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

[10] GAUTHIER, P. and POSSAMAI, D. (2009). Efficient simulation of the Wishart model. SSRN eLibrary.

[11] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering*: *Stochastic Modelling and Applied Probability*. *Applications of Mathematics* (*New York*) **53**. Springer, New York. MR1999614

[12] GLESER, L. J. (1976). A canonical representation for the noncentral Wishart distribution useful for simulation. *J*. *Amer*. *Statist*. *Assoc*. **71** 690–695. MR0433717

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

[14] GOURIEROUX, C. and SUFANA, R. (2003). Wishart quadratic term structure models. Working paper.

[15] GRASSELLI, M. and TEBALDI, C. (2008). Solvable affine term structure models. *Math. Finance* **18** 135–153. MR2380943

[16] HESTON, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* **6** 327–343.

[17] KABE, D. G. (1964). A note on the Bartlett decomposition of a Wishart matrix. *J. Roy. Statist. Soc. Ser. B* **26** 270–273. MR0171358

[18] KSHIRSAGAR, A. M. (1959). Bartlett decomposition and Wishart distribution. *Ann. Math. Statist.* **30** 239–241. MR0103552

[19] KUSUOKA, S. (2004). Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. In *Advances in Mathematical Economics. Vol.* 6. *Adv. Math. Econ.* **6** 69–83. Springer, Tokyo. MR2079333

[20] LEVIN, J. J. (1959). On the matrix Riccati equation. *Proc. Amer. Math. Soc.* **10** 519–524. MR0108628

[21] LYONS, T. and VICTOIR, N. (2004). Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460** 169–198. Stochastic analysis with applications to mathematical finance. MR2052260

[22] MAYERHOFER, E., PFAFFEL, O. and STELZER, R. (2011). On strong solutions for positive definite jump diffusions. *Stochastic Process. Appl.* **121** 2072–2086. MR2819242

[23] NINOMIYA, M. and NINOMIYA, S. (2009). A new higher-order weak approximation scheme for stochastic differential equations and the Runge–Kutta method. *Finance Stoch.* **13** 415–443. MR2519839

[24] NINOMIYA, S. and VICTOIR, N. (2008). Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance* **15** 107–121. MR2409419

[25] ODELL, P. L. and FEIVESON, A. H. (1966). A numerical procedure to generate a sample covariance matrix. *J. Amer. Statist. Assoc.* **61** 199–203. MR0192635

[26] SMITH, W. B. and HOCKING, R. R. (1972). Algorithm as 53: Wishart variate generator. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **21** 341–345.

[27] STRANG, G. (1968). On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.* **5** 506–517. MR0235754

[28] TALAY, D. and TUBARO, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8** 483–509. MR1091544

[29] TEICHMANN, J. (2010). Covariance matrix valued affine processes structure and numerics. Talk at the Bachelier Seminar in Paris.

CERMICS
UNIVERSITÉ PARIS-EST
PROJECT TEAM MATHFI ENPC-INRIA-UMLV
ECOLE DES PONTS
6-8 AVENUE BLAISE PASCAL
77455 MARNE LA VALLÉE
FRANCE
E-MAIL: ahdidaa@cermics.enpc.fr
alfonsi@cermics.enpc.fr