PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 27, Fasc. 1 (2007), pp. 3–43

POSITIVE-DEFINITE MATRIX PROCESSES OF FINITE VARIATION

BY

OLE EILER BARNDORFF-NIELSEN (ÅRHUS, DENMARK)
AND ROBERT STELZER (GARCHING, GERMANY)

Dedicated to the memory of Kazimierz Urbanik

Abstract. Processes of finite variation, which take values in the positive semidefinite matrices and are representable as the sum of an integral with respect to time and one with respect to an extended Poisson random measure, are considered. For such processes we derive conditions for the square root (and the r-th power with $0 < r < 1$) to be of finite variation and obtain integral representations of the square root. Our discussion is based on a variant of the Itô formula for finite variation processes.

Moreover, Ornstein–Uhlenbeck type processes taking values in the positive semidefinite matrices are introduced and their probabilistic properties are studied.

2000 AMS Mathematics Subject Classification: 15A52, 60H20, 60G52.

Key words and phrases: Finite variation process, Itô formula, Lévy process, matrix subordinator, Ornstein–Uhlenbeck type process, positive definite square root.

1. INTRODUCTION

The theory of self-decomposability, as developed by Lévy, Urbanik, Sato, Jurek and Mason, and others, has turned out to be of substantial interest for stochastic modelling in finance, turbulence and other fields. See, for instance, Barndorff-Nielsen (1998a), Barndorff-Nielsen and Shephard (2001) and Barndorff-Nielsen and Schmiegel (2004), where (positive) Lévy driven processes of Ornstein–Uhlenbeck type have a key role.

The focus of the present paper is on stochastic differential equation representations of square roots of positive definite matrix processes of Lévy or Ornstein–Uhlenbeck type. Such representations are, in particular, of interest in connection with the general theory of multipower variation, cf. Barndorff-Nielsen, Graversen, Jacod, Podolskij and Shephard (2006) and Barndorff-Nielsen, Graversen, Jacod and Shephard (2006).

In the present literature matrix-valued stochastic processes are not commonly used to model multivariate phenomena (see, for instance, the short discussion on multivariate stochastic volatility models at the end of Section 4). Our introduction of positive-definite Ornstein–Uhlenbeck processes and the discussion of the representations of square (and other) roots shows that matrix-valued models of considerable generality can be defined in a natural way and univariate results can very often be generalized by using notions and results from matrix analysis. Furthermore, several results of general interest regarding matrix-valued processes (semimartingales) and matrix analysis are obtained, as we proceed.

This paper is organized as follows. Section 2 establishes some notation, and in Section 3 we present a convenient version of Itô's formula for processes of finite variation. In Section 4 we introduce positive definite processes of Ornstein–Uhlenbeck type (OU processes), using the concept of matrix subordinators discussed by Barndorff-Nielsen and Pérez-Abreu (2007). The question of establishing tractable stochastic differential equations for roots of positive definite matrix processes is then addressed in Section 5, and in Section 6 the results are applied to the case of OU processes.

2. NOTATION

Throughout this paper we write $\mathbb{R}^+$ for the positive real numbers including zero and we denote the set of real $m \times n$ matrices by $M_{m,n}(\mathbb{R})$. If $m = n$, we simply write $M_n(\mathbb{R})$ and denote the group of invertible $n \times n$ matrices by $GL_n(\mathbb{R})$, the linear subspace of symmetric matrices by $\mathbb{S}_n$, the (closed) positive semidefinite cone by $\mathbb{S}_n^+$, and the open (in $\mathbb{S}_n$) positive definite cone by $\mathbb{S}_n^{++}$. Moreover, I_n stands for the $n \times n$ identity matrix and $\sigma(A)$ for the spectrum (the set of all eigenvalues) of a matrix $A \in M_n(\mathbb{R})$. The natural ordering on the symmetric $n \times n$ matrices will be denoted by $\leqslant$, i.e. for $A, B \in \mathbb{S}_n$ we have $A \leqslant B$ if and only if $B - A \in \mathbb{S}_n^+$. The tensor (Kronecker) product of two matrices A, B is written as $A \otimes B$. By vec we denote the well-known vectorisation operator that maps the $n \times n$ matrices to $\mathbb{R}^{n^2}$ by stacking the columns of the matrices below one another. Finally, A^* is the adjoint of a matrix $A \in M_n(\mathbb{R})$.

For a matrix A we denote by A_{ij} the element in the i-th row and j-th column and this notation is extended to processes in a natural way.

Regarding all random variables and processes we assume that they are defined on a given appropriate filtered probability space $(\Omega, \mathscr{F}, P, (\mathscr{F}_t))$ satisfying the usual hypotheses. With random functions we usually do not state the dependence on $\omega \in \Omega$ explicitly.

Furthermore, we employ an intuitive notation with respect to the integration with matrix-valued integrators. Let $A_t \in M_{m,n}$, $L_t \in M_{n,r}$ and $B_t \in M_{r,s}$ be three processes. Then we denote by $\int A_t \, dL_t \, B_t$ the matrix C in $M_{m,s}(\mathbb{R})$ which has the ij-th element $C_{ij} = \sum_{k=1}^{n} \sum_{l=1}^{r} \int a_{ik} b_{lj} dL_{kl}$. Moreover, we always denote by $\int_a^b$ with $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R}$ the integral over the half-open interval $(a, b]$ for notational convenience. If $b = \infty$, the integral is understood to be over (a, b).

3. ITÔ FORMULAE FOR FINITE VARIATION PROCESSES IN OPEN SETS

In this section we provide a univariate and a multivariate version of the Itô formula from stochastic analysis, which is especially suitable for the purposes of this paper. Actually, our version is a consequence of standard results, but not given in the usual references. Closely related versions for processes taking values in $\mathbb{R}^d$ instead of an open subset C can be found in Protter (2004), Theorem II.31, or Rogers and Williams (2000), pp. 28–29, for example.

As we are analysing stochastic processes in general open subsets C of $\mathbb{R}^d$, $M_d(\mathbb{R})$ or $\mathbb{S}_d$, we need an appropriate assumption that the process stays within C and does not hit the boundary, since this causes problems in general. To describe "good" behaviour we thus introduce "local boundedness within C". If C is the whole space, it is the same as "local boundedness".

DEFINITION 3.1. Let $(V, \|\cdot\|_V)$ be either $\mathbb{R}^d$, $M_d(\mathbb{R})$ or $\mathbb{S}_d$ with $d \in \mathbb{N}$ and equipped with the norm $\|\cdot\|_V$, let $a \in V$ and let $(X_t)_{t \in \mathbb{R}^+}$ be a V-valued stochastic process. We say that X_t is *locally bounded away from* a if there exists a sequence of stopping times $(T_n)_{n \in \mathbb{N}}$ increasing to infinity almost surely and a real sequence $(d_n)_{n \in \mathbb{N}}$ with $d_n > 0$ for all $n \in \mathbb{N}$ such that $\|X_t - a\|_V \geqslant d_n$ for all $0 \leqslant t < T_n$.

Likewise, we say that for some open set $C \subset V$ the process X_t is *locally bounded within* C if there exists a sequence of stopping times $(T_n)_{n \in \mathbb{N}}$ increasing to infinity almost surely and a sequence of compact convex subsets $D_n \subset C$ with $D_n \subset D_{n+1}$ for all $n \in \mathbb{N}$ such that $X_t \in D_n$ for all $0 \leqslant t < T_n$.

Obviously, if a process is locally bounded away from some a or is locally bounded within some C in one norm, then the same holds for all other norms. We will see in the following that these definitions play a central role for our Itô formulae and that they hold for many processes.

PROPOSITION 3.2 (Univariate Itô formula for processes of finite variation). *Let $(X_t)_{t \in \mathbb{R}^+}$ be a càdlàg process of finite variation (thus a semimartingale) with associated jump measure μ_X on $(\mathbb{R}^+ \times \mathbb{R}\backslash\{0\}, \mathscr{B}(\mathbb{R}^+ \times \mathbb{R}\backslash\{0\}))$* (see e.g. Jacod and Shiryaev (2003), Proposition II.1.16) *and let $f: C \to \mathbb{R}$ be continuously differentiable, where C is some open interval $C = (a, b)$ with $a, b \in \mathbb{R} \cup \{\pm\infty\}$, $a < b$. Assume that $(X_t)_{t \in \mathbb{R}^+}$ is locally bounded within C. Then the process X_t as well as its left limit process X_{t-} take values in C at all times $t \in \mathbb{R}^+$, the integral*

$$\int_0^t \int_{\mathbb{R}\backslash\{0\}} (f(X_{s-}+x) - f(X_{s-}))\, \mu_X(ds, dx)$$

exists a.s. for all $t \in \mathbb{R}$ *and*

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-})\,dX_s^c + \int_0^t \int_{\mathbb{R}\setminus\{0\}} (f(X_{s-}+x) - f(X_{s-}))\,\mu_X(ds, dx),$$

where $X_t^c = X_t - \int_0^t \int_{\mathbb{R}\setminus\{0\}} x\mu_X(ds, dx)$ *is the continuous part of* X.

(Strictly speaking, $f(X_{s-}+x)$ is not defined for all $x \in \mathbb{R}$, as f is only defined on C. But our assumptions assure that μ_X is concentrated on those x for which $X_{s-}+x \in C$. Therefore we can simply continue f arbitrarily outside of C.)

Proof. As X_t is locally bounded within C, the process X_t cannot get arbitrarily close to the boundary of C in finite time, and hence X_t and X_{t-} are in C at all times $t \in \mathbb{R}^+$.

Obviously,

$$\int_0^t \int_{\mathbb{R}\setminus\{0\}} (f(X_{s-}+x) - f(X_{s-}))\,\mu_X(ds, dx) = \sum_{0<s\leqslant t} \Delta f(X_s).$$

That X_t is locally bounded within C implies the existence of compact intervals $D_n \subset C$ such that $X_t \in D_n$ for all $t \in [0, T_n)$ for some sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times increasing to infinity a.s. However, f' is bounded on D_n, say by c_n, and the mean value theorem gives

$$\Delta f(X_s) = f(X_s) - f(X_{s-}) = f'(\zeta_s)(X_s - X_{s-}) = f'(\zeta_s)\Delta X_s \quad \text{with } \zeta_s \in D_n.$$

Therefore,

$$\int_0^t \int_{\mathbb{R}\setminus\{0\}} |f(X_{s-}+x) - f(X_{s-})|\,\mu_X(ds, dx)$$
$$= \sum_{0<s\leqslant t} |\Delta f(X_s)| \leqslant c_n \sum_{0<s\leqslant t} |\Delta X_s| \quad \text{for all } t \in [0, T_n),$$

which is finite due to the finite variation of X_t. Thus the almost sure existence of the integral is shown.

The standard Itô formula (see Bichteler (2002), Theorem 3.9.1 together with Proposition 3.10.10, for an appropriate version) gives

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-})\,dX_s + \int_0^t \int_{\mathbb{R}\setminus\{0\}} (f(X_{s-}+x) - f(X_{s-}) - f'(X_{s-})x)\,\mu_X(ds, dx)$$

on observing that, since X_t is a finite variation process, we can move from a twice continuously differentiable f to an only once continuously differentiable one, as in Protter (2004), Theorem II.31. Noting further that

$$\int_0^t f'(X_{s-})\,dX_s = \int_0^t f'(X_{s-})\,dX_s^c + \int_0^t \int_{\mathbb{R}\setminus\{0\}} f'(X_{s-})\,x\mu_X(ds, dx)$$

and that the integral $\int_0^t \int_{\mathbb{R}\setminus\{0\}} (f(X_{s-}+x)-f(X_{s-}))\,\mu_X(ds, dx)$ exists, we obtain

$$f(X_t) = f(X_0)+\int_0^t f'(X_{s-})\,dX_s^c+\int_0^t \int_{\mathbb{R}\setminus\{0\}} (f(X_{s-}+x)-f(X_{s-}))\,\mu_X(ds, dx). \blacksquare$$

Remark 3.3. (a) The assumption that X_t remains locally bounded within C ensures that $f'(X_t)$ is locally bounded. This reflects the boundedness of the derivative needed in the proof of Protter (2004), Theorem I.54, which is a special case of the above result.

(b) It is straightforward to see that X_t is locally bounded within $C = (a, b)$ if and only if X_t is in C at all times and locally bounded away from both a and b, where for $a = -\infty$ or $b = \infty$ this has to be understood as meaning locally bounded. Recall in this context that any finite variation process is locally bounded.

In the multivariate version we use the notion of (total) differentials, sometimes also called *Fréchet differentials* (see Rudin (1976), Chapter 9, or Bhatia (1997), Section X.4, for an overview focusing on the matrix case), rather than partial derivatives for notational convenience. Recall, however, that a function is continuously differentiable if and only if all partial derivatives exist and are continuous, and that the derivative, which is a linear operator, simply has the partial derivatives as entries. The derivative of a function f at a point x is denoted by $Df(x)$. In particular, we have the following multivariate version of Proposition 3.2. We state it only for processes in $\mathbb{R}^d$, but it should be obvious that $\mathbb{R}^d$ can be replaced by $M_d(\mathbb{R})$ or $\mathbb{S}_d$.

PROPOSITION 3.4 (Multivariate Itô formula for processes of finite variation). *Let $(X_t)_{t\in\mathbb{R}^+}$ be a càdlàg $\mathbb{R}^d$-valued process of finite variation (thus a semimartingale) with associated jump measure μ_X on $(\mathbb{R}^+\times\mathbb{R}^d\setminus\{0\}, \mathcal{B}(\mathbb{R}^+\times\mathbb{R}^d\setminus\{0\}))$ and let $f: C\to\mathbb{R}^m$ be continuously differentiable, where $C\subseteq\mathbb{R}^d$ is an open set. Assume that the process $(X_t)_{t\in\mathbb{R}^+}$ is locally bounded within C. Then the process X_t as well as its left limit process X_{t-} take values in C at all times $t\in\mathbb{R}^+$, the integral*

$$\int_0^t \int_{\mathbb{R}^d\setminus\{0\}} (f(X_{s-}+x)-f(X_{s-}))\,\mu_X(ds, dx)$$

exists a.s. for all $t\in\mathbb{R}$ and

$$f(X_t) = f(X_0)+\int_0^t Df(X_{s-})\,dX_s^c+\int_0^t \int_{\mathbb{R}^d\setminus\{0\}} (f(X_{s-}+x)-f(X_{s-}))\,\mu_X(ds, dx),$$

where $X_t^c = X_t-\int_0^t\int_{\mathbb{R}^d\setminus\{0\}} x\mu_X(ds, dx)$ is the continuous part of X.

Proof. The proof is a mere multivariate rephrasing of the one for Proposition 3.2 using an appropriate general multidimensional version of Itô's formula (e.g. Bichteler (2002), Proposition 3.10.10, Métivier (1982), Theorem 27.2, or Protter (2004), Theorem 7.33) and standard results from multivariate calculus. ∎

4. POSITIVE SEMIDEFINITE MATRIX PROCESSES OF OU TYPE

In this section we briefly review one-dimensional processes of Ornstein–Uhlenbeck (OU) type (cf. Applebaum (2004), Cont and Tankov (2004) or Barndorff-Nielsen and Shephard (2001, 2007), among many others) and then introduce Ornstein–Uhlenbeck processes taking values in the positive semidefinite matrices. For the necessary background on Lévy processes see Protter (2004), Section I.4, or Sato (1999).

In univariate financial modelling, it has become popular in recent years to specify the variance $(\sigma_t^2)_{t\in\mathbf{R}^+}$ as an Ornstein–Uhlenbeck process (see in particular the works of Barndorff-Nielsen and Shephard). We assume given a Lévy process $(L_t)_{t\in\mathbf{R}^+}$ and consider the SDE

$$d\sigma_t^2 = -\lambda\sigma_{t-}^2\,dt + dL_t \tag{4.1}$$

with some $\lambda\in\mathbf{R}$. The solution can be shown to be

$$\sigma_t^2 = e^{-\lambda t}\sigma_0^2 + \int_0^t e^{-\lambda(t-s)}\,dL_s \tag{4.2}$$

and is referred to as an OU (type) process. Note that for univariate OU type processes one often applies a time transformation on the Lévy process and then has $dL_{\lambda s}$ instead of dL_s above, but this is not possible in the multivariate case below. Provided the Lévy process L_t is a subordinator (a.s. non-decreasing Lévy process), the solution σ_t^2 is positive and thus can be used as a variance process. After extending the Lévy process to one, $(L_t)_{t\in\mathbf{R}}$, living on the whole real line in the usual way, one can show that (4.1) has a unique stationary solution given by

$$\sigma_t^2 = \int_{-\infty}^t e^{-\lambda(t-s)}\,dL_s$$

provided $\lambda > 0$ and the Lévy process has a finite logarithmic moment, i.e. $E(\log^+(L_t)) < \infty$.

There is a vast literature concerning the extension of OU processes to $\mathbf{R}^d$-valued processes (for instance, Sato and Yamazato (1984), Chojnowska-Michalik (1987) or Jurek and Mason (1993)). By identifying $M_d(\mathbf{R})$ with $\mathbf{R}^{d^2}$ one immediately obtains matrix-valued processes. So for a given Lévy process $(L_t)_{t\in\mathbf{R}}$ with values in $M_d(\mathbf{R})$ and a linear operator $\mathrm{A}: M_d(\mathbf{R})\to M_d(\mathbf{R})$ we call some solution to the SDE

$$dX_t = \mathrm{A}X_{t-}\,dt + dL_t \tag{4.3}$$

a *(matrix-valued) process of Ornstein–Uhlenbeck type.*

As in the univariate case one can show that for some given initial value X_0 the solution is unique and given by

$$X_t = e^{\mathrm{A}t}X_0 + \int_0^t e^{\mathrm{A}(t-s)}\,dL_s. \tag{4.4}$$

Provided $E(\log^+ \|L_t\|) < \infty$ and $\sigma(\mathrm{A}) \in (-\infty, 0) + i\mathbf{R}$, there exists a unique stationary solution given by

$$X_t = \int_{-\infty}^{t} e^{\mathrm{A}(t-s)}\, dL_s.$$

In order to obtain positive semidefinite Ornstein–Uhlenbeck processes we need to consider matrix subordinators as driving Lévy processes. An $M_d(\mathbf{R})$-valued Lévy process L_t is called "matrix subordinator" if $L_t - L_s \in S_d^+$ a.s. for all $t \geqslant s$; see Barndorff-Nielsen and Pérez-Abreu (2002, 2007), Rocha-Arteaga (2006) and the references therein for further details.

PROPOSITION 4.1. *Let L_t be a matrix subordinator, assume that the linear operator A satisfies $\exp(\mathrm{A}t)(S_d^+) \subseteq S_d^+$ for all $t \in \mathbf{R}^+$ and let $X_0 \in S_d^+$. Then the Ornstein–Uhlenbeck process $(X_t)_{t\in\mathbf{R}^+}$ with initial value X_0 satisfying* (4.3) *takes only values in S_d^+.*

If $E(\log^+ \|L_t\|) < \infty$ and $\sigma(\mathrm{A}) \in (-\infty, 0) + i\mathbf{R}$, then the unique stationary solution $(X_t)_{t\in\mathbf{R}}$ to (4.3) *takes values in S_d^+ only.*

Proof. The first term $e^{\mathrm{A}t} X_0$ in (4.4) is obviously positive semidefinite for all $t \in \mathbf{R}^+$ due to the assumption on A. Approximating the integral $\int_0^t e^{\mathrm{A}(t-s)}\, dL_s$ by sums in the usual way shows that also the second term is positive semidefinite, since all approximating sums are in S_d^+ due to the assumption on A and the S_d^+-increasingness of a Lévy subordinator.

The very same argument implies the positive semidefiniteness of the unique stationary solution. ∎

An important question arises now, namely, which linear operators A can one actually take to obtain both a unique stationary solution and ensure positive semidefiniteness. The condition $\exp(\mathrm{A}t)(S_d^+) \subseteq S_d^+$ means that for all $t \in \mathbf{R}^+$ the exponential operator $\exp(\mathrm{A}t)$ has to preserve positive definiteness. So one seems to need to know first which linear operators on $M_d(\mathbf{R})$ preserve positive definiteness. This problem has been studied for a long time in linear algebra in connection with the general topic "Linear Preserver Problems" (see, for instance, the overview articles Pierce et al. (1992) and Li and Pierce (2001)). We have the following:

PROPOSITION 4.2. *Let $\mathrm{A}: S_d \to S_d$ be a linear operator. Then $\mathrm{A}(S_d^+) = S_d^+$ if and only if there exists a matrix $B \in GL_d(\mathbf{R})$ such that A can be represented as $X \mapsto BXB^*$.*

Proof. This was initially proved in Schneider (1965). A more general proof in a Hilbert space context may be found in Li et al. (2003). ∎

Remark 4.3. No explicit characterization of the linear operators mapping S_d^+ into S_d^+, i.e. $\mathrm{A}(S_d^+) \subseteq S_d^+$, is known for general dimension d.

Naturally, all linear maps on S_d can be extended to mappings on M_d. From this linear algebraic result we obtain the following result, introducing the linear operators preserving positive semidefiniteness which we shall employ.

PROPOSITION 4.4. *Assume the operator* $\mathbb{A}: M_d(\mathbb{R}) \to M_d(\mathbb{R})$ *is representable as* $X \mapsto AX + XA^*$ *for some* $A \in M_d(\mathbb{R})$. *Then* $e^{\mathbb{A}t}$ *has the representation* $X \mapsto e^{At}Xe^{A^*t}$ *and* $e^{\mathbb{A}t}(\mathbb{S}_d^+) = \mathbb{S}_d^+$ *for all* $t \in \mathbb{R}$.

Proof. The equality $e^{\mathbb{A}t}X = e^{At}Xe^{A^*t}$ for all $X \in M_d(\mathbb{R})$ follows from Horn and Johnson (1991), pp. 255 and 440, and $e^{\mathbb{A}t}(\mathbb{S}_d^+) = \mathbb{S}_d^+$ for all $t \in \mathbb{R}$ is then implied by Proposition 4.2, since e^B is invertible for any matrix $B \in M_d(\mathbb{R})$. ∎

Note the close relation of this kind of operators to Kronecker sums and the so-called "Lyapunov equation" (see Horn and Johnson (1991), Chapter 4). For a linear operator $\mathbb{A}$ of the type specified in Proposition 4.4 formula (4.3) becomes

$$dX_t = (AX_{t-} + X_{t-}A^*)\,dt + dL_t \tag{4.5}$$

and the solution is

$$X_t = e^{At}X_0 e^{A^*t} + \int_0^t e^{A(t-s)}\,dL_s\, e^{A^*(t-s)}. \tag{4.6}$$

Confer also Horn and Johnson (1991), p. 440, for a related deterministic differential equation.

Using the vec transformation and Horn and Johnson (1991), Theorem 4.4.5, we see that $\sigma(\mathbb{A}) = \sigma(A) + \sigma(A)$, where the addition of two sets $A, B \subseteq \mathbb{R}$ is defined by $A + B = \{a + b : a \in A, b \in B\}$. Thus

THEOREM 4.5. *Let* $(L_t)_{t\in\mathbb{R}}$ *be a matrix subordinator with* $E(\log^+ \|L_t\|) < \infty$ *and* $A \in M_d(\mathbb{R})$ *such that* $\sigma(A) \subset (-\infty, 0) + i\mathbb{R}$. *Then the stochastic differential equation of Ornstein–Uhlenbeck type*

$$dX_t = (AX_{t-} + X_{t-}A^*)\,dt + dL_t$$

has a unique stationary solution

$$X_t = \int_{-\infty}^t e^{A(t-s)}\,dL_s\, e^{A^*(t-s)}$$

or, in vectorial representation,

$$\mathrm{vec}(X_t) = \int_{-\infty}^t \exp\left((I_d \otimes A + A \otimes I_d)(t-s)\right) d\,\mathrm{vec}(L_s).$$

Moreover, $X_t \in \mathbb{S}_d^+$ *for all* $t \in \mathbb{R}$.

Recall from Barndorff-Nielsen and Pérez-Abreu (2007) that any matrix subordinator $(L_t)_{t\in\mathbb{R}}$ has paths of finite variation and can be represented as

$$L_t = \gamma t + \int_0^t \int_{\mathbb{S}_d^+\setminus\{0\}} x\mu(ds, dx), \tag{4.7}$$

where $\gamma \in S_d^+$ is a deterministic drift and $\mu(ds, dx)$ an extended Poisson random measure on $\mathbb{R}^+ \times S_d^+$ (regarding the definitions of random measures and the integration theory with respect to them we refer to Jacod and Shiryaev (2003), Section II.1). Observe in particular that the integral exists without compensating. Moreover, the expectation of μ factorises, i.e. $E(\mu(ds, dx)) = \text{Leb}(ds)\, \nu(dx)$, Leb denoting the Lebesgue measure and ν the Lévy measure of L_t. The above equation (4.7) can be restated in a differential manner as

$$dL_t = \gamma dt + \int_{S_d^+ \setminus \{0\}} x\mu(dt, dx). \tag{4.8}$$

The obvious extension of this to a Lévy process $(L_t)_{t \in \mathbb{R}}$ having been started in the infinite past gives another representation of the above stationary OU process.

PROPOSITION 4.6. *The positive semidefinite Ornstein–Uhlenbeck process X_t as given in Theorem 4.5 can equivalently be represented as*

$$\begin{aligned} X_t &= \int_{-\infty}^{t} \int_{S_d^+ \setminus \{0\}} e^{A(t-s)} x e^{A^*(t-s)} \mu(ds, dx) + \int_{-\infty}^{t} e^{A(t-s)} \gamma e^{A^*(t-s)} ds \\ &= \int_{-\infty}^{t} \int_{S_d^+ \setminus \{0\}} e^{A(t-s)} x e^{A^*(t-s)} \mu(ds, dx) - B^{-1} \gamma, \end{aligned}$$

where B^{-1} is the inverse of the linear operator $B\colon M_d(\mathbb{R}) \to M_d(\mathbb{R})$, $X \mapsto AX + XA^*$ which can be represented as $\text{vec}^{-1} \circ \big((I_d \otimes A) + (A \otimes I_d)\big)^{-1} \circ \text{vec}$.

Proof. The invertibility of B and the positive semidefiniteness of $-B^{-1}\gamma$ follow immediately from the standard theory on the Lyapunov equations (Horn and Johnson (1991), Theorems 2.2.3 and 4.4.7). Now only the second equality remains to be shown, but this is immediate as

$$-B^{-1} \frac{d}{ds} e^{A(t-s)} \gamma e^{A^*(t-s)} = e^{A(t-s)} \gamma e^{A^*(t-s)} \quad \text{and} \quad \lim_{s \to -\infty} e^{A(t-s)} = 0. \ \blacksquare$$

The next proposition provides a characterization of the stationary distribution. To this end observe that $\text{tr}(XY)$ (with $X, Y \in M_d(\mathbb{R})$ and tr denoting the usual trace functional) defines a scalar product on $M_d(\mathbb{R})$. Moreover, the vec operator is a Hilbert space isometry between $M_d(\mathbb{R})$ equipped with this scalar product and $\mathbb{R}^{d^2}$ with the usual Euclidean scalar product. This, in particular, implies that the driving Lévy process L_t has characteristic function (cf. also Barndorff-Nielsen and Pérez-Abreu (2007))

$$\mu_{L_t}(Z) = \exp\Big(it\, \text{tr}(\gamma Z) + t \int_{S_d^+ \setminus \{0\}} \big(\exp\big(i\, \text{tr}(XZ)\big) - 1\big)\, \nu(dX)\Big). \tag{4.9}$$

Proposition 4.7. *The stationary distribution of the matrix Ornstein–Uhlenbeck process X_t is infinitely divisible with characteristic function*

$$\hat{\mu}_X(Z) = \exp\Big(i\,\mathrm{tr}(\gamma_X Z) + \int_{S_d^+\setminus\{0\}} \big(\exp(i\,\mathrm{tr}(YZ)) - 1\big)\,\nu_X(dY)\Big), \tag{4.10}$$

where $\gamma_X = -B^{-1}\gamma$ with B defined as in Proposition 4.6 and

$$\nu_X(E) = \int_0^\infty \int_{S_d^+\setminus\{0\}} I_E(e^{As} x e^{A^*s})\,\nu(dx)\,ds$$

for all Borel sets E in $S_d^+\setminus\{0\}$.

Assume that the driving Lévy process is square-integrable. Then the second order moment structure is given by

$$E(X_t) = \gamma_X - B^{-1} \int_{S_d^+\setminus\{0\}} y\nu(dy) = -B^{-1} E(L_1), \tag{4.11}$$

$$\begin{aligned} &\mathrm{Var}(\mathrm{vec}(X_t)) \\ &= \int_0^\infty \exp((A\otimes I_d + I_d\otimes A)\,t)\,\mathrm{Var}(\mathrm{vec}(L_1))\exp((A^*\otimes I_d + I_d\otimes A^*)\,t)\,dt \\ &= -\mathfrak{B}^{-1}\,\mathrm{Var}(\mathrm{vec}(L_1)), \end{aligned} \tag{4.12}$$

$$\mathrm{Cov}(\mathrm{vec}(X_{t+h}), \mathrm{vec}(X_t)) = \exp((A\otimes I_d + I_d\otimes A)\,h)\,\mathrm{Var}(\mathrm{vec}(X_t)), \tag{4.13}$$

where $t\in\mathbb{R}$ and $h\in\mathbb{R}^+$, and

$$\mathfrak{B}: M_{d^2}(\mathbb{R}) \to M_{d^2}(\mathbb{R}), \qquad X \mapsto (A\otimes I_d + I_d\otimes A)X + X(A^*\otimes I_d + I_d\otimes A^*).$$

The linear operator $\mathfrak{B}$ can be represented as

$$\mathrm{vec}^{-1}\circ\big((I_{d^2}\otimes(A\otimes I_d + I_d\otimes A)) + ((A\otimes I_d + I_d\otimes A)\otimes I_{d^2})\big)\circ\mathrm{vec}.$$

We used the vec operator above, as this clarifies the order of the elements of the (co)variance matrix.

Proof. The characteristic function is standard, cf. Barndorff-Nielsen, Pedersen and Sato (2001), p. 178, for instance. Regarding (4.11) a general result for infinitely divisible distributions implies that

$$E(X_t) = \gamma_X + \int_{S_d^+} y\nu_X(dy).$$

Using the explicit representation for ν_X and evaluating the integral as in the proof of the last proposition immediately establish (4.11). The proof of the first equality in (4.12) and of (4.13) is standard, see e.g. Marquardt and Stelzer (2006), Proposition 3.13, and the second equality in (4.12) follows by an explicit integration as before. ∎

Remark 4.8. In the existing literature for $\mathbb{R}^d$-valued processes only the analogue to the first equality in (4.12) is stated and an identity is given that becomes

$$-\operatorname{Var}(\operatorname{vec}(L_1)) = (A \otimes I_d + I_d \otimes A)\operatorname{Var}(\operatorname{vec}(X_t)) + \operatorname{Var}(\operatorname{vec}(X_t))(A^* \otimes I_d + I_d \otimes A^*)$$

in our case. That identity is, of course, equivalent to our second equality in (4.2), but usually obtained by a very different approach (cf. Arató (1982), for instance). Our version involving $\mathfrak{B}^{-1}$ stresses that the variance can be calculated by solving a standard linear equation and fits in nicely, as inverse operators of this type appear in many of our results.

Moreover, conditions ensuring that the stationary OU type process X_t is almost surely strictly positive definite can be obtained.

THEOREM 4.9. *If* $\gamma \in S_d^{++}$ *or* $\nu(S_d^{++}) > 0$, *then the stationary distribution* P_X *of* X_t *is concentrated on* S_d^{++}, *i.e.* $P_X(S_d^{++}) = 1$.

Proof. From Proposition 4.6 and its proof we have $X_t \geqslant -B^{-1}\gamma$. In the case $\gamma \in S_d^{++}$ this proves the theorem immediately, as then $-B^{-1}\gamma$ is strictly positive definite due to Horn and Johnson (1991), Theorem 2.2.3.

Assume now that $\nu(S_d^{++}) > 0$. From Proposition 4.6 we know that

$$X_0 \geqslant \sum_{-\infty < s \leqslant 0} e^{-As} \Delta(L_s) e^{-A^*s} \stackrel{d}{=} \sum_{0 \leqslant s < \infty} e^{As} \Delta(L_s) e^{A^*s}.$$

Since $Z \mapsto e^{As} Z e^{A^*s}$ preserves positive definiteness for all $s \in \boldsymbol{R}$, it is obviously sufficient to show that $(L_s)_{s \in \boldsymbol{R}^+}$ has at least one jump that is positive definite. Choose now $\varepsilon > 0$ such that

$$\nu(S_d^{++} \cap \{x \in S_d^+ : \|x\| \geqslant \varepsilon\}) > 0.$$

Then the process

$$L_{\varepsilon,s} := \sum_{0 \leqslant s \leqslant t} 1_{\{x \in S_d^+ : \|x\| \geqslant \varepsilon\}}(\Delta L_s)\, \Delta L_s$$

is a Lévy process with Lévy measure

$$\nu_\varepsilon(\cdot) = \nu(\cdot \cap \{x \in S_d^+ : \|x\| \geqslant \varepsilon\}),$$

where we denoted by $1_M(\cdot)$ the indicator function of a set M. The process L_ε is obviously a compound Poisson process and the probability that a jump of L_ε is in $S_d^+ \backslash S_d^{++}$ is given by

$$q := \nu_\varepsilon(S_d^+ \backslash S_d^{++}) / \nu_\varepsilon(S_d^+) < 1.$$

As the individual jump sizes and the jump times are independent and $(L_{\varepsilon,s})_{s \in \boldsymbol{R}^+}$ has a.s. infinitely many jumps in $\boldsymbol{R}^+$, this implies that with probability zero all jumps of $(L_{\varepsilon,s})_{s \in \boldsymbol{R}^+}$ are in $S_d^+ \backslash S_d^{++}$. In other words, $(L_{\varepsilon,s})_{s \in \boldsymbol{R}^+}$ and thus $(L_s)_{s \in \boldsymbol{R}^+}$ has a.s. at least one jump in S_d^{++}. ∎

The positive-definite Ornstein–Uhlenbeck processes introduced above can be used as a multivariate stochastic volatility model in finance, as an extension of the one-dimensional approach proposed in Barndorff-Nielsen and Shephard (2001). A different kind of generalization has been discussed by Hubalek and Nicolato (2005) and Lindberg (2005), who have specified different multivariate stochastic volatility models using factor models, where the individual factors are univariate positive Ornstein–Uhlenbeck type processes. The d-dimensional volatility model of Hubalek and Nicolato is of the form $\Sigma_t^2 = AS_t A^*$, where S_t is an Ornstein–Uhlenbeck process in S_m^+ (actually only on the diagonal matrices) and $A \in M_{d,m}(\mathbb{R})$. The results for the roots of positive definite processes which we obtain in Section 5 are with a minor obvious adaptation immediately applicable to processes of this type. Another proposal put forth in Gourieroux et al. (2004) specifies a $d \times d$ volatility process V_t as a sum $V_t = \sum_{i=1}^{K} x_{t,i} x_{t,i}^*$ with the processes $x_{t,i}$ being i.i.d. Gaussian Ornstein–Uhlenbeck processes in $\mathbb{R}^d$ and $K \in \mathbb{N}$. These processes are referred to as *Wishart autoregressive processes*, as the distribution of V_t is the Wishart distribution (see also Bru (1991)). This specification is not amenable to the type of SDE representations of the root processes that we shall discuss in Section 5, under a general set-up, and in Section 6 for positive definite OU processes. Note also, in this connection, that the Wishart law is not infinitely divisible, hence, in particular, not self-decomposable (see Lévy (1948)).

In stochastic volatility models the integrated variance process is of particular interest (see e.g. Barndorff-Nielsen and Shephard (2001, 2003)). The same reasoning as in the univariate case (Barndorff-Nielsen (1998b)) leads to the following explicit result for the integrated variance of a positive definite Ornstein–Uhlenbeck stochastic volatility process:

PROPOSITION 4.10. *Let X_t be a positive semidefinite Ornstein–Uhlenbeck process with initial value $X_0 \in S_d^+$ and driven by the Lévy process L_t. Then the integrated Ornstein–Uhlenbeck process X_t^+ is given by*

$$X_t^+ := \int_0^t X_t \, dt = B^{-1}(X_t - X_0 - L_t)$$

for $t \in \mathbb{R}^+$, where B is the linear operator defined in Proposition 4.6.

5. ROOTS OF POSITIVE SEMIDEFINITE PROCESSES

In this section we obtain stochastic representations of general roots of processes in $\mathbb{R}^+$ and later on of the square root of stochastic processes taking values in S_d^+. Recall that every positive semidefinite matrix A has a unique positive semidefinite square root $A^{1/2}$ defined by functional calculus (see, for instance, Horn and Johnson (1990, 1991) for a comprehensive introduction).

The interest in such representations comes, in particular, from the theoretical works on the properties of multipower variation; see Barndorff-Nielsen, Graversen, Jacod, Podolskij and Shephard (2006), for instance. In that paper the limit theorems are obtained under the hypothesis that the square root of the covariance matrix process is a semimartingale of a special type. Moreover, in many cases the additional assumption is needed that it takes values in the strictly positive definite matrices, as this ensures that the covariance matrix process is of the same type (and vice versa). However, as there are no formulae given relating the characteristics of the covariance matrix process with those of its square root, we shall derive the relations explicitly and discuss whether the invertibility assumption is indeed always necessary. Under the invertibility assumption Itô's lemma is the key tool, but as we see later on we can move away from this prerequisite. On the other hand, we restrict ourselves to the study of processes of finite variation. The reasons are that the processes we intend to apply our results to are naturally of finite variation and that in the infinite variation case it seems impossible to obtain results for processes that may reach the boundary $\partial S_d^+ = S_d^+ \backslash S_d^{++}$. As a consequence, all our "stochastic" integrals coming up can actually be computed pathwise as Lebesgue–Stieltjes integrals.

In the following we start by analysing univariate processes, where we study general r-th powers and then move on to multivariate processes.

5.1. The univariate case. Now we shall first present the univariate case, as it involves no advanced matrix analysis, but allows one to understand the behaviour of root processes. Due to the applications we have in mind, we state the following results for finite variation processes, whose discontinuous part is of the special form

$$\int_0^t \int_{\mathbb{R}^+\backslash\{0\}} g(s-, x)\,\mu(ds, dx)$$

with some extended Poisson random measure μ on $\mathbb{R}^+\backslash\{0\}$ (in the sense of Jacod and Shiryaev (2003), Definition 1.20). Moreover,

$$g(s, x) = g(\omega, s, x)\colon \Omega \times \mathbb{R}^+ \times \mathbb{R}^+\backslash\{0\} \to \mathbb{R}^+\backslash\{0\}$$

is a (random) function that is $\mathcal{F}_s \times \mathcal{B}(\mathbb{R}^+)$ measurable in (ω, x) and càdlàg in s. For such a process the jump measure is

$$\mu_X(ds, dx) = \mu\big(ds, g^{-1}(s-, \cdot)(dx)\big),$$

where $g^{-1}(s-, \cdot)$ is to be understood as taking the preimage of the set dx with respect to the map $\mathbb{R}^+\backslash\{0\} \to \mathbb{R}^+\backslash\{0\}$, $x \mapsto g(s-, x)$. We frequently refer to the dependence on $\omega \in \Omega$ in the following, but keep suppressing it in the notation.

THEOREM 5.1. *Let $(X_t)_{t\in\mathbb{R}^+}$ be a given adapted càdlàg process which takes values in $\mathbb{R}^+\backslash\{0\}$, is locally bounded away from zero and can be represented as*

$$dX_t = c_t\,dt + \int_{\mathbb{R}^+\backslash\{0\}} g(t-, x)\,\mu(dt, dx),$$

where c_t is a predictable and locally bounded process, μ an extended Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}^+ \setminus \{0\}$, and $g(s, x)$ is $\mathcal{F}_s \times \mathcal{B}(\mathbb{R}^+ \setminus \{0\})$ measurable in (ω, x) and càdlàg in s. Moreover, $g(s, x)$ takes only non-negative values.

Then for any $0 < r < 1$ the unique positive process $Y_t = X_t^r$ is representable as

$$Y_0 = X_0^r,$$

$$dY_t = a_t\,dt + \int_{\mathbb{R}^+\setminus\{0\}} w(t-, x)\,\mu(dt, dx),$$

where the drift $a_t := rX_{t-}^{r-1} c_t$ is predictable and locally bounded and where

$$w(s, x) := (X_s + g(s, x))^r - (X_s)^r$$

is $\mathcal{F}_s \times \mathcal{B}(\mathbb{R}^+)$ measurable in (ω, x) and càdlàg in s. Moreover, $w(s, x)$ takes only non-negative values.

Proof. Remark 3.3 implies the local boundedness of X_t within $\mathbb{R}^+$ and restating Proposition 3.2 in a differential manner gives

$$dX_t^r = rX_{t-}^{r-1} c_t d_t + \int_{\mathbb{R}^+\setminus\{0\}} ((X_{t-} + x)^r - X_{t-}^r)\,\mu_X(dt, dx).$$

Using the relation between μ_X and μ stated before the theorem, we obtain

$$dX_t^r = rX_{t-}^{r-1} c_t d_t + \int_{\mathbb{R}^+\setminus\{0\}} \left((X_{t-} + g(t-, x))^r - X_{t-}^r\right)\mu(dt, dx).$$

The positivity of $w(s, x)$ is a consequence of an elementary inequality recalled in the following lemma and the additional properties stated are now straightforward. ∎

For the sake of completeness and since it is essential to our results, we recall the following elementary inequality and give a proof.

Lemma 5.2. *For $a, x \in \mathbb{R}^+$ and $0 < r < 1$ it follows that $(a+x)^r - a^r$ is monotonically decreasing in a and*

$$(a+x)^r - a^r \leqslant x^r.$$

In particular, for $a, b \in \mathbb{R}^+$ the inequality $|a^r - b^r| \leqslant |a-b|^r$ holds true.

Proof. Define for fixed x the function $f: \mathbb{R}^+ \to \mathbb{R}$, $a \mapsto (a+x)^r - a^r$. Then

$$f'(a) = r\left((a+x)^{r-1} - a^{r-1}\right) \leqslant 0$$

since the $(r-1)$-st power is monotonically decreasing. Hence, f is monotonically decreasing and $f(a) = (a+x)^r - a^r \leqslant f(0) = x^r$. For the second inequality we assume without loss of generality that $a \geqslant b$. Then

$$|a^r - b^r| = \left(b + (a-b)\right)^r - b^r \leqslant (a-b)^r = |a-b|^r,$$

due to the first inequality. ∎

Remark 5.3. Actually the representation stated in Theorem 5.1 holds for arbitrary powers X_t^r with $r \in \mathbb{R}$. If $r \geqslant 1$, the assumption that X_t is locally bounded away from zero is no longer necessary.

For processes that start at zero or may become zero, we obviously cannot use Itô's formula in the above manner, since there is no way to extend the r-th power for $0 < r < 1$ to an open set containing $[0, \infty)$ in a continuously differentiable manner. Likewise, all advanced extensions of Itô's formula we know of (e.g. Bardina and Jolis (1997), Ghomrasni and Peskir (2003), Peskir (2005)) cannot be applied. For instance, the Boleau–Yor formula (Protter (2004), Theorem IV.77) allows for a non-continuous derivative, but still demands it to be bounded, however for r-th roots it is unbounded at zero. The Meyer–Itô formula (Protter (2004), Theorem IV.70) needs a left derivative, which again cannot be defined at zero. But by using the very standard Itô formula and applying a tailor-made limiting procedure, we can indeed verify an extension to processes that may become zero:

THEOREM 5.4. *Let $(X_t)_{t \in \mathbb{R}^+}$ be a given adapted càdlàg process which takes values in $\mathbb{R}^+$ and can be represented as*

$$dX_t = c_t\,dt + \int_{\mathbb{R}^+\setminus\{0\}} g(t-, x)\,\mu(dt, dx),$$

where c_t is a predictable and locally bounded process, μ an extended Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}^+\setminus\{0\}$, and $g(s, x)$ is $\mathcal{F}_s \times \mathcal{B}(\mathbb{R}^+\setminus\{0\})$ measurable in (ω, x) and càdlàg in s. Moreover, $g(s, x)$ takes only non-negative values. Assume that the integrals

$$\int_0^t rX_{s-}^{r-1} c_s\,ds \text{ (in the Lebesgue sense)}$$

and

$$\int_0^t \int_{\mathbb{R}^+\setminus\{0\}} (X_{s-} + g(s-, x))^r - (X_{s-})^r\,\mu(ds, dx)$$

exist a.s. for all $t \in \mathbb{R}^+$.

Then for any $0 < r < 1$ the unique positive process $Y_t = X_t^r$ is representable as

$$\text{(5.1)} \qquad \begin{aligned} Y_0 &= X_0^r, \\ dY_t &= a_t\,dt + \int_{\mathbb{R}^+\setminus\{0\}} w(t-, x)\,\mu(dt, dx), \end{aligned}$$

where the drift $a_t := rX_{t-}^{r-1} c_t$ is predictable and where

$$w(s, x) = (X_s + g(s, x))^r - (X_s)^r$$

is $\mathcal{F}_s \times \mathcal{B}(\mathbb{R}^+)$ measurable in (ω, x) and càdlàg in s. Moreover, $w(s, x)$ takes only non-negative values and Y_t is a.s. of finite variation.

Note that $c_t = 0$ implies $a_t = 0$ above, even if $X_{t-} = 0$, using the conventions of Lebesgue integration theory.

Proof. We first show that $Y_t = X_t^r$ is representable by (5.1). Recall below that all integrals can be viewed as pathwise Lebesgue–Stieltjes ones.

For any $\varepsilon > 0$ the process $X_{\varepsilon,t} := X_t + \varepsilon$ is bounded away from zero and

$$X_{\varepsilon,t} = X_0 + \varepsilon + \int_0^t c_s\,ds + \int_0^t \int_{\mathbf{R}^+\setminus\{0\}} g(s-, x)\,\mu(ds, dx).$$

From Theorem 5.1 we obtain

$$(5.2)\qquad (X_t+\varepsilon)^r = X_{\varepsilon,t}^r = (X_0+\varepsilon)^r + \int_0^t r(X_{s-}+\varepsilon)^{r-1} c_s\,ds$$
$$+\int_0^t \int_{\mathbf{R}^+\setminus\{0\}} \big((X_{s-}+\varepsilon+g(s-, x))^r - (X_{s-}+\varepsilon)^r\big)\,\mu(ds, dx).$$

For $s \in \mathbf{R}^+$ we clearly see that $(X_{s-}+\varepsilon)^r \to X_{s-}^r$ pointwise as $\varepsilon \to 0$. Moreover, since $r-1 \in (-1, 0)$, it follows that $(X_{s-}+\varepsilon)^{r-1}$ is decreasing in ε. Thus,

$$|r(X_{s-}+\varepsilon)^{r-1} c_s| \leqslant |rX_{s-}^{r-1} c_s| \quad \text{for all } \varepsilon > 0.$$

By assumption, $|rX_{s-}^{r-1} c_s|$ is Lebesgue-integrable over $[0, t]$, and so majorized convergence gives

$$\int_0^t r(X_{s-}+\varepsilon)^{r-1} c_s\,ds \to \int_0^t rX_{s-}^{r-1} c_s\,ds \quad \text{as } \varepsilon \to 0.$$

From Lemma 5.2 we see that $(X_{s-}+\varepsilon+g(s-, x))^r - (X_{s-}+\varepsilon)^r$ is positive and also decreasing in ε. So our assumptions and majorized convergence ensure that

$$\lim_{\varepsilon\to 0} \int_0^t \int_{\mathbf{R}^+\setminus\{0\}} \big((X_{s-}+\varepsilon+g(s-, x))^r - (X_{s-}+\varepsilon)^r\big)\,\mu(ds, dx)$$
$$= \int_0^t \int_{\mathbf{R}^+\setminus\{0\}} \big((X_{s-}+g(s-, x))^r - X_{s-}^r\big)\,\mu(ds, dx).$$

Combining these results we obtain, from (5.2) and by letting $\varepsilon \to 0$,

$$X_t^r = X_0^r + \int_0^t rX_{s-}^{r-1} c_s\,ds + \int_0^t \int_{\mathbf{R}^+\setminus\{0\}} \big((X_{s-}+g(s-, x))^r - X_{s-}^r\big)\,\mu(ds, dx),$$

which concludes the proof of the representation for Y_t.

To establish the finite variation of the process Y_t it suffices now to argue that both integral processes

$$\int_0^t rX_{s-}^{r-1} c_s\,ds \quad \text{and} \quad \int_0^t \int_{\mathbf{R}^+\setminus\{0\}} (X_{s-}+g(s-, x))^r - (X_{s-})^r\,\mu(ds, dx)$$

are of finite variation. For the second this is immediately clear and for the first we only need to observe that the existence in the Lebesgue sense implies the existence of $\int_0^t |rX_{s-}^{r-1} c_s|\, ds$. The latter is strictly increasing (thus of finite variation) when viewed as a process in t and its total variation is an upper bound for the total variation of the first integral. ∎

Remark 5.5. (a) Inspecting the proof it is clear that Theorem 5.1 remains valid when replacing the square root with any continuously differentiable function $f: \mathbf{R}^+ \to \mathbf{R}$. If additionally

$$|f'(x+\varepsilon)| \leqslant K\,|f'(x)| \quad \text{and} \quad |f(x+\varepsilon+y)-f(x+\varepsilon)| \leqslant \tilde{K}\,|f(x+y)-f(x)|$$

for all $x, y, \varepsilon \in \mathbf{R}^+$, where K and $\tilde{K}$ are some constants, the same is true for Theorem 5.4.

Then $f(X_t)$ is representable by (5.1) with

$$a_t = f'(X_{t-})\,c_t \quad \text{and} \quad w(t, x) = f\big(X_t+g(t, x)\big)-f(X_t).$$

(b) In general, r-th powers with $0 < r < 1$ of finite variation processes do not have to be of finite variation, as the following deterministic example exhibits. Let X_t be given by:

$$X_t = \frac{1}{n^2}-\left(1+\frac{1}{n}\right)\left(t-1+\frac{1}{n}\right) \quad \text{for } t \in \left[1-\frac{1}{n}, 1-\frac{1}{n+1}\right),\ n \in \mathbf{N},$$

$$X_t = 0 \qquad \text{for } t \in [1, \infty).$$

Then $X_{1-1/n} = 1/n^2$ and $X_{(1-1/(n+1))-} = 0$ for all $n \in \mathbf{N}$ and in each interval $[1-1/n, 1-1/(n+1))$ the process X_t is linearly decreasing. From this it is immediate to see that the total variation of $(X_t)_{t \in \mathbf{R}^+}$ is given by $2\sum_{n=1}^{\infty} 1/n^2-1$, which is finite. Likewise, we see that for $0 < r < 1$ the process X_t^r has jumps of size $1/n^{2r}$ at the times $1-1/n$. As $\sum_{n=1}^{\infty} 1/n^{\alpha}$ is infinite for all $\alpha \leqslant 1$, this shows that for $r \leqslant 1/2$ the process X_t^r is not of finite variation. Note, moreover, that X_t is of the form studied in Theorem 5.4 where $c_t = -(1+1/n)$ for $t \in [1-1/n, 1-1/(n+1))$, which is trivially predictable and locally bounded, $g(s, x) = x$ and

$$\mu(ds, dx) = \sum_{n=1}^{\infty} \delta_{(1-1/n)}(ds)\,\delta_{1/n^2}(dx)$$

with δ_v denoting the Dirac measure with respect to v.

Naturally, the next step is to give some readily checkable conditions for the existence of the integrals.

LEMMA 5.6. *The integral $\int_0^t \int_{\mathbf{R}^+\setminus\{0\}} w(s-, x)\,\mu(ds, dx)$ exists a.s. in the usual sense, if the integral $\int_0^t \int_{\mathbf{R}^+\setminus\{0\}} (g(s-, x))^r\,\mu(ds, dx)$ exists a.s. or there is some a.s. finite random variable $C > 0$ such that $X_t \geqslant C$ for all $t \in \mathbf{R}^+$.*

Proof. In the first case the existence follows by a standard majorization argument from $0 \leqslant w(s, x) = (X_s + g(s, x))^r - (X_s)^r \leqslant (g(s, x))^r$ (Lemma 5.2).

Likewise, we observe in the second case that we can argue ω-wise and the function $x \mapsto x^r$ is Lipschitz on any interval of the form $[a, \infty)$ with $a \in \mathbb{R}^+ \setminus \{0\}$. Thus there is a (possibly random) $K \in \mathbb{R}^+$ such that $0 \leqslant (X_s + g(s, x))^r - (X_s)^r \leqslant Kg(s, x)$. Hence, the claim follows by a dominated convergence argument, since the integral $\int_0^t \int_{\mathbb{R}^+ \setminus \{0\}} g(s-, x) \mu(ds, dx)$ exists. ∎

The condition $X_t \geqslant C$ actually means that Theorem 5.1 applies.

LEMMA 5.7. *The integral $\int_0^t \int_{\mathbb{R}^+ \setminus \{0\}} \left((X_{s-} + g(s-, x))^r - X_{s-}^r\right) \mu(ds, dx)$ exists in the usual sense, provided $c_t \geqslant 0$ for all $t \in \mathbb{R}^+$. In particular, the process X_t is monotonically increasing then.*

Proof. The monotonicity of X_t is obvious. We assume $c_t = 0$ for all $t \in \mathbb{R}^+$ first. As the mapping $x \mapsto x^r$ is monotone, also the process X_t^r has càdlàg monotonically increasing paths. Thus X_t^r is necessarily of finite variation. Denoting the variation of a function f over a time interval $[t_1, t_2]$ with $0 \leqslant t_1 \leqslant t_2$ by $\operatorname{var}(f; t_1, t_2)$, one deducts that

$$\operatorname{var}(X_t^r, t_1, t_2) = X_{t_2}^r - X_{t_1}^r = \sum_{t_1 < s \leqslant t_2} \Delta(X_s^r) = \sum_{t_1 < s \leqslant t_2} |\Delta(X_s^r)|.$$

But, obviously,

$$\sum_{t_1 \leqslant s \leqslant t_2} |\Delta(X_s^r)| = \int_{t_1}^{t_2} \int_{\mathbb{R}^+ \setminus \{0\}} \left|(X_{s-} + g(s-, x))^r - X_{s-}^r\right| \mu(ds, dx),$$

and hence the finite variation of X_t^r implies the existence of the integral.

If c_t does not vanish, we obtain $X_{t_2}^r - X_{t_1}^r \geqslant \sum_{t_1 < s \leqslant t_2} \Delta(X_s^r)$ and can then basically argue as before. ∎

LEMMA 5.8. *Suppose the function $g(s, x) = g(x)$ is deterministic and independent of s and the extended Poisson random measure μ is the jump measure of a Lévy subordinator with Lévy measure ν. Then the integral*

$$\int_0^t \int_{\mathbb{R}^+ \setminus \{0\}} \left((X_{s-} + g(x))^r - X_{s-}^r\right) \mu(ds, dx)$$

is a.s. defined for all $t \in \mathbb{R}^+$ provided $\int_{0 \leqslant x \leqslant 1} g(x)^r \nu(dx)$ is finite.

Proof. Recall that $E(\mu(ds, dx)) = ds \times \nu(dx)$ in the given set-up. The existence of the integral follows immediately by combining Lemma 5.6 and the fact that $\int_{0 \leqslant x \leqslant 1} g(x)^r \nu(dx) < \infty$ implies the existence of $\int_0^t \int_{\mathbb{R}^+ \setminus \{0\}} g(x)^r \mu(ds, dx)$ for all $t \in \mathbb{R}^+$ (cf. Marcus and Rosiński (2005), p. 113). ∎

Regarding the existence of the integral with respect to the Lebesgue measure, we only present the following criterion (a standard consequence of dominated convergence), which is applicable to many processes of interest.

LEMMA 5.9. *Assume that there exists a (possibly random) function $f: \mathbf{R}^+ \to \mathbf{R}^+$ with $\int_0^t f(t)\,dt < \infty$ a.s. such that $|rX_{t-}^{r-1} c_t| \leqslant f(t)$ for all $t \in \mathbf{R}^+$. Then the integral $\int_0^t rX_{t-}^{r-1} c_t\,dt$ exists in the Lebesgue sense. The latter is in particular the case if there are (possibly random) constants $C \geqslant 0$ and $\alpha > -1$ such that $|rX_{t-}^{r-1} c_t| \leqslant Ct^\alpha$.*

For positive Lévy processes, i.e. Lévy subordinators, one can immediately apply the above results and obtain the following

COROLLARY 5.10. *Let $(L_t)_{t \in \mathbf{R}^+}$ be a Lévy subordinator with initial value $L_0 \in \mathbf{R}^+$, associated drift γ and jump measure μ. Then for $0 < r < 1$ the unique positive root process L_t^r is of finite variation and*

$$dL_t^r = r\gamma L_{t-}^{r-1}\,dt + \int_{\mathbf{R}^+ \setminus \{0\}} ((L_{t-} + x)^r - L_{t-}^r)\,\mu(dt, dx),$$

where the drift $r\gamma L_{t-}^{r-1}$ is predictable. Moreover, the drift is locally bounded if and only if $L_0 > 0$ or $\gamma = 0$.

Proof. If γ is zero, the integrability condition imposed on the drift in Theorem 5.4 is trivially satisfied and in the case of a non-vanishing γ we know that $L_t \geqslant \gamma t$ for all $t \in \mathbf{R}^+$. The latter gives $r\gamma L_t^{r-1} \leqslant r\gamma^r t^{r-1}$, and so an application of Lemma 5.9 establishes the existence of $\int_0^t r\gamma L_t^{r-1}\,dt$ in the Lebesgue sense. Finally, noting that Lévy subordinators are monotonically increasing and using Lemma 5.7, we infer the corollary immediately from Theorem 5.4. The result on the local boundedness of the drift is immediate. ∎

5.2. The multivariate case. The aim of this section is to generalise the above univariate results to processes taking values in the cone of positive semidefinite $d \times d$ matrices. For reasons becoming clear later we only take square roots, but generalizations to general roots are straightforward and we shall indicate them. Before giving rigorous results and proofs, we want to give intuitive but non-rigorous arguments showing what the results should be. The reason is that for the rigorous proof we will need the multidimensional Itô formula and the derivative of the matrix square root, whereas the following two elementary lemmata immediately allow for an intuitive argument implying what the result should be. Though these lemmata are rather elementary, we decided to give complete proofs, as they seem to be unavailable in the standard literature, but should be useful in many situations.

The first result generalizes the representation for the product of two one-dimensional semimartingales (confer e.g. Protter (2004), p. 68) to matrix products of semimartingales and is briefly stated, without proof, in Karandikar (1991) (for the continuous case already in Karandikar (1982a, b)).

LEMMA 5.11. *Let $m, n, d \in N$ and $A_t \in M_{d,m}(\mathbb{R})$, $B_t \in M_{m,n}(\mathbb{R})$ be semimartingales. Then the matrix product $A_t B_t \in M_{d,n}(\mathbb{R})$ is a semimartingale and*

$$A_t B_t = \int_0^t A_{t-}\, dB_t + \int_0^t dA_t\, B_{t-} + [A, B]_t^M,$$

where $[A, B]_t^M \in M_{d,n}(\mathbb{R})$ is defined by

$$[A, B]_{t,ij}^M = \sum_{k=1}^m [A_{ik}, B_{kj}]_t.$$

If the continuous part of the quadratic covariation of A and B is zero, we have

$$[A, B]_t^M = A_0 B_0 + \sum_{0<s\leqslant t} \Delta A_s \Delta B_s.$$

Proof. Applying the univariate result componentwise to $A_t B_t$ we obtain for $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant n$:

$$(A_t B_t)_{ij} = \sum_{k=1}^m A_{t,ik} B_{t,kj} = \sum_{k=1}^m \Big(\int_0^t A_{s-,ik}\, dB_{kj,s} + \int_0^t B_{s-,kj}\, dA_{ik,s} + [A_{ik}, B_{kj}]_t\Big)$$

$$= \Big(\int_0^t A_{s-}\, dB_s + \int_0^t dA_s\, B_{s-} + [A, B]_t^M\Big)_{ij}.$$

In particular, we see immediately that all components of $A_t B_t$ are semimartingales being sums of products of semimartingales. Thus $A_t B_t$ is a matrix-valued semimartingale.

If the continuous quadratic covariation is zero, we have

$$[A, B]_t^M = \sum_{k=1}^m [A_{ik}, B_{kj}]_t = \sum_{k=1}^m \big(A_{0,ik} B_{0,kj} + \Delta A_{s,ik} \Delta B_{s,kj}\big)$$

$$= \big(A_0 B_0 + \sum_{0<s\leqslant t} \Delta A_s \Delta B_s\big)_{ij},$$

since $\Delta A_s = (\Delta A_{s,kl})_{1\leqslant k\leqslant d, 1\leqslant l\leqslant m}$ and likewise for B. ■

Remark 5.12. Obviously, the operator $[\cdot, \cdot]^M$ plays the same role for the matrix multiplication of matrix-valued semimartingales as the quadratic variation does for ordinary multiplication of one-dimensional semimartingales. Therefore we call the operator $[\cdot, \cdot]^M$ the *matrix covariation.* Note that in general it can be decomposed into

$$[A, B]_t^M = A_0 B_0 + [A, B]_t^{M,c} + \sum_{0<s\leqslant t} \Delta A_s \Delta B_s,$$

where $[A, B]_{t,ij}^{M,c} := \sum_{k=1}^m [A_{ik}, B_{kj}]_t^c$, i.e. into a continuous part and a pure jump part.

Our next result concerns quadratic equations of positive semidefinite matrices.

LEMMA 5.13. *Let $A, B \in S_d^+$. The equation*

$$X^2+AX+XA-B=0$$

has a unique positive semidefinite solution given by

$$X=\sqrt{A^2+B}-A.$$

Proof. We start by establishing the positive semidefiniteness of $\sqrt{A^2+B}-A$. It is clear that $A^2+B \geqslant A^2$. Observing that the matrix square root is a matrix monotone function (i.e. preserves the ordering on S_d^+, see e.g. Bhatia (1997), Proposition V.1.8), we have $\sqrt{A^2+B} \geqslant A$, which is equivalent to the claim.

Solving the equation can actually be done using the standard trick for complex quadratic equations:

$$X^2+AX+XA-B=(X+A)^2-A^2-B=0 \Leftrightarrow (X+A)^2=A^2+B.$$

Taking any "square root" on the right-hand side equation would now lead to a solution X. However, we consider only positive semidefinite solutions, and thus $X+A$ has to be in S_d^+, which is the case if and only if we take the unique positive semidefinite square root. Therefore there is one and only one solution in S_d^+ which is given by $X=\sqrt{A^2+B}-A$. ∎

Let now a positive semidefinite process X_t be given by

$$dX_t=c_t\,dt+\int_{S_d^+\setminus\{0\}} g(t-,x)\,\mu(dt,dx),$$

where c_t is an S_d-valued, predictable and locally bounded process, μ an extended Poisson random measure on $\mathbf{R}^+\times S_d^+\setminus\{0\}$ and $g(s,x)$ is $\mathscr{F}_s\times\mathscr{B}(S_d^+\setminus\{0\})$ measurable in (ω, x) and càdlàg in s. Moreover, $g(s,x)$ assumes only values in S_d^+. Suppose $Y_t:=\sqrt{X_t}$ is representable as

$$dY_t=a_t\,dt+\int_{S_d^+\setminus\{0\}} w(t-,x)\,\mu(dt,dx)$$

for some appropriate a_t and $w(t,x)$ being of the same type as c_t and $g(t,x)$. Using a differential version of Lemma 5.11 we obtain

$$\begin{aligned}
dY_t^2 &= Y_{t-}\,dY_t+dY_t\,Y_{t-}+d[Y,Y]_t^M = Y_{t-}\,dY_t+dY_t\,Y_{t-}+(\Delta Y_t)^2\\
&= Y_{t-}\Big(a_t\,dt+\int_{S_d^+\setminus\{0\}} w(t-,x)\,\mu(dt,dx)\Big)+\Big(a_t\,dt+\int_{S_d^+\setminus\{0\}} w(t-,x)\,\mu(dt,dx)\Big)Y_{t-}\\
&\quad+\int_{S_d^+\setminus\{0\}} w^2(t-,x)\,\mu(dt,dx)\\
&=\big(\sqrt{X_{t-}}\,a_t+a_t\sqrt{X_{t-}}\big)\,dt\\
&\quad+\int_{S_d^+\setminus\{0\}}\big(\sqrt{X_{t-}}\,w(t-,x)+w(t-,x)\sqrt{X_{t-}}+w^2(t-,x)\big)\,\mu(dt,dx).
\end{aligned}$$

As one clearly needs to have $dY_t^2 = dX_t$, the equations

$$c_t = \sqrt{X_{t-}}\, a_t + a_t \sqrt{X_{t-}}$$

and

$$\sqrt{X_{t-}}\, w(t-, x) + w(t-, x)\sqrt{X_{t-}} + w^2(t-, x) = g(t-, x)$$

have to hold. Assuming the necessary invertibility, we obtain $a_t = \mathbf{X}_{t-}^{-1} c_t$, where $\mathbf{X}_{t-}: M_d(\mathbf{R}) \to M_d(\mathbf{R})$ is the linear operator $Z \mapsto \sqrt{X_{t-}}\, Z + Z\sqrt{X_{t-}}$, and $w(s-, x) = \sqrt{X_{s-} + g(s-, x)} - \sqrt{X_{s-}}$ by Lemma 5.13. In the following we show that this representation for $\sqrt{X_t}$ is indeed true. It will also turn out that we implicitly obtained the derivative of the positive definite matrix square root, which is given in the next lemma. Here and in the following we regard S_d^{++} as a subset of the vector space S_d, which we identify with $\mathbf{R}^{d(d+1)/2}$.

LEMMA 5.14. *The positive definite square root $\sqrt{\cdot}: S_d^{++} \to S_d^{++}$ is continuously differentiable and the derivative $D\sqrt{X}$ is given by the inverse of the linear operator $Z \mapsto \sqrt{X}\, Z + Z\sqrt{X}$.*

Proof. The square root is the inverse of the bijective function $f: S_d^{++} \to S_d^{++}$, $X \mapsto X^2$. It is easy to see that $Df(X)$ is the linear operator $Z \mapsto XZ + ZX$ (see also Bhatia (1997), Example X.4.2). Using the relation $\sigma(Df(X)) = \sigma(X) + \sigma(X) \subset \mathbf{R}^+ \backslash \{0\}$, we see that $Df(X)$ is invertible for all $X \in S_d^{++}$. Thus, Rudin (1976), Theorem 9.24, shows that the square root is continuously differentiable and the derivative is given by the claimed linear operator. ∎

With the above considerations, we can now generalize our results on the behaviour of univariate square roots in a straightforward manner to the multivariate case.

THEOREM 5.15. *Let $(X_t)_{t\in\mathbf{R}^+}$ be a given adapted càdlàg process which takes values in S_d^{++}, is locally bounded within S_d^{++} and can be represented as*

$$dX_t = c_t\, dt + \int_{S_d^+\backslash\{0\}} g(t-, x)\, \mu(dt, dx), \tag{5.3}$$

where c_t is an S_d-valued, predictable and locally bounded process, μ an extended Poisson random measure on $\mathbf{R}^+ \times S_d^+\backslash\{0\}$, and $g(s, x)$ is $\mathcal{F}_s \times \mathcal{B}(S_d^+\backslash\{0\})$ measurable in (ω, x) and càdlàg in s. Moreover, $g(s, x)$ takes only values in S_d^+.

Then the integral $\int_0^t \int_{S_d^+\backslash\{0\}} (\sqrt{X_{s-} + g(s-, x)} - \sqrt{X_{s-}})\, \mu(ds, dx)$ exists a.s. for all $t \in \mathbf{R}^+$ and the unique positive definite square root process $Y_t = \sqrt{X_t}$ is given by

$$Y_0 = \sqrt{X_0},$$

$$dY_t = a_t\, dt + \int_{S_d^+\backslash\{0\}} w(t-, x)\, \mu(dt, dx),$$

with $a_t = \mathbf{X}_{t-}^{-1} c_t$, where $\mathbf{X}_{t-}$ is the linear operator $Z \mapsto \sqrt{X_{t-}}\, Z + Z\sqrt{X_{t-}}$ on $M_d(\mathbb{R})$. The drift process a_t is predictable and locally bounded and

$$w(s, x) := \sqrt{X_s + g(s, x)} - \sqrt{X_s}$$

is $\mathcal{F}_s \times \mathcal{B}(S_d^+ \backslash \{0\})$ measurable in (ω, x) and càdlàg in s. Moreover, $w(s, x)$ takes only positive semidefinite values.

Proof. The representation of Y_t follows from Proposition 3.4 and Lemma 5.14 by the same arguments as used for Theorem 5.1.

Using the vec-transformation and the Kronecker product, we easily see that the linear operator $\mathbf{X}_{t-}$ is symmetric (self-adjoint) and has a spectrum that is positive and locally bounded away from 0, since $\sigma(\mathbf{X}_{t-}) = \sigma(\sqrt{X_{t-}}) + \sigma(\sqrt{X_{t-}})$, the function $f: S_d^{++} \to S_d^{++}$, $Z \mapsto \min(\sigma(Z))$ is continuous and $\sqrt{X_{t-}}$ is locally bounded within S_d^{++}. The variational characterizations of the eigenvalues of a self-adjoint operator (cf. Horn and Johnson (1990), Section 4.2, for a matrix formulation) imply that

$$\min(\sigma(\mathbf{X}_{t-})) = \min_{\|x\|_{\tilde{2}} \neq 0} \left(\frac{\|\mathbf{X}_{t-} x\|_{\tilde{2}}}{\|x\|_{\tilde{2}}} \right).$$

Hence, $\|\mathbf{X}_{t-}^{-1}\|_{\tilde{2}} \leqslant \left(\min(\sigma(\mathbf{X}_{t-}))\right)^{-1}$ is locally bounded. Here $\|\cdot\|_{\tilde{2}}$ denotes the norm on $M_d(\mathbb{R})$ given by

$$\|x\|_{\tilde{2}} = \|\mathrm{vec}(x)\|_2 = \sqrt{\mathrm{tr}(xx^T)},$$

with $\|\cdot\|_2$ being the Euclidean norm on $\mathbb{R}^{d^2}$, and the associated operator norm on the linear operators over $M_d(\mathbb{R})$. This establishes the local boundedness of a_t.

That $w(s, x)$ takes only positive semidefinite values follows from Lemma 5.13, and the additional properties stated are straightforward. ∎

Remark 5.16. In principle we could immediately extend the above result to arbitrary r-th powers with $0 < r < 1$ again. Yet, this would mean that we need to calculate Df_r, where f_r denotes the unique positive definite r-th power and a_t would become $Df_r(X_{t-})c_t$. In general, there seems to be no useful formula for Df_r. Arguing as in Lemma 5.14 was possible for $r = 1/n$ with $n \in \mathbb{N}$, but then $Df_r(X)$ would be characterized as the inverse of the linear operator

$$Z \mapsto \sum_{j+k=n-1;\, j,k \in \mathbb{N}_0} X^{jr} Z X^{kr}.$$

Although in principle this can be applied, it appears to be infeasible for general n.

Assuming the existence of the relevant integrals, the strict positivity condition can again be relaxed. To be able to argue as in the univariate case we need two new technical results, the first one involving the so-called trace norm

$\|\cdot\|_{\mathrm{tr}}$ of matrices. For $A \in M_d(\mathbb{R})$ it is defined as $\|A\|_{\mathrm{tr}} = \mathrm{tr}\left((AA^*)^{1/2}\right)$ and it is easy to see that $\|A\|_{\mathrm{tr}} = \mathrm{tr}(A)$ for $A \in \mathbb{S}_d^+$.

LEMMA 5.17. *Let $A, B \in \mathbb{S}_d^+$ and $0 < r < 1$. Then the function $\mathbb{R}^+ \to \mathbb{R}^+$,*

$$\varepsilon \mapsto \|(A + \varepsilon I_d + B)^r - (A + \varepsilon I_d)^r\|_{\mathrm{tr}}$$

is monotonically decreasing. In particular,

$$\|(A + \varepsilon I_d + B)^r - (A + \varepsilon I_d)^r\|_{\mathrm{tr}} \leqslant \|(A + B)^r - A^r\|_{\mathrm{tr}} \quad \text{for all } \varepsilon \in \mathbb{R}^+.$$

Proof. Denote for some matrix $Z \in \mathbb{S}_d^+$ by $\lambda_1(Z), \lambda_2(Z), \ldots, \lambda_d(Z)$ the eigenvalues of Z sorted in ascending order.

Choose now some arbitrary $\varepsilon, \tilde{\varepsilon} \in \mathbb{R}^+$ with $\varepsilon \geqslant \tilde{\varepsilon}$. From Horn and Johnson (1990), Corollary 4.3.3, we obtain $\lambda_i(A+B) \geqslant \lambda_i(A)$ for $i = 1, 2, \ldots, d$. This implies, by Lemma 5.2, that

$$\begin{aligned}
&\sum_{i=1}^{d} \left((\lambda_i(A+B)+\varepsilon)^r - (\lambda_i(A)+\varepsilon)^r\right) \\
&\qquad = \sum_{i=1}^{d} \left((\lambda_i(A)+\varepsilon+\lambda_i(A+B)-\lambda_i(A))^r - (\lambda_i(A)+\varepsilon)^r\right) \\
&\qquad \leqslant \sum_{i=1}^{d} \left((\lambda_i(A)+\tilde{\varepsilon}+\lambda_i(A+B)-\lambda_i(A))^r - (\lambda_i(A)+\tilde{\varepsilon})^r\right) \\
&\qquad = \sum_{i=1}^{d} \left((\lambda_i(A+B)+\tilde{\varepsilon})^r - (\lambda_i(A)+\tilde{\varepsilon})^r\right).
\end{aligned}$$

Noting that the trace of a matrix is the sum of its eigenvalues and that $\lambda_i(Z + \varepsilon I_d) = \lambda_i(Z) + \varepsilon$ and $\lambda_i(Z)^r = \lambda_i(Z^r)$ for all $Z \in \mathbb{S}_d^+$ and $\varepsilon > 0$, we conclude that

$$\mathrm{tr}\left((A+\varepsilon I_d+B)^r\right) - \mathrm{tr}\left((A+\varepsilon I_d)^r\right) \leqslant \mathrm{tr}\left((A+\tilde{\varepsilon} I_d+B)^r\right) - \mathrm{tr}\left((A+\tilde{\varepsilon} I_d)^r\right).$$

This immediately implies

$$\|(A+\varepsilon I_d+B)^r - (A+\varepsilon I_d)^r\|_{\mathrm{tr}} \leqslant \|(A+\tilde{\varepsilon} I_d+B)^r - (A+\tilde{\varepsilon} I_d)^r\|_{\mathrm{tr}}.$$

This shows the claimed monotonicity and inequality, choosing $\tilde{\varepsilon} = 0$. ■

LEMMA 5.18. *Let $A \in \mathbb{S}_d^+$, $\varepsilon \in \mathbb{R}^+$ and denote by $\mathbf{A}_\varepsilon$ the linear operator*

$$M_d(\mathbb{R}) \to M_d(\mathbb{R}) \colon X \mapsto \sqrt{A + \varepsilon I_d}\, X + X \sqrt{A + \varepsilon I_d}.$$

Then $\|\mathbf{A}_\varepsilon^{-1} \mathrm{x}\|_{\tilde{2}}$ is decreasing in ε for every $\mathrm{x} \in M_d(\mathbb{R})$.

Here $\|\cdot\|_{\tilde{2}}$ denotes again the norm on $M_d(\mathbb{R})$ given by $\|\mathrm{x}\|_{\tilde{2}} = \|\mathrm{vec}(\mathrm{x})\|_2 = \sqrt{\mathrm{tr}(\mathrm{x}\mathrm{x}^T)}$, with $\|\cdot\|_2$ being the Euclidean norm on $\mathbb{R}^{d^2}$, and the associated operator norm on the linear operators over $M_d(\mathbb{R})$.

We understand $\|\mathbf{A}_0^{-1} \mathrm{x}\|_{\tilde{2}} = \infty$ in the case $A \in \mathbb{S}_d^+ \setminus \mathbb{S}_d^{++}$ above.

Proof. Note first that

$$\|\mathbf{A}_\varepsilon^{-1} x\|_{\tilde{2}} = \|(\sqrt{A+\varepsilon I_d}\otimes I_d + I_d\otimes\sqrt{A+\varepsilon I_d})^{-1}\operatorname{vec}(x)\|_2$$

and that

$$\sqrt{A+\varepsilon I_d}\otimes I_d + I_d\otimes\sqrt{A+\varepsilon I_d} \in S_{d^2}^+$$

and in particular is self-adjoint. Thus we have

$$\|\mathbf{A}_\varepsilon^{-1} x\|_{\tilde{2}} = \langle\operatorname{vec}(x), (\sqrt{A+\varepsilon I_d}\otimes I_d + I_d\otimes\sqrt{A+\varepsilon I_d})^{-2}\operatorname{vec}(x)\rangle^{1/2}.$$

Since taking the inverse reverses the ordering on $S_{d^2}^+$, this implies that it is sufficient to show that $(\sqrt{A+\varepsilon I_d}\otimes I_d + I_d\otimes\sqrt{A+\varepsilon I_d})^2$ is increasing in ε in the ordering on S_d. But let now $U\in M_d(\mathbb{R})$ be a unitary matrix such that U^*AU is diagonal; then

$$(U^*\otimes U^*)(\sqrt{A+\varepsilon I_d}\otimes I_d + I_d\otimes\sqrt{A+\varepsilon I_d})^2(U\otimes U)$$

is diagonal and obviously increasing in ε. Observing that $U\otimes U$ is again unitary and that such transformations preserve the ordering on S_d^+ concludes the proof. ■

PROPOSITION 5.19. *Let $(X_t)_{t\in\mathbb{R}^+}$ be a given adapted càdlàg process which takes values in S_d^+ and can be represented as*

$$dX_t = c_t\,dt + \int_{S_d^+\setminus\{0\}} g(t-, x)\,\mu(dt, dx),$$

where c_t is an S_d-valued, predictable and locally bounded process, μ an extended Poisson random measure on $\mathbb{R}^+\times S_d^+\setminus\{0\}$ and $g(s, x)$ is $\mathcal{F}_s\times\mathcal{B}(S_d^+\setminus\{0\})$ measurable in (ω, x) and càdlàg in s. Moreover, $g(s, x)$ takes values in S_d^+. Let $\mathbf{X}_{t-}$ be the linear operator $M_d(\mathbb{R})\to M_d(\mathbb{R})$, $Z\mapsto\sqrt{X_{t-}}\,Z+Z\sqrt{X_{t-}}$ and assume that the integrals

$$\int_0^t \mathbf{X}_{s-}^{-1} c_s\,ds \text{ (in the Lebesgue sense)}$$

and

$$\int_0^t \int_{S_d^+\setminus\{0\}} (\sqrt{X_{s-}+g(s-, x)}-\sqrt{X_{s-}})\,\mu(ds, dx)$$

exist a.s. for all $t\in\mathbb{R}^+$.

Then the unique positive semidefinite square root process $Y_t=\sqrt{X_t}$ is representable as

$$Y_0 = \sqrt{X_0},$$

$$dY_t = a_t\,dt + \int_{S_d^+\setminus\{0\}} w(t-, x)\,\mu(dt, dx), \tag{5.4}$$

where the drift $a_t = \mathbf{X}_{t-}^{-1} c_t$ *is predictable and where*

$$w(s, x) := \sqrt{X_s + g(s, x)} - \sqrt{X_s}$$

is $\mathscr{F}_s \times \mathscr{B}(S_d^+ \setminus \{0\})$ *measurable in* (ω, x) *and càdlàg in s. Moreover,* $w(s, x)$ *takes only positive semidefinite values and* Y_t *is a.s. of finite variation.*

Due to the conventions of Lebesgue integration theory we always have $a_t = 0$ if $c_t = 0$ above.

Proof. We first show that $Y_t = \sqrt{X_t}$ is representable by (5.4). Recall below that the integral of an $M_d(\mathbf{R})$-valued function exists if and only if the integral of the norm exists for one and hence all norms on $M_d(\mathbf{R})$.

For any $\varepsilon > 0$ we define the process $X_{\varepsilon,t} := X_t + \varepsilon I_d$. Obviously, $X_{\varepsilon,t} \geqslant \varepsilon I_d$ for all $t \in \mathbf{R}^+$ and the process $X_{\varepsilon,t}$ is of finite variation, and hence locally bounded. Observing that for all $\delta, K > 0$ the set $\{x \in S_d^{++} : x \geqslant \delta I_d, \|x\| \leqslant K\}$ is convex and compact, we infer that $X_{\varepsilon,t}$ is locally bounded within S_d^{++} and

$$X_{\varepsilon,t} = X_0 + \varepsilon I_d + \int_0^t c_s\, ds + \int_0^t \int_{S_d^+ \setminus \{0\}} g(s-, x)\, \mu(ds, dx).$$

From Theorem 5.15 we obtain

$$\sqrt{X_t + \varepsilon I_d} = \sqrt{X_{\varepsilon,t}} = \sqrt{X_0 + \varepsilon I_d} + \int_0^t \mathbf{X}_{\varepsilon,s-}^{-1} c_s\, ds + \int_0^t \int_{S_d^+ \setminus \{0\}} \left(\sqrt{X_{s-} + \varepsilon I_d + g(s-, x)} - \sqrt{X_{s-} + \varepsilon I_d}\right) \mu(ds, dx), \tag{5.5}$$

where $\mathbf{X}_{\varepsilon,s-}$ denotes the linear operator

$$M_d(\mathbf{R}) \to M_d(\mathbf{R}):\ Z \mapsto \sqrt{X_{s-} + \varepsilon I_d}\, Z + Z \sqrt{X_{s-} + \varepsilon I_d}.$$

For $s \in \mathbf{R}^+$ we clearly see that $\sqrt{X_{s-} + \varepsilon} \to \sqrt{X_{s-}}$ and $\mathbf{X}_{\varepsilon,s-} \to \mathbf{X}_{s-}$ pointwise as $\varepsilon \to 0$. Moreover, Lemma 5.18 ensures $\|\mathbf{X}_{\varepsilon,s-}^{-1} c_s\|_2 \leqslant \|\mathbf{X}_{s-}^{-1} c_s\|_2$ for all $\varepsilon > 0$. By assumption, $\|\mathbf{X}_{\varepsilon,s-}^{-1} c_s\|_2$ is Lebesgue-integrable over $[0, t]$, and so majorized convergence gives

$$\int_0^t \mathbf{X}_{\varepsilon,s-}^{-1} c_s\, ds \to \int_0^t \mathbf{X}_{s-}^{-1} c_s\, ds \quad \text{as } \varepsilon \to 0.$$

From Lemma 5.17 we see that $\|\sqrt{X_{s-} + \varepsilon I_d + g(s-, x)} - \sqrt{X_{s-} + \varepsilon I_d}\|_{\mathrm{tr}}$ is decreasing in ε. So our assumptions and majorized convergence ensure that

$$\lim_{\varepsilon \to 0} \int_0^t \int_{S_d^+ \setminus \{0\}} \left(\sqrt{X_{s-} + \varepsilon I_d + g(s-, x)} - \sqrt{X_{s-} + \varepsilon I_d}\right) \mu(ds, dx) = \int_0^t \int_{S_d^+ \setminus \{0\}} \left(\sqrt{X_{s-} + g(s-, x)} - \sqrt{X_{s-}}\right) \mu(ds, dx).$$

Combining these results we obtain, from (5.5) and by letting $\varepsilon \to 0$,

$$\sqrt{X_t} = \sqrt{X_0} + \int_0^t \mathrm{X}_{s-}^{-1} c_s\, ds + \int_0^t \int_{S_d^+\setminus\{0\}} \left(\sqrt{X_{s-}+g(s-,x)} - \sqrt{X_{s-}}\right) \mu(ds, dx),$$

which concludes the proof of the representation for Y_t.

To establish the finite variation of the process Y_t it suffices now to argue that both integral processes

$$\int_0^t \mathrm{X}_{s-}^{-1} c_s\, ds \quad \text{and} \quad \int_0^t \int_{S_d^+\setminus\{0\}} \left(\sqrt{X_{s-}+g(s-,x)} - \sqrt{X_{s-}}\right) \mu(ds, dx)$$

are of finite variation. For the second this is immediately clear and for the first we only need to observe that the existence in the Lebesgue sense implies the existence of $\int_0^t \|\mathrm{X}_{s-}^{-1} c_s\|\, ds$ for any norm $\|\cdot\|$. The latter is strictly increasing (thus of finite variation) when viewed as a process in t and its total variation is an upper bound for the total variation of the first integral calculated using the same norm $\|\cdot\|$. ∎

Remark 5.20. When replacing the square root with an arbitrary continuously differentiable function $f: S_d^+ \to S_d$, the above proposition remains valid if $\|Df(x+\varepsilon I_d)z\| \leqslant K\|Df(x)z\|$ and

$$\|f(x+\varepsilon I_d+y) - f(x+\varepsilon I_d)\| \leqslant \tilde{K}\|f(x+y)-f(x)\| \tag{5.6}$$

for all $x, y \in S_d^+$, $z \in S_d$ and $\varepsilon \in \boldsymbol{R}^+$, where K and $\tilde{K}$ are some constants. Then $f(X_t)$ is representable by (5.4) with $a_t = Df(X_{t-})c_t$ and $w(t, x) = f(X_t + g(t, x)) - f(X_t)$.

For general r-th powers with $0 < r < 1$ condition (5.6) holds due to Lemma 5.17. In particular, this implies that the above theorem applies immediately to the r-th power if $c_t = 0$ for all $t \in \boldsymbol{R}^+$. Furthermore, the square root can be replaced by the r-th power in the following Lemma 5.23, Corollary 5.24, and Lemmata 5.25–5.27.

Before giving criteria for the existence of the integrals assumed in the above theorem, we establish some auxiliary results. The first one establishes that S_d^+-increasing functions are always of finite variation.

LEMMA 5.21. *Let $f: \boldsymbol{R}^+ \to S_d^+$ be an S_d^+-increasing function, i.e. $f(a) \leqslant f(b)$ for all $a, b \in \boldsymbol{R}^+$ with $a \leqslant b$. Then f is of finite variation on compacts.*

Proof. Obviously, we are free to choose any norm on $M_d(\boldsymbol{R})$. Let thus $\|\cdot\|_{\mathrm{tr}}$ again denote the trace norm and recall that $\|A\|_{\mathrm{tr}} = \mathrm{tr}(A)$ for all $A \in S_d^+$. For $s, t \in \boldsymbol{R}^+$, $t \geqslant s$, we obtain

$$\|f(t)-f(s)\|_{\mathrm{tr}} = \mathrm{tr}\left(f(t)-f(s)\right) = \mathrm{tr}\left(f(t)\right) - \mathrm{tr}\left(f(s)\right),$$

due to the linearity of the trace. From this we can immediately conclude that the total variation of f over any interval $[a, b]$ with $a, b \in \mathbf{R}^+$, $a \leqslant b$, calculated in the trace norm is given by $\operatorname{tr}(f(b)) - \operatorname{tr}(f(a))$, which is finite. Hence, f is of finite variation on compacts. ■

The trace norm has also been used in Pérez-Abreu and Rocha-Arteaga (2005) and Barndorff-Nielsen and Pérez-Abreu (2007) and thus seems to be very well adapted to the structure of matrix subordinators. The lemma could alternatively be easily established using the theory for general cones developed in Duda (2005) and the properties of the trace functional/norm.

Moreover, we need to consider an appropriate matrix extension of the inequality $\sqrt{a+b} - \sqrt{a} \leqslant \sqrt{b}$ for all $a, b \in \mathbf{R}^+$. Actually, the question whether $\sqrt{A+B} - \sqrt{A} \leqslant \sqrt{B}$ for $A, B \in \mathbf{S}_d^+$ seems not to have been discussed in the literature yet. However, the following norm version suffices for our purposes.

DEFINITION 5.22. Let $A, B \in M_d(\mathbf{R})$. Then $|A| = (A^* A)^{1/2}$ is called the *modulus* (*absolute value*) of A.

A norm $\|\cdot\|$ on $M_d(\mathbf{R})$ is said to be *unitarily invariant* if $\|UAV\| = \|A\|$ for all unitary matrices $U, V \in M_d(\mathbf{R})$.

For more information see e.g. Bhatia (1997) and for unitarily invariant norms also Horn and Johnson (1990).

LEMMA 5.23 (Ando (1988), Corollary 2). *Let $A, B \in \mathbf{S}_d^+$ and $\|\cdot\|$ be any unitarily invariant norm. Then*

$$\|\sqrt{A} - \sqrt{B}\| \leqslant \|\sqrt{|A-B|}\|.$$

This result has originally been obtained in Birman et al. (1975). We can simplify the result somewhat by using the operator norm associated to the usual Euclidean norm on $\mathbf{R}^d$.

COROLLARY 5.24 (cf. Bhatia (1997), Section X.1). *Let $A, B \in \mathbf{S}_d^+$ and let $\|\cdot\|_2$ denote the operator norm associated with the Euclidean norm. Then*

$$\|\sqrt{A} - \sqrt{B}\|_2 \leqslant \sqrt{\|(|A-B|)\|_2}.$$

In particular, $\|\sqrt{A+B} - \sqrt{A}\|_2 \leqslant \sqrt{\|B\|_2}$.

Armed with these prerequisites we can now state criteria for the existence of the integrals in Proposition 5.19.

LEMMA 5.25. *The integral $\int_0^t \int_{\mathbf{S}_d^+ \setminus \{0\}} w(s-, x)\, \mu(ds, dx)$ exists a.s. for all $t \in \mathbf{R}^+$ in the usual sense if the integrals*

$$\int_0^t \int_{\mathbf{S}_d^+ \setminus \{0\}} \sqrt{\|g(s-, x)\|_2}\, \mu(ds, dx) \quad \text{or} \quad \int_0^t \int_{\mathbf{S}_d^+ \setminus \{0\}} \sqrt{g(s-, x)}\, \mu(ds, dx)$$

exist a.s. for all $t \in \mathbf{R}^+$ or there is some $\mathbf{S}_d^{+\,+}$-valued random variable C such that $X_t \geqslant C$ for all $t \in \mathbf{R}^+$.

Due to the equivalence of all norms one can actually use any other norm instead of $\|\cdot\|_2$. Moreover, the second case corresponds to Theorem 5.15.

Proof. First of all we note that $\int_0^t \int_{S_d^+\setminus\{0\}} \sqrt{\|g(s-,x)\|_2}\,\mu(ds,dx)$ exists if and only if the integral $\int_0^t \int_{S_d^+\setminus\{0\}} \sqrt{g(s-,x)}\,\mu(ds,dx)$ exists. This follows immediately, since according to the definition of integration with respect to Poisson random measures the integral $\int_0^t \int_{S_d^+\setminus\{0\}} \sqrt{g(s-,x)}\,\mu(ds,dx)$ exists if and only if $\int_0^t \int_{S_d^+\setminus\{0\}} \|\sqrt{g(s-,x)}\|\,\mu(ds,dx)$ exists for one and hence all norms $\|\cdot\|$, and $\|\sqrt{x}\|_2 = \sqrt{\|x\|_2}$ for all $x \in S_d^+$.

Noting that Corollary 5.24 gives $\|w(s-,x)\|_2 \leqslant \sqrt{\|g(s-,x)\|_2}$, a simple majorization argument establishes the existence of $\int_0^t \int_{S_d^+\setminus\{0\}} w(s-,x)\,\mu(ds,dx)$ in the first case.

Assume now that $X_t \geqslant C$ for all $t \in \boldsymbol{R}^+$ holds with some $C \in S_d^+$. Then we once again argue ω-wise. The square root function is Lipschitz on any set $\mathscr{A} \subset S_d^{++}$ for which there is some $C_0 \in S_d^{++}$ such that $C \geqslant C_0$ for all $C \in \mathscr{A}$ (see, for instance, Bhatia (1997), p. 305). Thus there exists a constant K (possibly depending on C) such that

$$\|\sqrt{X_{s-}+g(s-,x)} - \sqrt{X_{s-}}\| \leqslant K\|g(s-,x)\|.$$

This implies the existence of the integral, as $\int_0^t \int_{S_d^+\setminus\{0\}} g(s-,x)\,\mu(ds,dx)$ exists due to our assumptions on the process X_t. ∎

Lemma 5.26. *The integral $\int_0^t \int_{S_d^+\setminus\{0\}} w(s-,x)\,\mu(ds,dx)$ exists a.s. for all $t \in \boldsymbol{R}^+$ in the usual sense provided $c_t \in S_d^+$ for all $t \in \boldsymbol{R}^+$, i.e. the process X_t is S_d^+-increasing.*

Proof. The S_d^+-increasingness of X_t is clear. Since the square root preserves the ordering on S_d^+, the process $\sqrt{X_t}$ is S_d^+-increasing, as well. Thus, Lemma 5.21 ensures that $\sqrt{X_t}$ is of finite variation.

Now, we first assume $c_t = 0$ for all $t \in \boldsymbol{R}^+$. Denoting the variation (in the trace norm) of a function f over a time interval $[t_1, t_2]$ with $0 \leqslant t_1 \leqslant t_2$ by $\operatorname{var}(f; t_1, t_2)$, one deducts that

$$\operatorname{var}(\sqrt{X}, t_1, t_2) = \operatorname{tr}(\sqrt{X_{t_2}}) - \operatorname{tr}(\sqrt{X_{t_1}}) = \sum_{t_1 < s \leqslant t_2} \|\Delta(\sqrt{X_s})\|_{\mathrm{tr}}.$$

But, obviously,

$$\sum_{t_1 < s \leqslant t_2} \|\Delta(\sqrt{X_s})\|_{\mathrm{tr}} = \int_{t_1}^{t_2} \int_{S_d^+\setminus\{0\}} \|\sqrt{X_{s-}+g(s-,x)} - \sqrt{X_{s-}}\|_{\mathrm{tr}}\,\mu(ds,dx),$$

and hence the finite variation of $\sqrt{X_t}$ implies the existence of the integral.

If c_t does not vanish, then we obtain

$$\operatorname{tr}(\sqrt{X_{t_2}})-\operatorname{tr}(\sqrt{X_{t_1}}) \geqslant \sum_{t_1 < s \leqslant t_2} \|\Delta(\sqrt{X_s})\|_{\mathrm{tr}}$$

and can argue as before. ∎

For the following recall that we refer to $\mathbb{S}_d^+$-increasing Lévy processes as matrix subordinators.

LEMMA 5.27. *Suppose the function $g(s, x) = g(x)$ is deterministic and independent of s and the extended Poisson random measure μ is the jump measure of a matrix subordinator with Lévy measure ν. Then the integral*

$$\int_0^t \int_{\mathbb{S}_d^+\setminus\{0\}} (\sqrt{X_{s-}+g(x)}-\sqrt{X_{s-}})\,\mu(ds, dx)$$

is indeed a.s. defined for all $t \in \mathbb{R}^+$ provided $\int_{0\leqslant \|x\|_2\leqslant 1, x\in\mathbb{S}_d^+\setminus\{0\}} \sqrt{\|g(x)\|_2}\,\nu(dx)$ is finite.

Again we can use any other norm instead of $\|\cdot\|_2$.

Proof. Recall that $E(\mu(ds, dx)) = ds \times \nu(dx)$ in the given set-up. The existence of the integral follows immediately by combining Lemma 5.25 and the fact that

$$\int_{\|x\|_2\leqslant 1} \sqrt{\|g(x)\|_2}\,\nu(dx) = \int_{\|x\|_2\leqslant 1} \|\sqrt{g(x)}\|_2\,\nu(dx) < \infty$$

implies the existence of $\int_0^t \int_{\mathbb{S}_d^+\setminus\{0\}} \sqrt{g(x)}\,\mu(ds, dx)$ for all $t \in \mathbb{R}^+$ (cf. Marcus and Rosiński (2005), p. 113). Here we note that

$$\int_0^t \int_{\mathbb{S}_d^+\setminus\{0\}} \min(\|g(x)\|_2, 1)\,\nu(dx)\,ds \leqslant t\Big(\nu(\{x\in\mathbb{S}_d^+ : \|x\|_2 > 1\})+\int_{\|x\|_2\leqslant 1} \sqrt{\|g(x)\|_2}\,\nu(dx)\Big)$$

is finite. ∎

Regarding the existence of the integral with respect to the Lebesgue measure, we only restate the criterion of Lemma 5.9 for the multivariate case.

LEMMA 5.28. *Assume that there exists a (possibly random) function $f: \mathbb{R}^+ \to \mathbb{R}^+$ with $\int_0^t f(t)\,dt < \infty$ a.s. such that $\|X_{t-}^{-1} c_t\| \leqslant f(t)$ for all $t \in \mathbb{R}^+$. Then the integral $\int_0^t X_{t-}^{-1} c_t\,dt$ exists in the Lebesgue sense. The latter is in particular the case if there are (possibly random) constants $C \geqslant 0$ and $\alpha > -1$ such that $\|X_{t-}^{-1} c_t\| \leqslant Ct^\alpha$.*

After these general considerations we shall now turn to studying the roots of matrix subordinators.

COROLLARY 5.29. *Let $(L_t)_{t\in\mathbb{R}^+}$ be a matrix subordinator with initial value $L_0 \in \mathbb{S}_d^+$, associated drift γ and jump measure μ. Then the unique positive semi-*

definite process $\sqrt{L_t}$ *is of finite variation and, provided that either* $L_0 \in S_d^{++}$ *or* $\gamma \in S_d^{++} \cup \{0\}$,

$$d\sqrt{L_t} = \mathbf{L}_{t-}^{-1} \gamma dt + \int_{S_d^+ \setminus \{0\}} (\sqrt{L_{t-} + x} - \sqrt{L_{t-}}) \mu(dt, dx),$$

where $\mathbf{L}_{t-}$ *is the linear operator on* $M_d(\mathbf{R})$ *with* $Z \mapsto \sqrt{L_{t-}}\, Z + Z \sqrt{L_{t-}}$. *The drift* $\mathbf{L}_{t-}^{-1} \gamma$ *is predictable, and additionally locally bounded provided* $L_0 \in S_d^{++}$ *or* $\gamma = 0$.

Proof. As the square root preserves the ordering on S_d^+, $\sqrt{L_t}$ is S_d^+-increasing, and thus of finite variation by Lemma 5.21.

In the case $L_0 \in S_d^{++}$ the corollary follows from Theorem 5.15.

Else we know from Lemma 5.26 that the integral

$$\int_0^t \int_{S_d^+ \setminus \{0\}} (\sqrt{L_{s-} + x} - \sqrt{L_{s-}}) \mu(ds, dx)$$

exists a.s. for all $t \in \mathbf{R}^+$. Next we show that the integral $\int_0^t \mathbf{L}_{s-}^{-1} \gamma ds$ exists for all $t \in \mathbf{R}^+$. For $\gamma = 0$ this is trivial. For $\gamma \in S_d^{++}$, we have $L_s \geqslant \gamma s \in S_d^{++}$. Using the variational characteristics of the eigenvalues as in the proof of Theorem 5.15 we get

$$\min_{\|x\|_2 \neq 0} \left(\frac{\|\mathbf{L}_{s-} x\|_2}{\|x\|_2} \right) = \min(\sigma(\mathbf{L}_{s-})) = 2 \min(\sigma(\sqrt{L_{s-}})) \geqslant 2\sqrt{s}\sqrt{\min(\sigma(\gamma))}.$$

Therefore

$$\|\mathbf{L}_{s-}^{-1}\|_2 \leqslant \left(\min(\sigma(\mathbf{L}_{s-}))\right)^{-1} \leqslant \left(2\sqrt{\min(\sigma(\gamma))}\right)^{-1} s^{-1/2}.$$

Hence, $\|\mathbf{L}_{s-}^{-1} \gamma\| \leqslant C s^{-1/2}$ for all $s \in \mathbf{R}^+$ with some constant $C \in \mathbf{R}^+$, and so Lemma 5.28 establishes the existence of $\int_0^t \mathbf{L}_{s-} c_s ds$ for all $t \in \mathbf{R}^+$ in the Lebesgue sense. Therefore Proposition 5.19 concludes the proof. ■

Remark 5.30. If the Lévy process is supposed to have initial value in ∂S_d^+ (e.g. zero, as is usual) and non-zero drift $\gamma \in \partial S_d^+$, then there appears to be basically no hope to obtain a representation of the above type.

6. ROOTS OF ORNSTEIN-UHLENBECK PROCESSES

Now we turn to studying the behaviour of the roots of positive Ornstein–Uhlenbeck processes as defined in Section 4. Recall in particular that the driving Lévy process L_t is assumed to be a (matrix) subordinator.

Straightforward calculations based on Theorems 5.1 and 5.4 establish the following result for a univariate OU process $dX_t = -\lambda X_{t-} dt + dL_t$.

PROPOSITION 6.1. *Let $(X_t)_{t\in\mathbb{R}^+}$ be a positive univariate process of Ornstein–Uhlenbeck type driven by a Lévy subordinator L_t with drift γ and associated Poisson random measure μ. Then for $0<r<1$ the unique positive r-th power $Y_t = X_t^r$ is of finite variation and has the following representation:*

$$dY_t = (-\lambda r X_{t-}^r + \gamma r X_{t-}^{r-1})\,dt + \int_{\mathbb{R}^+\setminus\{0\}} \left((X_{t-}+x)^r - (X_{t-})^r\right)\mu(dt, dx)$$

$$= (-\lambda r Y_{t-} + \gamma r Y_{t-}^{1-1/r})\,dt + \int_{\mathbb{R}^+\setminus\{0\}} \left((Y_{t-}^{1/r}+x)^r - Y_{t-}\right)\mu(dt, dx),$$

provided that the process X_t is locally bounded away from zero or the integrals

$$\int_0^t \gamma r X_{s-}^{r-1}\,ds \quad \text{and} \quad \int_0^t \int_{\mathbb{R}^+\setminus\{0\}} \left((X_{s-}+x)^r - X_{s-}^r\right)\mu(ds, dx)$$

exist a.s. for all $t\in\mathbb{R}$.

Before showing that the conditions are actually satisfied for all positive OU processes, we show this for stationary ones, as this case is of particular interest and the proof is very straightforward. Recall in particular that a stationary OU process can be represented as $\int_{-\infty}^t e^{-\lambda(t-s)}\,dL_s$, where the driving Lévy process has a finite logarithmic moment.

PROPOSITION 6.2. *Let X_t be a stationary positive process of Ornstein–Uhlenbeck type with driving Lévy process L_t (having drift γ and non-zero Lévy measure ν). Then it is locally bounded away from zero.*

The same holds for any positive Ornstein–Uhlenbeck process X_t with $X_0 > 0$ a.s.

Proof. Let us first consider the stationary case. If $\gamma > 0$, we see from Proposition 4.6 that $X_t \geqslant \gamma/\lambda > 0$ for all t, which implies that X_t is locally bounded away from 0. Otherwise note first that $X_t \geqslant e^{-\lambda t} X_0$ for all $t \geqslant 0$ and that the stationary distribution is self-decomposable (cf. Sato (1999), Theorem 17.5). As the driving Lévy process has a non-zero Lévy measure, the stationary distribution must be non-trivial, and thus, by Sato (1999), Example 27.8, absolutely continuous with respect to the Lebesgue measure. Therefore we have $X_0 > 0$ a.s. Hence, there is a.s. a sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ increasing to infinity such that $X_t \geqslant 1/n$ for all $t\in[0, T_n)$ (actually we can set $T_n = \ln(X_0 n)/\lambda$), which implies that X_t is locally bounded away from the origin.

Obviously, the same arguments apply in the non-stationary case. ■

PROPOSITION 6.3. *Let $(X_t)_{t\in\mathbb{R}^+}$ be a positive univariate process of Ornstein–Uhlenbeck type driven by a Lévy subordinator L_t with drift γ and associated Poisson random measure μ. Then for $0<r<1$ the integrals*

$$\int_0^t \gamma r X_{s-}^{r-1}\,ds \quad \text{and} \quad \int_0^t \int_{\mathbb{R}^+\setminus\{0\}} \left((X_{s-}+x)^r - X_{s-}^r\right)\mu(ds, dx)$$

exist for all $t\in\mathbb{R}$.

Proof. To show this we introduce the auxiliary process

$$Z_t = X_0 + \int_0^t e^{\lambda s}\, dL_s \quad \text{for } t \in \mathbf{R}^+.$$

It holds that $Z_t = e^{\lambda t} X_t$ for all $t \in \mathbf{R}^+$, the process is monotonically increasing and

$$dZ_t = e^{\lambda t} \gamma dt + \int_{\mathbf{R}^+ \setminus \{0\}} e^{\lambda t} x \mu(dt, dx).$$

The increasingness implies the existence of the integral

$$\int_0^t \int_{\mathbf{R}^+ \setminus \{0\}} ((Z_{s-} + e^{\lambda s} x)^r - Z_{s-}^r)\, \mu(ds, dx) = \int_0^t \int_{\mathbf{R}^+ \setminus \{0\}} e^{\lambda r s} ((X_{s-} + x)^r - X_{s-}^r)\, \mu(ds, dx).$$

Since $0 < \min\{1, e^{\lambda r t}\} \leqslant e^{\lambda r s} \leqslant \max\{1, e^{\lambda r t}\}$ for all $s \in [0, t]$, this shows that the integral

$$\int_0^t \int_{\mathbf{R}^+ \setminus \{0\}} ((X_{s-} + x)^r - X_{s-}^r)\, \mu(ds, dx)$$

exists for all $t \in \mathbf{R}$.

Obviously,

$$Z_t \geqslant \int_0^t e^{\lambda s} \gamma ds = \frac{\gamma}{\lambda}(e^{\lambda t} - 1).$$

Assuming first $\lambda \geqslant 0$, we obtain

$$\int_0^t \gamma r X_{s-}^{r-1}\, ds = \int_0^t r\gamma e^{-\lambda(r-1)s} Z_{s-}^{r-1}\, ds \leqslant \int_0^t r\gamma^r \lambda^{1-r} e^{-\lambda(r-1)s} (e^{\lambda s} - 1)^{r-1}\, ds$$

$$= r\gamma^r \lambda^{1-r} e^{-\lambda(r-1)t} \int_0^t (e^{\lambda s} - 1)^{r-1}\, ds.$$

Noting that $e^{\lambda s} - 1 \geqslant s$ for all $s \in \mathbf{R}^+$, we infer the existence of $\int_0^t \gamma r X_{s-}^{r-1}\, ds$ for all $t \in \mathbf{R}^+$ immediately. In the case $\lambda < 0$ one calculates

$$\int_0^t \gamma r X_{s-}^{r-1}\, ds \leqslant r\gamma^r |\lambda|^{1-r} \int_0^t (e^{-\lambda s} - 1)^{r-1}\, ds,$$

which likewise implies the existence of the integral for all $t \in \mathbf{R}^+$. ∎

Remark 6.4. For a driftless driving Lévy process we see from

$$dY_t = -\lambda r Y_{t-}\, dt + \int_{\mathbf{R}^+ \setminus \{0\}} ((Y_{t-}^{1/r} + x)^r - Y_{t-})\, \mu(dt, dx) \tag{6.1}$$

that the drift part is again that of an Ornstein–Uhlenbeck process.

Moreover, observe that (6.1) gives a stochastic differential equation (cf. Applebaum (2004) for information on this type of SDEs) for the r-th power (with $0 < r < 1$) of the OU process. Since the derivative of $y \mapsto (y^{1/r} + x)^r$

is given by $y \mapsto \left(y^{1/r}/(y^{1/r}+x)\right)^{1-r}$ and is thus obviously bounded by one for all $x \in \mathbb{R}^+$, the function $y \mapsto (y^{1/r}+x)^r$ is (globally) Lipschitz. This implies that for any initial value Y_0 the SDE (6.1) has a unique solution.

If $\gamma > 0$, one likewise has the SDE

$$dY_t = (-\lambda r Y_{t-} + \gamma r Y_{t-}^{1-1/r})\,dt + \int_{\mathbb{R}^+\setminus\{0\}} \left((Y_{t-}^{1/r}+x)^r - Y_{t-}\right)\mu(dt, dx)$$

for the r-th power of the OU process. In this case one has only local Lipschitz continuity in $\mathbb{R}^+$ for $y \mapsto \gamma r y^{1-1/r}$. In such a set-up results on the existence of unique solutions are still obtainable, but as these would require a rather lengthy discussion, we refrain from giving any details.

From the following proposition we see that the r-th power of a positive OU process X_t with $\gamma = 0$ has a representation quite similar to the one for the OU process given by

$$X_t = e^{-\lambda t} X_0 + \int_0^t \int_{\mathbb{R}^+\setminus\{0\}} e^{-\lambda(t-s)} x \mu(ds, dx).$$

Proposition 6.5. *Assume that $\gamma = 0$ and $X_0 \geqslant 0$ a.s. Then the process $Y_t = X_t^r$ can be represented as*

$$\begin{aligned} Y_t &= e^{-\lambda r t} X_0^r + \int_0^t \int_{\mathbb{R}^+\setminus\{0\}} \left((e^{-\lambda(t-s)} X_{s-} + e^{-\lambda(t-s)} x)^r - (e^{-\lambda(t-s)} X_{s-})^r\right)\mu(ds, dx) \\ &= e^{-\lambda r t} X_0^r + \int_0^t \int_{\mathbb{R}^+\setminus\{0\}} e^{-\lambda r(t-s)} \left((X_{s-}+x)^r - X_{s-}^r\right)\mu(ds, dx). \end{aligned}$$

Proof. As in the proof of Proposition 6.3 we use the auxiliary process

$$Z_t = X_0 + \int_0^t \int_{\mathbb{R}^+\setminus\{0\}} e^{\lambda s} x \mu(ds, dx).$$

For the process Z_t^r we obtain from Proposition 5.4

$$\begin{aligned} dZ_t^r &= \int_{\mathbb{R}^+\setminus\{0\}} \left((Z_{s-} + e^{\lambda s} x)^r - Z_{s-}^r\right)\mu(ds, dx) \\ &= \int_{\mathbb{R}^+\setminus\{0\}} \left((e^{\lambda s} X_{s-} + e^{\lambda s} x)^r - (e^{\lambda s} X_{s-})^r\right)\mu(ds, dx). \end{aligned}$$

Thus,

$$Z_t^r = X_0^r + \int_0^t \int_{\mathbb{R}^+\setminus\{0\}} \left((e^{\lambda s} X_{s-} + e^{\lambda s} x)^r - (e^{\lambda s} X_{s-})^r\right)\mu(ds, dx).$$

This implies the assertion via $Y_t = X_t^r = e^{-\lambda r t} Z_t^r$. ∎

Finally, let us improve the representation of Proposition 6.5 for a stationary Ornstein–Uhlenbeck process.

PROPOSITION 6.6. *Let X_t be a stationary process of Ornstein–Uhlenbeck type with driving Lévy subordinator L_t (having non-zero Lévy measure) with a vanishing drift γ. Then for $0 < r < 1$ the stationary process $Y_t = X_t^r$ can be represented as*

$$Y_t = \int_{-\infty}^{t} \int_{\mathbb{R}^+\setminus\{0\}} \left((e^{-\lambda(t-s)} X_{s-} + e^{-\lambda(t-s)} x)^r - (e^{-\lambda(t-s)} X_{s-})^r\right) \mu(ds, dx)$$

$$= \int_{-\infty}^{t} \int_{\mathbb{R}^+\setminus\{0\}} e^{-\lambda r(t-s)} \left((X_{s-}+x)^r - X_{s-}^r\right) \mu(ds, dx).$$

Proof. Note that, as in Proposition 6.5, the equality

$$Y_t = e^{-\lambda r(t-\tau)} \sqrt{X_\tau} + \int_{\tau}^{t} \int_{\mathbb{R}^+\setminus\{0\}} e^{-\lambda r(t-s)} \left((X_{s-}+x)^r - X_{s-}^r\right) \mu(ds, dx)$$

holds for all $\tau \in (-\infty, 0]$. Letting τ go to $-\infty$ we see that $e^{-\lambda r(t-\tau)} X_\tau^r$ goes to zero, since for any stationary OU process $e^{-\lambda(t-\tau)} X_\tau$ converges to zero. As, moreover, the left-hand side is independent of τ, the integral

$$\int_{\tau}^{t} \int_{\mathbb{R}^+\setminus\{0\}} e^{-\lambda r(t-s)} \left((X_{s-}+x)^r - X_{s-}^r\right) \mu(ds, dx)$$

exists for all $\tau \in (-\infty, 0]$ and is increasing for decreasing τ, the limit of the integrals for $\tau \to -\infty$ exists. This implies the result immediately. ∎

Having analysed the univariate positive Ornstein–Uhlenbeck processes in depth, let us now turn to multivariate positive definite ones and see which results can be extended. Here we state all results again only for the square root, but extensions to more general powers are immediate. The general result on the representation of the square root follows immediately from the results of Section 5.2.

PROPOSITION 6.7. *Let $(X_t)_{t\in\mathbb{R}^+}$ be an $\mathbb{S}_d^+$-valued process of Ornstein–Uhlenbeck type driven by a matrix subordinator L_t with drift $\gamma \in \mathbb{S}_d^+$ and associated Poisson random measure μ. Then the unique positive square root $Y_t = \sqrt{X_t}$ is of finite variation and has the following representation:*

$$dY_t = \mathbf{X}_{t-}^{-1}(AX_{t-} + X_{t-}A^* + \gamma)\,dt + \int_{\mathbb{S}_d^+\setminus\{0\}} (\sqrt{X_{t-}+x} - \sqrt{X_{t-}})\,\mu(dt, dx)$$

$$= \mathbf{Y}_{t-}^{-1}(AY_{t-}^2 + Y_{t-}^2 A^* + \gamma)\,dt + \int_{\mathbb{S}_d^+\setminus\{0\}} (\sqrt{Y_{t-}^2 + x} - Y_{t-})\,\mu(dt, dx),$$

provided that the process X_t is locally bounded within $\mathbb{S}_d^{++}$ or the integrals

$$\int_0^t \mathbf{X}_{s-}^{-1}(AX_{s-} + X_{s-}A^* + \gamma)\,ds \quad \text{and} \quad \int_0^t \int_{\mathbb{S}_d^+\setminus\{0\}} (\sqrt{X_{s-}+x} - \sqrt{X_{s-}})\,\mu(ds, dx)$$

exist a.s. for all $t \in \mathbb{R}$. Here, $\mathbf{X}_{t-}$ is the linear operator $Z \mapsto \sqrt{X_{t-}}\, Z + Z\sqrt{X_{t-}}$ and $\mathbf{Y}_{t-}$ the map $Z \mapsto Y_{t-} Z + Z Y_{t-}$.

For stationary OU processes one can again establish local boundedness, provided the driving Lévy process is non-degenerate.

PROPOSITION 6.8. *Let X_t be a stationary positive semidefinite Ornstein–Uhlenbeck process and assume that the driving Lévy process L_t has drift $\gamma \in S_d^{++}$ or Lévy measure ν such that $\nu(S_d^{++}) > 0$. Then the process X_t is locally bounded within S_d^{++}.*

The same holds for any positive definite Ornstein–Uhlenbeck process with initial value $X_0 \in S_d^{++}$ a.s.

Proof. In the stationary case Theorem 4.9 implies $X_0 \in S_d^{++}$ a.s. From (4.6) we thus always obtain $X_t \geqslant e^{At} X_0 e^{A^*t} \in S_d^{++}$ for all $t \in \mathbb{R}$. As $\min(\sigma(e^{At} X_0 e^{A^*t}))$ is continuous in t and strictly positive, $\min(\sigma(e^{At} X_0 e^{A^*t}))$ is locally bounded away from 0; in particular,

$$T_n := \inf\left\{t \in \mathbb{R}^+ : e^{At} X_0 e^{A^*t} < \frac{1}{n} I_d\right\}$$

defines a sequence of stopping times that a.s. increases to infinity. But this implies $X_t \geqslant n^{-1} I_d$ for all $t \in [0, T_n)$. Together with the local boundedness of X_t and the fact that sets of the form $\{x \in S_d^+ : x \geqslant \varepsilon I_d, \|x\| \leqslant K\}$ with $\varepsilon, K > 0$ are convex and compact, this establishes the local boundedness of X_t within S_d^{++}. ∎

In general we cannot obtain the existence of the relevant integrals for all positive definite OU processes, but the following proposition covers many cases of interest.

PROPOSITION 6.9. *Let X_t be a positive definite Ornstein–Uhlenbeck process driven by a matrix subordinator L_t with drift γ and Lévy measure ν. Then the integral*

$$\int_0^t \mathbf{X}_{s-}^{-1}(AX_{s-} + X_{s-}A^* + \gamma)\, ds$$

exists a.s. for all $t \in \mathbb{R}$ provided $\gamma \in S_d^{++}$ or $\gamma = 0$, $X_0 = 0$ and L_t is a compound Poisson process with $\nu(S_d^+ \backslash S_d^{++}) = 0$. Furthermore, the integral

$$\int_0^t \int_{S_d^+ \backslash \{0\}} (\sqrt{X_{s-} + x} - \sqrt{X_{s-}})\, \mu(ds, dx)$$

exists a.s. for all $t \in \mathbb{R}$, provided L_t is compound Poisson (with drift) or $\int_{0 \leqslant \|x\|_2 \leqslant 1} \sqrt{\|x\|_2}\, \nu(dx)$ is finite.

Proof. Let us first consider the second integral. Then

$$\int_{0 \leqslant \|x\|_2 \leqslant 1} \sqrt{\|x\|_2}\, \nu(dx) < \infty$$

is trivially satisfied for any compound Poisson process, and so Lemma 5.27 gives the result.

If $\gamma = 0$, $X_0 = 0$ and L_t is a compound Poisson process, $X_t = 0$ for all $t \in [0, T)$, where T denotes the first jump time of L_t. So the integral $\int_0^t \mathbf{X}_{s-}^{-1}(AX_{s-} + X_{s-}A^* + \gamma)\, ds$ exists a.s. for all $t \in [0, T)$. The condition $\nu(S_d^+ \backslash S_d^{++}) = 0$ ensures that the first jump ΔL_T is a.s. strictly positive definite, and hence $X_T \in S_d^{++}$ a.s. Using basically the same arguments as in Proposition 6.8, we infer that the integral $\int_0^t \mathbf{X}_{s-}^{-1}(AX_{s-} + X_{s-}A^* + \gamma)\, ds$ exists also a.s for all $t \in [T, \infty)$, which concludes the proof of this case.

Assume now that $\gamma \in S_d^{++}$. We have

$$X_t \geqslant \int_0^t e^{A(t-s)} \gamma e^{A^*(t-s)}\, ds \geqslant \int_0^t \min\left(\sigma\left(e^{A(t-s)} \gamma e^{A^*(t-s)}\right)\right) I_d\, ds.$$

But $e^{A(t-s)} \gamma e^{A^*(t-s)} \in S_d^{++}$ for all $t, s \in \mathbf{R}^+$, and so, for any $M \in \mathbf{R}^+$, continuity and compactness ensure the existence of a constant $k_M > 0$ such that

$$\left(\sigma\left(e^{A(t-s)} \gamma e^{A^*(t-s)}\right)\right) \geqslant k_M \quad \text{for all } t, s \in [0, M].$$

Hence $X_t \geqslant k_M t$ for all $t \in [0, M]$. Using the same matrix analytical arguments as in the proof of Corollary 5.29, this implies

$$\|\mathbf{X}_{t-}^{-1}\|_2 \leqslant \frac{1}{2\sqrt{k_M}} t^{-1/2} \quad \text{for all } t \in [0, M].$$

Moreover, as X_t is locally bounded, there is a.s. a constant K_M such that $\|X_t\|_2 \leqslant K_M$ for all $t \in [0, M]$. Here we have fixed $\omega \in \Omega$, but recall that we can argue pathwise). Since the integral

$$\int_0^t \frac{\|\mathbf{A}\|_2\, K_M + \|\gamma\|_2}{2\sqrt{k_M}} s^{-1/2}\, ds$$

is finite for all $t \in [0, M]$, where $\mathbf{A}$ is the linear operator $M_d(\mathbf{R}) \to M_d(\mathbf{R})$, $Z \mapsto AZ + ZA^*$, majorized convergence implies that the integral $\int_0^t \mathbf{X}_{s-}^{-1}(AX_{s-} + X_{s-}A^* + \gamma)\, ds$ exists a.s. for all $t \in [0, M]$. As $M \in \mathbf{R}^+$ was arbitrary, this concludes the proof. ∎

However, one can again show that the square root of a positive definite OU process X_t with $\gamma = 0$ has a representation similar to the one for the OU process given by

$$X_t = e^{At} X_0 e^{A^*t} + \int_0^t \int_{S_d^+ \backslash \{0\}} e^{A(t-s)} x e^{A^*(t-s)} \mu(ds, dx).$$

PROPOSITION 6.10. *Assume that $\gamma = 0$ and $X_0 \geqslant 0$ a.s. Then the process $Y_t = \sqrt{X_t}$ can be represented as*

$$Y_t = \sqrt{e^{At} X_0 e^{A^*t}}$$
$$+ \int_0^t \int_{S_d^+ \backslash \{0\}} \left(\sqrt{e^{A(t-s)}(X_{s-} + x) e^{A^*(t-s)}} - \sqrt{e^{A(t-s)} X_{s-} e^{A^*(t-s)}}\right) \mu(ds, dx).$$

Proof. Let $(Z_u)_{u\in\mathbf{R}^+}$ be the auxiliary process given by

$$Z_u = e^{A(t-u)} X_u e^{A^*(t-u)},$$

where $t \in \mathbf{R}^+$ is fixed. Then

$$Z_u = e^{At} X_0 e^{A^* t} + \int_0^u \int_{S_d^+\setminus\{0\}} e^{A(t-s)} x e^{A^*(t-s)} \mu(ds, dx)$$

is S_d^+-increasing. Using Proposition 5.19 and Lemma 5.26, we obtain

$$\sqrt{Z_u} = \sqrt{e^{At} X_0 e^{A^* t}} + \int_0^u \int_{S_d^+\setminus\{0\}} \left(\sqrt{Z_{s-} + e^{A(t-s)} x e^{A^*(t-s)}} - \sqrt{Z_{s-}}\right) \mu(ds, dx).$$

Since $X_t = Z_t$ and $Z_{s-} = e^{A(t-s)} X_{s-} e^{A^*(t-s)}$, this immediately concludes the proof. ∎

Finally, let us improve the above representation for a stationary positive definite Ornstein–Uhlenbeck process.

Proposition 6.11. *Let X_t be a stationary process of Ornstein–Uhlenbeck type with driving matrix subordinator L_t with a vanishing drift γ. Then the stationary process $Y_t = \sqrt{X_t}$ can be represented as*

$$Y_t = \int_{-\infty}^t \int_{S_d^+\setminus\{0\}} \left(\sqrt{e^{A(t-s)}(X_{s-}+x) e^{A^*(t-s)}} - \sqrt{e^{A(t-s)} X_{s-} e^{A^*(t-s)}}\right) \mu(dx, ds).$$

Proof. The proposition follows from Proposition 6.10 by the same arguments as in the proof of Proposition 6.6. ∎

Acknowledgements. The research leading to this paper was carried out during mutual visits to the authors' home institutions. For these visits the first author was supported by a research award from the Alexander von Humboldt Foundation and the second author was supported by the Thiele Centre for Applied Mathematics in Natural Science at the University of Århus. Moreover, financial support by the Deutsche Forschungsgemeinschaft through the graduate programme "Angewandte Algorithmische Mathematik" at the Munich University of Technology is gratefully acknowledged by the second author.

REFERENCES

[1] T. Ando, *Comparison of norms* $|||f(a)-f(b)|||$ *and* $|||f(|a-b|)|||$, Math. Z. 197 (1988), pp. 403–409.

[2] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge Stud. Adv. Math., Vol. 93, Cambridge University Press, Cambridge 2004.

[3] M. Arató, *Linear Stochastic Systems with Constant Coefficients*, Lecture Notes in Control and Inform. Sci., Vol. 45, Springer, Berlin 1982.

[4] X. Bardina and M. Jolis, *An extension of Itô's formula for elliptic diffusion processes*, Stochastic Process. Appl. 69 (1997), pp. 83–109.

[5] O. E. Barndorff-Nielsen, *Probability and statistics; selfdecomposability, finance and turbulence*, in: *Proceedings of the conference "Probability towards 2000" held at Columbia University, New York, 2–6 October 1995*, L. Accardi and C. C. Heyde (Eds.), Springer, Berlin 1998a, pp. 47–57.

[6] O. E. Barndorff-Nielsen, *Processes of normal inverse Gaussian type*, Finance Stoch. 2 (1998b), pp. 41–68.

[7] O. E. Barndorff-Nielsen, S. E. Gaversen, J. Jacod, M. Podolskij and N. Shephard, *A central limit theorem for realised power and bipower variations of continuous semimartingales*, in: *From Stochastic Calculus to Mathematical Finance — The Shiryaev Festschrift*, Y. Kabanov, R. Lipster and J. Stoyanov (Eds.), Springer, Heidelberg 2006, pp. 33–68.

[8] O. E. Barndorff-Nielsen, S. E. Graversen, J. Jacod and N. Shephard, *Limit theorems for bipower variation in financial econometrics*, Econometric Theory 22 (2006), pp. 677–719.

[9] O. E. Barndorff-Nielsen, J. Pedersen and K. Sato, *Multivariate subordination, self-decomposability and stability*, Adv. in Appl. Probab. 33 (2001), pp. 160–187.

[10] O. E. Barndorff-Nielsen and V. Pérez-Abreu, *Extensions of type G and marginal infinite divisibility*, Theory Probab. Appl. 47 (2002), pp. 202–218.

[11] O. E. Barndorff-Nielsen and V. Pérez-Abreu, *Matrix subordinators and related Upsilon transformations*, Theory Probab. Appl. (2007), to appear.

[12] O. E. Barndorff-Nielsen and J. Schmiegel, *Lévy-based tempo-spatial modelling; with applications to turbulence*, Uspekhi Mat. Nauk 59 (2004), pp. 65–91.

[13] O. E. Barndorff-Nielsen and N. Shephard, *Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial econometrics (with discussion)*, J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (2001), pp. 167–241.

[14] O. E. Barndorff-Nielsen and N. Shephard, *Integrated OU processes and non-Gaussian OU-based stochastic volatility models*, Scand. J. Statist. 30 (2003), pp. 277–295.

[15] O. E. Barndorff-Nielsen and N. Shephard, *Financial Volatility in Continuous Time*, Cambridge University Press, Cambridge 2007, to appear.

[16] R. Bhatia, *Matrix Analysis*, Grad. Texts in Math., Vol. 169, Springer, New York 1997.

[17] K. Bichteler, *Stochastic Integration with Jumps*, Encyclopedia Math. Appl., Vol. 89, Cambridge University Press, Cambridge 2002.

[18] M. S. Birman, L. S. Koplienko and M. Z. Solomjak, *Estimates of the spectrum of a difference of fractional powers of selfadjoint operators*, Izv. Vysh. Uchebn. Zaved. Mat. 154 (1975), pp. 3–10. English translation: Soviet Math. (Iz. VUZ) 19 (1975), No 3, 1–6 (1976).

[19] M.-F. Bru, *Wishart processes*, J. Theoret. Probab. 4 (1991), pp. 725–751.

[20] A. Chojnowska-Michalik, *On processes of Ornstein–Uhlenbeck type in Hilbert space*, Stochastics 21 (1987), pp. 251–286.

[21] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, CRC Financial Mathematical Series, Chapman and Hall, London 2004.

[22] J. Duda, *Cone monotone mappings: Continuity and differentiability*, Nonlinear Anal. (2007), to appear.

[23] R. Ghomrasni and G. Peskir, *Local time-space calculus and extensions of Itô's formula*, in: *High Dimensional Probability III*, J. Hoffmann-Jørgensen, M. B. Marcus and J. A. Wellner (Eds.), Progr. Probab., Vol. 55, Birkhäuser, Basel 2003, pp. 177–192.

[24] C. Gourieroux, J. Jasiak and R. Sufana, *The Wishart autoregressive process of multi-*

variate stochastic volatility, Document de travail du CREST 2004-32, CREST, Malakoff, France, 2004.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge 1990.

[26] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge 1991.

[27] F. Hubalek and E. Nicolato, *On multivariate extensions of Lévy driven Ornstein–Uhlenbeck type stochastic volatility models and multi-asset options*, 2005, in preparation.

[28] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Grundlehren Math. Wiss., Vol. 288, 2nd edition, Springer, Berlin 2003.

[29] Z. J. Jurek and D. J. Mason, *Operator-limit Distributions in Probability Theory*, Wiley, New York 1993.

[30] R. L. Karandikar, *A.s. approximation results for multiplicative stochastic integrals*, in: *Séminaire de Probabilités (Strasbourg)*, tome 16, Lecture Notes in Math. No 920, Springer, Berlin 1982a, pp. 384–391.

[31] R. L. Karandikar, *Multiplicative decomposition of non-singular matrix valued continuous semimartingales*, Ann. Probab. 10 (1982b), pp. 1088–1091.

[32] R. L. Karandikar, *Multiplicative decomposition of nonsingular matrix valued semimartingales*, in: *Séminaire de Probabilités (Strasbourg)*, tome 25, Lecture Notes in Math. No 1485, Springer, Berlin 1991, pp. 262–269.

[33] P. Lévy, *The arithmetical character of the Wishart distribution*, Proc. Cambridge Philos. Soc. 44 (1948), pp. 123–124.

[34] C.-K. Li and S. Pierce, *Linear preserver problems*, Amer. Math. Monthly 108 (2001), pp. 591–605.

[35] C.-K. Li, L. Rodman and P. Semrl, *Linear maps on selfadjoint operators preserving invertibility, positive definiteness, numerical range*, Canad. Math. Bull. 46 (2003), pp. 216–228.

[36] C. Lindberg, *Portfolio Optimization and Statistics in Stochastic Volatility Markets*, PhD thesis, Department of Mathematical Sciences, Division of Mathematical Statistics, Chalmers University of Technology and Göteborg University, Göteborg, Sweden, 2005.

[37] M. B. Marcus and J. Rosiński, *Continuity and boundedness of infinitely divisible processes: a Poisson point process approach*, J. Theoret. Probab. 18 (2005), pp. 109–160.

[38] T. Marquardt and R. Stelzer, *Multivariate CARMA processes*, Stochastic Process. Appl. 117 (2007), pp. 96–120.

[39] M. Métivier, *Semimartingales: a Course on Stochastic Processes*, de Gruyter Stud. Math., Vol. 2, Walter de Gruyter, Berlin 1982.

[40] V. Pérez-Abreu and A. Rocha-Arteaga, *Covariance-parameter Lévy processes in the space of trace-class operators*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 8 (2005), pp. 33–54.

[41] G. Peskir, *A change-of-variables formula with local time on curves*, J. Theoret. Probab. 18 (2005), pp. 499–535.

[42] S. Pierce, M. H. Lim, R. Loewy, C. K. Li, N. K. Tsing, B. McDonald and L. Beasley, *A survey of linear preserver problems*, Linear Multilinear Algebra 33 (1992), pp. 1–129.

[43] P. Protter, *Stochastic Integration and Differential Equations*, Stoch. Model. Appl. Probab., Vol. 21, 2nd edition, Springer, New York 2004.

[44] A. Rocha-Arteaga, *Subordinators in a class of Banach spaces*, Random Oper. Stochastic Equations 14 (2006), pp. 1–14.

[45] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales — Volume 2: Itô Calculus*, 2nd edition, Cambridge University Press, 2000.

[46] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, Singapore 1976.

[47] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Stud. Adv. Math., Vol. 68, Cambridge University Press, 1999.

[48] K. Sato and M. Yamazato, *Operator-selfdecomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type*, Stochastic Process. Appl. 17 (1984), pp. 73–100.

[49] H. Schneider, *Positive operators and an inertia theorem*, Numer. Math. 7 (1965), pp. 11–17.

Ole Eiler Barndorff-Nielsen
Department of Mathematical Sciences
University of Århus
Ny Munkegade
DK-8000 Århus C, Denmark
E-mail: oebn@imf.au.dk

Robert Stelzer
Graduate Programme
Applied Algorithmic Mathematics
Centre for Mathematical Sciences
Munich University of Technology
Boltzmannstraße 3
D-85747 Garching, Germany
E-mail: rstelzer@ma.tum.de
www.ma.tum.de/stat

Received on 20.6.2006