First Jump Approximation of a Lévy Driven SDE and an Application to Multivariate ECOGARCH Processes

Robert Stelzer*

The first jump approximation of a pure jump Lévy process, which converges to the Lévy process in the Skorokhod topology in probability, is generalised to a multivariate setting and an infinite time horizon. It is shown that it can generally be used to obtain "first jump approximations" of Lévy-driven stochastic differential equations, by establishing that it has uniformly controlled variations.

Applying this general result to multivariate exponential continuous time GARCH(1,1) processes, it is shown that there exists a sequence of piecewise constant processes determined by multivariate exponential GARCH$(1,1)$ processes in discrete time which converge in probability in the Skorokhod topology to the continuous time process.

AMS Subject Classification 2000: Primary: 60G51; 60H10 *Secondary:* 60H35, 91B28

Keywords: first jump approximation, Lévy process, multivariate exponential COGARCH, Skorokhod topology, stochastic differential equation, uniform tightness, uniformly controlled variations

1 Introduction

Stochastic Differential Equations (SDEs) driven by Lévy processes are widely used for stochastic modelling in various areas of applications. For simulations and analysing discretely observed data (discrete time) approximations are usually called for. Such approximations are e.g. the Euler scheme (see [6, 7, 17]), methods based on series representations ([18]) or normal approximations ([2, 21]). Recently, Szimayer and Maller [22] suggested a new approximation scheme for pure jump Lévy processes, the first jump approximation. The idea is to approximate the Lévy process on a given time grid by considering only the first jump of size greater than some minimal size and shifting this jump to the next grid point. Provided the grid size and minimal jump size converge to zero at suitable rates, the corresponding first jump approximations converge to the Lévy process in probability in the Skorokhod topology. Szimayer and Maller [22] considered this approximation scheme for univariate Lévy processes over a finite time horizon and used it to construct an approximation scheme for American option pricing. Recently the first jump approximation was used in [15] to show that for

*Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, D-85747 Garching, Germany.
Email: rstelzer@ma.tum.de, http://www-m4.ma.tum.de

a COGARCH(1,1) process, as introduced in [10], there exists a sequence of GARCH(1,1) processes converging to it in probability in the Skorokhod topology.

In this paper we extend the first jump approximation to multivariate Lévy processes and an infinite time horizon. Thereafter we show that it has uniformly controlled variations in the sense of [11, 13]. This property is in our context equivalent to uniform tightness and implies that one obtains an approximation of the solution of a pure jump Lévy driven SDE when replacing the Lévy process with its first jump approximation. The solutions of the approximating SDEs again converge in the Skorokhod topology in probability.

Furthermore, we use the first jump approximation of Lévy driven SDEs to derive an interesting property of a concrete model, viz. multivariate exponential continuous time GARCH(1,1) (henceforth ECOGARCH(1,1)) processes (see [4, 5]). These processes form a multivariate stochastic volatility model and are thus of particular interest regarding the modelling of financial data. We show that for a multivariate ECOGARCH(1,1) process there exists a sequence of piecewise constant processes determined by multivariate exponential GARCH(1,1) (henceforth EGARCH(1,1)) processes which converge to the ECOGARCH(1,1) process in the Skorokhod topology in probability. The importance of this result is that it provides a link between discrete and continuous time modelling and a possibility for the estimation of ECOGARCH(1,1) processes like for COGARCH(1,1) processes in [15], where a similar convergence result for COGARCH(1,1) processes was obtained using a tailor-made approach instead of general results on SDEs.

This paper is organised as follows. In the next Section 2 we summarise notation and some preliminaries regarding Lévy processes and convergence in the Skorokhod topology. Section 3 discusses the first jump approximation of Lévy driven SDEs. Finally, in Section 4 we first review discrete time multivariate EGARCH(1,1) processes and multivariate ECOGARCH(1,1) processes briefly. Thereafter it is shown that an ECOGARCH(1,1) process can be approximated arbitrarily well in the Skorokhod topology in probability by piece-wise constant processes determined by discrete time EGARCH(1,1) processes.

2 Preliminaries

Before presenting our results we summarise below the notation to be used and some basic facts on Lévy processes and convergence in the Skorokhod space.

2.1 Notation

Throughout this paper we write $\mathbb{R}^+$ for the positive real numbers including zero, $\mathbb{R}^{++}$ when zero is excluded. Moreover, we denote the set of real $d \times m$ matrices by $M_{d,m}(\mathbb{R})$, the $d \times d$ matrices by $M_d(\mathbb{R})$, the group of invertible $d \times d$ matrices by $GL_d(\mathbb{R})$, the linear subspace of symmetric matrices by $\mathbb{S}_d$, the (closed) positive semi-definite cone by $\mathbb{S}_d^+$ and the open (in $\mathbb{S}_d$) positive definite cone by $\mathbb{S}_d^{++}$. I_d stands for the $d \times d$ identity matrix, $\det(A)$ for the determinant and $\sigma(A)$ for the spectrum (the set of all eigenvalues) of a matrix $A \in M_d(\mathbb{R})$. Moreover, $\mathrm{vech} : \mathbb{S}_d \to \mathbb{R}^{d(d+1)/2}$ denotes the "vector-half" operator that stacks the columns of the lower triangular part of a symmetric matrix below another. Finally, A^* is the adjoint of a matrix $A \in M_d(\mathbb{R})$.

Norms of vectors and matrices are denoted by $\|\cdot\|$. If the norm is not specified then it is irrelevant which particular norm is used.

The exponential of a matrix A is denoted by $\exp(A)$ or e^A and the unique positive semi-definite square root of a matrix $A \in \mathbb{S}_d^+$ by $A^{1/2}$.

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 $(\Omega, \mathscr{F}, P, (\mathscr{F}_t)_{t\in\mathbb{R}^+})$ satisfying the usual hypotheses (complete and right continuous filtration). L^p denotes as usual the space of all random variables with a finite p-th moment, i.e. all random variables X with $E(\|X\|^p) < \infty$ in a multivariate setting.

The indicator function of some set A is denoted by 1_A.

2.2 Multivariate Lévy processes

Now we state some elementary properties of multivariate Lévy processes that will be needed. For a more general treatment and proofs we refer to [1, 16, 20].

We consider a pure jump Lévy process $L = (L_t)_{t\in\mathbb{R}^+}$ (where $L_0 = 0$ a.s.) in $\mathbb{R}^d$ determined by its characteristic function $E\left[e^{i\langle u,L_t\rangle}\right] = \exp\{t\psi_L(u)\},\ t \geq 0$, in the Lévy-Khintchine form where

$$\psi_L(u) = i\langle\gamma_L, u\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle 1_{[0,1]}(\|x\|)\right)\nu_L(dx) \text{ for } u \in \mathbb{R}^d \tag{2.1}$$

with $\gamma_L \in \mathbb{R}^d$ and ν_L being a measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ that satisfies $\nu_L(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(\|x\|^2 \wedge 1)\,\nu_L(dx) < \infty$. The measure ν_L is referred to as the Lévy measure of L. Here $\|\cdot\|$ may be any fixed norm. Different norms simply correspond to different truncation functions and thus for a given Lévy process γ_L changes when the norm is changed. Implicitly we presume throughout this paper that given a Lévy process γ_L is set to the value such that (2.1) holds with the currently employed norm.

It is a well-known fact that to every càdlàg Lévy process L on $\mathbb{R}^d$ one can associate a random measure N_L on $\mathbb{R}^+ \times \mathbb{R}^d \setminus \{0\}$ describing the jumps of L (see e.g. [8, Section II.1]). For any measurable set $B \subset \mathbb{R}^+ \times \mathbb{R}^d \setminus \{0\}$,

$$N_L(B) = \sharp\{s \geq 0 :\ (s, L_s - L_{s-}) \in B\}.$$

The jump measure N_L is an extended Poisson random measure (as defined in [8, Definition II.1.20]) on $\mathbb{R}^+ \times \mathbb{R}^d \setminus \{0\}$ with intensity measure $n_L(ds,dx) = ds\,\nu_L(dx)$. By the Lévy-Itô decomposition we can rewrite L almost surely as

$$L_t = \gamma_L t + \int_{\|x\|\geq 1}\int_0^t x N_L(ds,dx) + \lim_{\varepsilon\downarrow 0}\int_{\varepsilon\leq\|x\|\leq 1}\int_0^t x\tilde{N}_L(ds,dx). \tag{2.2}$$

for every $t \geq 0$. Here $\tilde{N}_L(ds,dx) = N_L(ds,dx) - ds\nu_L(dx)$ is the compensated jump measure, the terms in (2.2) are independent and the convergence in the last term is a.s. and locally uniform in $t \geq 0$.

Assuming that ν_L satisfies additionally $\int_{\|x\|>1}\|x\|^2\,\nu_L(dx) < \infty$, L has finite mean and covariance matrix Σ_L given by $\Sigma_L = C_L + \int_{\mathbb{R}^d} xx^*\,\nu_L(dx)$.

For the theory of stochastic integration and SDEs (with respect to Lévy processes and/or random measures) we refer to any of the standard texts, e.g. [1, 8, 16].

2.3 Convergence in the Skorokhod topology

For a complete and separable normed space $(E, \|\cdot\|_E)$ denote by D_E the set of all functions $f : \mathbb{R}^+ \to E$ that are right continuous and have left limits. Let further Λ be the set of all time change functions, i.e. all continuous and strictly increasing functions $\lambda : \mathbb{R}^+ \to \mathbb{R}^+$ that satisfy $\lambda(0) = 0$ and $\lim_{t\to\infty}\lambda(t) = \infty$. Denote by $e_{\mathbb{R}^+}$ the function $\mathbb{R}^+ \to \mathbb{R}^+, t \mapsto t$ and by $\|x\|_{E,[a,b)} := \sup_{t\in[a,b)}\|x(t)\|_E$ for $x \in D_E$. A sequence $(x_n)_{n\in\mathbb{N}}$ in D_E converges to $x \in D_E$ in the Skorokhod topology if there exists a sequence

$(\lambda_n)_{n\in\mathbb{N}}$ in Λ such that $(\lambda_n)_{n\in\mathbb{N}}$ converges uniformly to $e_{\mathbb{R}^+}$, i.e. $\lim_{n\to\infty}\|\lambda_n - e_{\mathbb{R}^+}\|_{\mathbb{R}^+,[0,\infty)} = 0$, and $(x_n \circ \lambda_n)_{n\in\mathbb{N}}$ converges uniformly on compacts to x, i.e. $\lim_{n\to\infty}\|x_n \circ \lambda_n - x\|_{E,[0,N]} = 0$ for all $N \in \mathbb{N}$. A separable (yet not complete) metric inducing the Skorokhod topology is given by

$$d_E(x,y) = \inf_{\lambda\in\Lambda} d_{\lambda,E}(x,y) \text{ for } x,y \in D_E \tag{2.3}$$

where

$$d_{\lambda,E}(x,y) = \|\lambda - e_{\mathbb{R}^+}\|_{\mathbb{R}^+,[0,\infty)} + \sum_{n=1}^{\infty}\frac{1}{2^n}\min\left(1, \sup_{t\le n}\|x(\min(n,\lambda(t))) - y(\min(n,t))\|_E\right),$$

see e.g. [13]. For equivalent norms on E the above definition leads to equivalent metrics on D_E. Convergence of a sequence $(X^{(n)})_{n\in\mathbb{N}}$ of E-valued càdlàg random processes, i.e. random variables in D_E, in probability in the Skorokhod topology to a càdlàg random process X means for us in the following that $\operatorname{plim}_{n\to\infty} d_E(X^{(n)},X) = 0$ with plim denoting the limit in probability. Moreover, we note that uniform convergence on compacts in probability (ucp convergence, cf. [16, Chapter II.4]) obviously implies convergence in the Skorokhod topology in probability and that in a metric space convergence in probability is metrisable (follows e.g. by replacing the absolute value with the metric in [14, p. 160 (3), p.175 9.]).

For a more comprehensive introduction to the Skorokhod topology we refer the reader to [3] or [8, Chapter VI], for instance.

3 First jump approximation of Lévy driven SDEs

In this section we present our main result, the first jump approximation of Lévy driven SDEs. We start by giving a d-dimensional and infinite time extension of the first jump approximation for a Lévy process presented in [22] and an interesting refinement of it for a Lévy process with zero mean. Below the infimum over an empty set is taken to be ∞ as usual.

Theorem 3.1. *Let L be a d-dimensional pure jump Lévy process with drift γ_L and Lévy measure ν_L. Further let $(m^{(n)})_{n\in\mathbb{N}}$ be a positive sequence, which is bounded by* 1 *and monotonically decreases to* 0*, and $(t_i^{(n)})_{i\in\mathbb{N}_0}$ be for each $n \in \mathbb{N}$ a strictly increasing sequence with $t_0^{(n)} = 0$ and $\lim_{i\to\infty} t_i^{(n)} = \infty$. Setting $\delta^{(n)} := \sup_{i\in\mathbb{N}}\left\{t_i^{(n)} - t_{i-1}^{(n)}\right\}$, assume further that $\lim_{n\to\infty}\delta^{(n)} = 0$ and*

$$\lim_{n\to\infty}\delta^{(n)}\left(\nu_L\left(J^{(n)}\right)\right)^2 = 0, \tag{3.1}$$

where $J^{(n)} := \{x \in \mathbb{R}^d : \|x\| > m^{(n)}\}$.

(a) *Define for all $n \in \mathbb{N}$*

$$\gamma^{(n)} := \gamma_L - \int_{J^{(n)}\cap\{\|x\|\le 1\}} x\nu_L(dx),$$

$$\tau_i^{(n)} := \inf\{t : t_{i-1}^{(n)} < t \le t_i^{(n)}, \|\Delta L_t\| \in J^{(n)}\} \quad \text{for all } i \in \mathbb{N},$$

$$\tilde{L}_t^{(n)} := \gamma^{(n)} t + \sum_{i\in\mathbb{N}:\,\tau_i^{(n)}\le t} \Delta L_{\tau_i^{(n)}} \quad \text{for } t \in \mathbb{R}^+, \quad \tilde{L}^{(n)} := (\tilde{L}_t^{(n)})_{t\in\mathbb{R}^+},$$

$$\bar{L}_t^{(n)} := \tilde{L}_{t_{i-1}^{(n)}}^{(n)}, \quad \text{for all } t \in [t_{i-1}^{(n)}, t_i^{(n)}),\ i \in \mathbb{N}, \quad \bar{L}^{(n)} := (\bar{L}_t^{(n)})_{t\in\mathbb{R}^+}.$$

Then it holds that

$$\tilde{L}^{(n)} \to L \text{ in ucp as } n \to \infty \tag{3.2}$$

and

$$\mathrm{plim}_{n\to\infty} d_{\mathbb{R}^d}(\bar{L}^{(n)}, L) = 0. \tag{3.3}$$

(b) *If L has finite expectation and $E(L_1) = 0$, define $L^{(n)} := (L_t^{(n)})_{t\in\mathbb{R}^+}$ by setting*

$$L_t^{(n)} := \sum_{i\in\mathbb{N}:\, t_i^{(n)}\le t} \left(1_{(0,\infty)}(\tau_i^{(n)})\Delta L_{\tau_i^{(n)}} - \frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \int_{J^{(n)}} x\nu_L(dx) \right).$$

Then

$$\mathrm{plim}_{n\to\infty} d_{\mathbb{R}^d}(L^{(n)}, L) = 0. \tag{3.4}$$

and $E\left(L_t^{(n)}\right) = 0$ for all $t \in \mathbb{R}^+$ as well as $E\left(\Delta L^{(n)}_{t_i^{(n)}}\right) = 0$ for all $i \in \mathbb{N}$ and each $n \in \mathbb{N}$.

(c) *Provided $E(\|L_1\|^2) < \infty$, it holds that $E(\|\tilde{L}_t^{(n)}\|^2)$, $E(\|\bar{L}_t^{(n)}\|^2)$, $E(\|L_t^{(n)}\|^2)$ and $E(\|\Delta\tilde{L}_t^{(n)}\|^2)$, $E(\|\Delta\bar{L}_t^{(n)}\|^2)$, $E(\|\Delta L_t^{(n)}\|^2)$ are finite for all $t \in \mathbb{R}^+$ and $n \in \mathbb{N}$.*

The definition of $\tau_i^{(n)}$ above means that it is the first time in the grid interval $(t_{i-1}^{(n)}, t_i^{(n)}]$ at which L has a jump bigger than $m^{(n)}$ in norm. If there is no such jump, $\tau_i^{(n)} = \infty$. $\tilde{L}^{(n)}$ approximates the Lévy process L by a drift and the first jumps of size greater than $m^{(n)}$ in the grid intervals. Since the jumps are left at their original time, we obtain ucp convergence. In the approximation $\bar{L}^{(n)}$ both these jumps and the increment caused by the drift are shifted to the grid points, so that $\bar{L}^{(n)}$ is constant in between the grid times. Due to shifting the jumps we obtain only convergence in the Skorokhod topology. Finally, the approximation $L^{(n)}$ is a modification of $\bar{L}^{(n)}$ when L has a finite and vanishing first moment. It ensures that also the approximation has a vanishing mean, as it will often be desirable to reproduce this property of L.

Proof. (a) An inspection of the proof of [22, Theorem 3.1] shows that it immediately generalises to our multivariate set-up with no upper bound on the jump sizes and no binning of the jump sizes. This proves (3.2) which implies

$$\mathrm{plim}_{n\to\infty} d_{\mathbb{R}^d}(\tilde{L}^{(n)}, L) = 0.$$

Using the time change λ which is the obvious extension to $[0,\infty)$ of the time change employed in the proof of [22, Theorem 3.2] and arguments analogous to theirs give $d_{\mathbb{R}^d}(\tilde{L}^{(n)}, \bar{L}^{(n)}) \le d_{\lambda,\mathbb{R}^d}(\tilde{L}^{(n)}, \bar{L}^{(n)}) \le \delta^{(n)} + o(\sqrt{\delta^{(n)}}) \to 0$ as $n \to \infty$ a.s. The triangle inequality thus establishes (3.3).

(b) Assume now $E(L_1) = 0$. Then $E\left(L_t^{(n)}\right) = 0$ for all $t \in \mathbb{R}^+$ as well as $E\left(\Delta L^{(n)}_{t_i^{(n)}}\right) = 0$ for all $i \in \mathbb{N}$ and each $n \in \mathbb{N}$ follows from

$$E\left(1_{(0,\infty)}(\tau_i^{(n)})\Delta L_{\tau_i^{(n)}}\right) = \frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \int_{J^{(n)}} x\nu_L(dx).$$

Since $E(L_1) = 0$, we have $\gamma_L = -\int_{\|x\|>1} x\nu_L(dx)$. Hence, straightforward calculations give for all $t \in \mathbb{R}^+$

$$\begin{aligned}\|\bar{L}_t^{(n)} - L_t^{(n)}\| &= \left\| \sum_{i\in\mathbb{N}:\, t_i^{(n)}\le t} \left(\frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \int_{J^{(n)}} x\nu_L(dx) - \left(t_i^{(n)} - t_{i-1}^{(n)}\right) \int_{J^{(n)}} x\nu_L(dx) \right) \right\| \\ &\le tC^{(n)},\end{aligned}$$

where $C^{(n)} = \sqrt{\delta^{(n)}}\int_{J^{(n)}} \|x\|\nu_L(dx)\sum_{k=1}^{\infty} \frac{\left(\nu_L(J^{(n)})\sqrt{\delta^{(n)}}\right)^k (\delta^{(n)})^{(k-1)/2}}{k+1!}$.
Since (3.1) and $E(\|L_1\|) < \infty$ imply that $\sqrt{\delta^{(n)}}\int_{J^{(n)}} \|x\|\nu_L(dx)$ and $\nu_L(J^{(n)})\sqrt{\delta^{(n)}}$ converge to zero as $n \to \infty$, we have $\lim_{n\to\infty} C^{(n)} = 0$ and hence $\lim_{n\to\infty} \|\bar{L}^{(n)} - L^{(n)}\|_{\mathbb{R}^d,[0,T]} = 0$ a.s. for all $T \in \mathbb{R}^+$. Combining this with (3.3) shows (3.4).

Finally, (c) is easily seen using

$$E\left(1_{(0,\infty)}(\tau_i)\Delta L_{\tau_i^{(n)}}\left(\Delta L_{\tau_i^{(n)}}\right)^*\right) = \frac{1 - e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)} - t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \int_{J^{(n)}} xx^*\nu_L(dx),$$

as $E(\|L_1\|^2) < \infty$ is equivalent to the finiteness of $\int_{\mathbb{R}^d} \|x\|^2\nu_L(dx)$ or of all elements of $\int_{\mathbb{R}^d} xx^*\nu_L(dx)$. □

The next Lemma shows that when one of the sequences $(\delta^{(n)})_{n\in\mathbb{N}}$ or $(m^{(n)})_{n\in\mathbb{N}}$ is given one can always choose the other one such that (3.1) holds.

Lemma 3.2. *Let L be a Lévy process in $\mathbb{R}^d$. Assume that $(\delta^{(n)})_{n\in\mathbb{N}}$ is a monotonically decreasing sequence in $\mathbb{R}^{++}$ with $\lim_{n\to\infty}\delta^{(n)} = 0$ or $(m^{(n)})_{n\in\mathbb{N}}$ is a monotonically decreasing sequence in $\mathbb{R}^{++}$ with $\lim_{n\to\infty} m^{(n)} = 0$ and $m^{(n)} \le 1\,\forall n \in \mathbb{N}$, respectively. Then a monotonically decreasing sequence $(m^{(n)})_{n\in\mathbb{N}}$ in $\mathbb{R}^{++}$ with $\lim_{n\to\infty} m^{(n)} = 0$ and $m^{(n)} \le 1\,\forall n \in \mathbb{N}$ or a monotonically decreasing sequence $(\delta^{(n)})_{n\in\mathbb{N}}$ in $\mathbb{R}^{++}$ with $\lim_{n\to\infty}\delta^{(n)} = 0$, respectively, can be chosen such that (3.1) is satisfied for all norms $\|\cdot\|$.*

Proof. We have that $\int_{\mathbb{R}^d} \min(\|x\|^2, 1)\nu_L(dx) < \infty$ for all norms $\|\cdot\|$, because ν_L is a Lévy measure. Thus $\lim_{n\to\infty}\delta^{(n)}\nu_L(J^{(n)}\backslash U_1(0)) = 0$ for all norms where $U_1(0)$ denotes the open ball around zero with radius 1 and $J^{(n)} = \{x \in \mathbb{R}^d : \|x\| > m^{(n)}\}$. Moreover,

$$m(n)^2\nu_L(J^{(n)} \cap U_1(0)) \le \int_{J^{(n)}\cap U_1(0)} \|x\|^2\nu_L(dx) \le \int_{U_1(0)} \|x\|^2\nu_L(dx) < \infty.$$

This implies that (3.1) holds if $\delta^{(n)} = o\left(\left(m^{(n)}\right)^4\right)$. Hence, the lemma is shown by choosing for instance $m^{(n)} = (\delta^{(n)})^{1/5} \wedge 1$ or $\delta^{(n)} = (m^{(n)})^5$, respectively, for all $n \in \mathbb{N}$. □

Remark 3.3. *As an inspection of our proofs shows, the sets $J^{(n)}$ need not be of the form $\{x : \|x\| > m^{(n)}\}$. All our results remain valid for arbitrary increasing sequences $(J^{(n)})_{n\in\mathbb{N}}$ of measurable sets with $\cup_{n=1}^{\infty} J^{(n)} = \mathbb{R}^d\backslash\{0\}$, $\inf\{\|x\| : x \in J^{(n)}\} > 0$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty}\inf\{\|x\| : x \in J^{(n)}\} = 0$.*

Next we show that the first jump approximation of a Lévy process has uniformly controlled variations (UCV) which is an important property regarding the convergence of stochastic integrals and solutions to stochastic differential equations (cf. [11, 13]). Equivalently one could use a condition called "uniform tightness" (cf. [8, Section VI.6] or [11, 13]).

In the following we need to transform any semi-martingale to one with bounded jumps in a suitable way. To this end we define for a semi-martingale Z and $\kappa \in \mathbb{R}^{++}\cup\{\infty\}$ the semi-martingale $Z^{[\kappa]}$ by setting $Z_t^{[\kappa]} = Z_t - \sum_{0<s\le t} r_\kappa(\Delta Z_s)$ with $r_\kappa(z) := \max(0, 1 - \kappa/\|z\|)z$ for finite κ and $Z^{[\infty]} = Z$. Furthermore, for a finite variation process A we denote by $(TV(A)_t)_{t\in\mathbb{R}^+}$ the process giving the total variation of A over the interval $[0,t]$ for $t \in \mathbb{R}^+$ and for a d-dimensional martingale $M = (M_1, M_2, \ldots, M_d)^*$ the quadratic variation $[M,M]$ is understood to be defined by $[M,M]_t = \sum_{i=1}^{d} [M_i, M_i]_t$.

Definition 3.4 (Kurtz and Protter [11]). *Let $(Z^{(n)})_{n\in\mathbb{N}}$ be a sequence of semi-martingales in $\mathbb{R}^d$ each defined on its own filtered probability space $(\Omega^{(n)},\mathscr{F}^{(n)},P^{(n)},(\mathscr{F}_t^{(n)})_{t\in\mathbb{R}^+})$ satisfying the usual hypothesis. If there exists a $\kappa\in\mathbb{R}^{++}\cup\{\infty\}$ such that for each $\alpha>0$ and $n\in\mathbb{N}$ there exist $(\mathscr{F}_t^{(n)})$-local martingales $M^{(n)}$, $(\mathscr{F}_t^{(n)})$-adapted finite variation processes $A^{(n)}$ in $\mathbb{R}^d$ and $(\mathscr{F}_t^{(n)})$-stopping times $T^{(n,\alpha)}$ satisfying $(Z^{(n)})^{[\kappa]}=M^{(n)}+A^{(n)}$, $P^{(n)}(T^{(n,\alpha)}\le\alpha)\le(1/\alpha)$ and*

$$\sup_{n\in\mathbb{N}}E_{P^{(n)}}\left([M^{(n)},M^{(n)}]_{\min(t,T^{(n,\alpha)})}+TV(A^{(n)})_{\min(t,T^{(n,\alpha)})}\right)<\infty$$

for all $t\in\mathbb{R}^+$ then the sequence $(Z^{(n)})_{n\in\mathbb{N}}$ is said to have uniformly controlled variations (UCV).

In the following our processes are defined on the same probability space, but the filtrations are different.

Theorem 3.5. *Let L be a d-dimensional Lévy process without a Brownian part and $(\bar{L}^{(n)})_{n\in\mathbb{N}}$ the first jump approximation of Theorem 3.1 (a). Let $(\mathscr{F}_t^{(n)})_{t\in\mathbb{R}}$ be for each $n\in\mathbb{N}$ the completed filtration generated by $\bar{L}^{(n)}$. Then the usual conditions are satisfied and $\bar{L}^{(n)}$ is for each $n\in\mathbb{N}$ a semi-martingale on $(\Omega,\mathscr{F},P,(\mathscr{F}_t^{(n)})_{t\in\mathbb{R}})$. Moreover, $(\bar{L}^{(n)})_{n\in\mathbb{N}}$ has UCV.*

If L has finite mean and $E(L_1)=0$, let $L^{(n)}$ be the first jump approximation of Theorem 3.1 (b). Then $L^{(n)}$ is for each $n\in\mathbb{N}$ a martingale on $(\Omega,\mathscr{F},P,(\mathscr{F}_t^{(n)})_{t\in\mathbb{R}})$ and $(L^{(n)})_{n\in\mathbb{N}}$ has UCV.

Proof. (i) Since $\bar{L}^{(n)}$ is piecewise constant, it is clear that $(\mathscr{F}_t^{(n)})_{t\in\mathbb{R}}$ is right continuous. Thus the usual conditions are satisfied. The semi-martingale property is also immediate. To see that also $L^{(n)}$ is a semi-martingale with respect to this filtration, provided $E(L_1)=0$, it suffices to note that $L^{(n)}-\bar{L}^{(n)}$ is a deterministic process of finite variation on compacts. That $L^{(n)}$ is even a martingale is then straightforward as the jumps have zero expectation and are independent of the past.

(ii) We now show UCV for $(\bar{L}^{(n)})_{n\in\mathbb{N}}$.

Choose $\kappa\in(2,\infty)$ such that $\kappa/2>\sup_{n\in\mathbb{N}}\left\{\delta^{(n)}\|\gamma^{(n)}\|\right\}+1$ and

$$\kappa/2>\sup_{i,n\in\mathbb{N}}\left\{\left\|\frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)}\int_{J_0^{(n)}}x\nu_L(dx)\right\|\right\}$$

where $J_0^{(n)}:=J^{(n)}\cap\{x:\|x\|\le1\}$.

The finiteness of the first supremum is a consequence of (3.1) and the finiteness of the second one follows from

$$\left\|\frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)}\int_{J_0^{(n)}}x\nu_L(dx)\right\|\le\sqrt{\delta^{(n)}}\nu_L(J_0^{(n)})\sum_{k=0}^{\infty}\frac{\left(\sqrt{\delta^{(n)}}\nu_L(J^{(n)})\right)^k(\delta^{(n)})^{(k+1)/2}}{k+1!}$$

for all $i,n\in\mathbb{N}$, since the right hand side goes to zero as $n\to\infty$.

Define $(M^{(n)})_{t\in\mathbb{R}^+}$ by

$$M_t^{(n)}=\sum_{i\in\mathbb{N}:\,t_i^{(n)}\le t}\left(1_{(0,\infty)}(\tau_i^{(n)})1_{(0,1]}(\|\Delta L_{\tau_i^{(n)}}\|)\Delta L_{\tau_i^{(n)}}-\frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)}\int_{J_0^{(n)}}x\nu_L(dx)\right)$$

and $(A^{(n)})_{t\in\mathbb{R}^+}$ by

$$A_t^{(n)} = \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} \left(\tilde{r}_\kappa \left(1_{(0,\infty)}(\tau_i^{(n)}) 1_{(1,\infty)}(\|\Delta L_{\tau_i^{(n)}}\|) \Delta L_{\tau_i^{(n)}} + \gamma^{(n)} \left(t_i^{(n)} - t_{i-1}^{(n)} \right) \right) \right.$$
$$\left. + \frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \int_{J_0^{(n)}} x\nu_L(dx) \right),$$

where $\tilde{r}_\kappa(x) = x - r_\kappa(x)$. Then $M^{(n)}$ is a $(\mathscr{F}_t^{(n)})$-martingale with expectation zero and $A^{(n)}$ an $(\mathscr{F}_t^{(n)})$-adapted finite variation process for all $n\in\mathbb{N}$. By the choice of κ we have

$$(L^{(n)})^{[\kappa]} = M^{(n)} + A^{(n)}.$$

Since $[M^{(n)},M^{(n)}]_t = \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} \left(\Delta M_{t_i^{(n)}}^{(n)}\right)^* \Delta M_{t_i^{(n)}}^{(n)}$, it follows that

$$\begin{aligned}
E\left([M^{(n)},M^{(n)}]_t\right) &= \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} E\left(1_{(0,\infty)}(\tau_i^{(n)}) 1_{(0,1]}(\|\Delta L_{\tau_i^{(n)}}\|)(\Delta L_{\tau_i^{(n)}})^* \Delta L_{\tau_i^{(n)}}\right) \\
&\quad - \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} \left(\frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \right)^2 \int_{J_0^{(n)}} x^*\nu_L(dx) \int_{J_0^{(n)}} x\nu_L(dx) \\
&\le \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} E\left(1_{(0,\infty)}(\tau_i^{(n)}) 1_{(0,1]}(\|\Delta L_{\tau_i^{(n)}}\|)(\Delta L_{\tau_i^{(n)}})^* \Delta L_{\tau_i^{(n)}}\right) \\
&= \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} \frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \int_{J_0^{(n)}} x^* x\nu_L(dx).
\end{aligned}$$

Denoting by $\|\cdot\|_2$ the Euclidean norm and using the elementary inequality $1-e^{-x}\le x$ for all $x\in\mathbb{R}^+$, this implies

$$\sup_{n\in\mathbb{N}} E([M^{(n)},M^{(n)}]_t) \le t \int_{\|x\|\le 1} \|x\|_2^2 \nu_L(dx) < \infty \tag{3.5}$$

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

Turning to $A^{(n)}$ we have that

$$\begin{aligned}
\|\Delta A_{t_i^{(n)}}^{(n)}\| \le 3 &\left\| \tilde{r}_\kappa \left(1_{(0,\infty)}(\tau_i^{(n)}) 1_{(1,\infty)}(\|\Delta L_{\tau_i^{(n)}}\|) \Delta L_{\tau_i^{(n)}} \right) \right\| + \|\gamma_L\| \left(t_i^{(n)} - t_{i-1}^{(n)} \right) \\
&+ \left\| \left(\frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} - \left(t_i^{(n)}-t_{i-1}^{(n)}\right) \right) \int_{J_0^{(n)}} x\nu_L(dx) \right\|
\end{aligned}$$

for all $i,n\in\mathbb{N}$. The inequality is immediate provided

$$\left\| 1_{(0,\infty)}(\tau_i^{(n)}) 1_{(1,\infty)}(\|\Delta L_{\tau_i^{(n)}}\|) \Delta L_{\tau_i^{(n)}} + \gamma^{(n)} \left(t_i^{(n)} - t_{i-1}^{(n)} \right) \right\| \le \kappa \qquad \text{and}$$
$$\left\| 1_{(0,\infty)}(\tau_i^{(n)}) 1_{(1,\infty)}(\|\Delta L_{\tau_i^{(n)}}\|) \Delta L_{\tau_i^{(n)}} \right\| \le \kappa.$$

Otherwise the choice of κ ensures $\left\| 1_{(0,\infty)}(\tau_i^{(n)}) 1_{(1,\infty)}(\|\Delta L_{\tau_i^{(n)}}\|)\Delta L_{\tau_i^{(n)}} \right\| > \kappa/2$ and that $\left\| \Delta A^{(n)}_{t_i^{(n)}} \right\| \leq (3/2)\kappa$ which implies the validity of the inequality. We have

$$\left\| \left(\frac{1 - e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)} - t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} - \left(t_i^{(n)} - t_{i-1}^{(n)}\right) \right) \int_{J_0^{(n)}} x \nu_L(dx) \right\| \leq C_A^{(n)} \left(t_i^{(n)} - t_{i-1}^{(n)}\right)$$

with $C_A^{(n)} := \sqrt{\delta^{(n)}} \nu_L(J_0^{(n)}) \sum_{k=1}^{\infty} \frac{\left(\nu_L(J^{(n)})\sqrt{\delta^{(n)}}\right)^k (\delta^{(n)})^{(k-1)/2}}{k+1!}$, which converges to zero as $n \to \infty$. Hence,

$$TV(A^{(n)})_t \leq 3 \int_0^t \int_{\|x\|>1} \min(\|x\|, \kappa) N_L(ds, dx) + t \left(\|\gamma_L\| + \sup_{n \in \mathbb{N}} \left\{ C_A^{(n)} \right\} \right)$$

and thus

$$\sup_{n \in \mathbb{N}} E\left(TV(A^{(n)})_t \right) \leq t \left(3 \int_{\|x\|>1} \min(\|x\|, \kappa) \nu_L(dx) + \|\gamma_L\| + \sup_{n \in \mathbb{N}} \left\{ C_A^{(n)} \right\} \right) < \infty \, \forall t \in \mathbb{R}^+. \tag{3.6}$$

Combining (3.5) and (3.6) and choosing $T^{(n,\alpha)} = \alpha + 1$ for all $n \in \mathbb{N}$ and $\alpha \in \mathbb{R}^{++}$ shows that $(\bar{L}^{(n)})_{n \in \mathbb{N}}$ has UCV.

(iii) It remains to verify that $(L^{(n)})_{n \in \mathbb{N}}$ has UCV, provided L has a finite expectation and $E(L_1) = 0$. The arguments presented in the proof of Theorem 3.1 (b) imply that

$$TV(L^{(n)} - \bar{L}^{(n)})_t \leq C^{(n)} t.$$

for all $n \in \mathbb{N}$ and $t \in \mathbb{R}^+$ with $\lim_{n \to \infty} C^{(n)} = 0$. This immediately shows that $(L^{(n)} - \bar{L}^{(n)})_{n \in \mathbb{N}}$ has UCV. As $(L^{(n)} - \bar{L}^{(n)})_{n \in \mathbb{N}}$ and $(\bar{L}^{(n)})_{n \in \mathbb{N}}$ converge in probability in the Skorokhod topology, [13, Theorem 7.6] ensures that both sequences are uniformly tight. Property 6.4 on page 377 of [8] therefore implies that $(L^{(n)})_{n \in \mathbb{N}}$ is uniformly tight and so [13, Theorem 7.6] shows that it has UCV. □

The UCV property has far reaching implications regarding the convergence of sequences of stochastic differential equations and of stochastic integrals in the Skorokhod topology in probability or weakly. These follow from the general results to be found in [11, 13] or [8, Sections VI.6, IX.6]. To avoid very lengthy technical statements we consider only the following simple setting, since in the general case one only allows for more general coefficients (e.g. locally Lipschitz ones or functional Lipschitz ones depending on the whole past) in the SDEs which necessitates very technical conditions for them, whereas the main condition on the driving processes remains that they have UCV. Note, however, that the generalisation to more general coefficients is straightforward.

Theorem 3.6. *(a) Let L be a d-dimensional pure jump Lévy process, $\bar{L}^{(n)}$ its first jump approximation given in Theorem 3.1 (a) and $f : \mathbb{R}^m \to M_{m,d}(\mathbb{R})$ a globally Lipschitz function. Moreover, assume given a sequence of m-dimensional càdlàg processes $Z^{(n)}$ adapted to $(\mathscr{F}_t^{(n)})_{t \in \mathbb{R}^+}$ and an m-dimensional càdlàg process Z adapted to $(\mathscr{F}_t)_{t \in \mathbb{R}^+}$ such that*

$$\operatorname{plim}_{n \to \infty} d_{\mathbb{R}^{d+m}} \left(\begin{pmatrix} \bar{L}^{(n)} \\ Z^{(n)} \end{pmatrix}, \begin{pmatrix} \bar{L} \\ Z \end{pmatrix} \right) = 0.$$

Denote by $X^{(n)}$ for each $n \in \mathbb{N}$ the unique càdlàg solution to

$$X_t^{(n)} = Z_t^{(n)} + \int_0^t f(X_{s-}^{(n)}) d\bar{L}_s^{(n)} \tag{3.7}$$

and by X the unique càdlàg solution to

$$X_t = Z_t + \int_0^t f(X_{s-})dL_s. \tag{3.8}$$

Then $\operatorname{plim}_{n\to\infty} d_{\mathbb{R}^m}(X^{(n)}, X) = 0$.

(b) If the Lévy process L has finite first moments and $E(L_1) = 0$, then $\bar{L}^{(n)}$ can be replaced by the first jump approximation $L^{(n)}$ given in Theorem 3.1 (b).

Proof. Combine Theorem 3.5 with [11, Corollary 5.6] or [8, Theorem IX.6.9]. □

Remark 3.7. *Our above theorems remain valid when replacing for all $n \in \mathbb{N}$ the filtrations $(\mathscr{F}_t^{(n)})_{t\in\mathbb{R}^+}$ with the completed filtrations generated by the original initial information $\mathscr{F}_0$ and $(L_t^{(n)})_{t\in\mathbb{R}^+}$.*

Moreover, obvious analogues of all our convergence theorems hold when one replaces convergence in the Skorokhod topology in probability by the weaker notion of convergence in the Skorokhod topology in distribution.

Usually $Z^{(n)}$ will simply be constant over time and equal the initial value of the SDE under consideration. Then (3.7) gives for fixed $n \in \mathbb{N}$ an approximation of the solution to (3.8) which is piecewise constant and thus an essentially discrete approximation on the time grid $(t_i^{(n)})_{i\in\mathbb{N}}$. Provided one component of L simply is the time our above approximation equals the Euler scheme for this component. However, in general our approximation scheme is different from the Euler approach. In particular, in our case one needs to be able to draw from the "first jump distributions" of the Lévy process in order to obtain simulations, whereas in the Euler scheme one needs to be able to draw from the distribution of the increments. Thus, the first jump approximation may be considerably easier to simulate than the Euler scheme in some cases (cf. Remark 4.2).

A natural question arising now are convergence rates for the first jump approximation of Lévy driven SDEs. Unfortunately, this seems to be a very intricate issue beyond the scope of the present paper. So far mainly convergence rates for Euler schemes and/or compound Poisson approximations have been obtained in the literature (see [6, 7, 17, 19]). There are two main ways to obtain convergence rates. One is to look at the error of $E(g(X_T))$ for a fixed (terminal) time T and suitable functions g (see [6, 17]). This is very relevant in applications, as often the aim of simulating the solution of an SDE is simply to obtain an approximation of $E(g(X_T))$ for a certain function. The other is of a "functional central limit type" (see [7, 12, 19]) and especially relevant when one is interested in path functionals. Here one is seeking for convergence rates ρ_n such that $\rho_n(X^{(n)} - X)$ has a non-trivial limit in the Skorokhod topology. The latter approach seems to fit well with our results at a first glance. Yet, as a first step one would need to obtain a "functional central limit type" result for the first jump approximation of a Lévy process similar to [19, Theorem 3.1].

The existing results on convergence rates for approximations of SDEs are not applicable to our setting nor seem the methods used to obtain them adaptable to our setting without considerable further work, because the shift in the jump times and the fact that both the time grid size and the minimal jump size have to go to zero at connected rates make the first jump approximation markedly different from the previously studied approximations. Moreover, severe technical conditions, e.g. differentiability conditions for the function f which we demand only to be Lipschitz above, will be needed to obtain convergence rates. However, it is clearly intended to study functional central limit type results and convergence rates of the first jump approximation in future work.

Note also that [22] provides convergence rates for the first jump approximation of a Lévy process directly in terms of the Skorokhod metric by considering explicit expressions for the error. As far as we know, this cannot be translated into convergence rates for our SDE approximations, because no appropriate stability results for SDEs are available.

4 Approximation of ECOGARCH(1,1) processes by EGARCH(1,1) processes

The first jump approximation can be used to obtain very interesting convergence results. In the following we use it to show that a multivariate ECOGARCH(1,1) process is the limit of piece-wise constant processes which are determined by discrete time EGARCH(1,1) processes.

Before we show this, we briefly review multivariate EGARCH(1,1) and ECOGARCH(1,1) processes referring to [4, 5] for more details and a discussion of the relevance in applications.

4.1 Multivariate EGARCH(1,1) processes

Multivariate EGARCH processes have been introduced recently in [9] and the following more general definition can be found in [5].

Let $d \in \mathbb{N}, \mu \in \mathbb{S}_d, \alpha, \beta \in M_m(\mathbb{R})\backslash\{0\}$ with $m = \frac{d(d+1)}{2}$, $\varepsilon = (\varepsilon_n)_{n\in\mathbb{Z}}$ an i.i.d. sequence of $\mathbb{R}^d$-valued random variables with $E(\varepsilon_1) = 0$ and $\mathrm{var}(\varepsilon_1) = I_d$ and $f : \mathbb{R}^d \to \mathbb{R}^m$ a measurable function such that $f(\varepsilon_1) \in L^2$. If $\sigma(\alpha) \subset (-1,1) + i\mathbb{R}$, then the process $Y = (Y_t)_{t\in\mathbb{Z}}$, where

$$Y_t = \exp((\mu + H_t)/2)\varepsilon_t$$

and the vectorised log volatility $X = \mathrm{vech}(H)$ with initial value X_0 is given by

$$X_t = \alpha X_{t-1} + \beta f(\varepsilon_{t-1})$$

for all $t \in \mathbb{N}$, is called an *EGARCH(*$1,1$*) process*.

Above we have considered a general transformation f of the noise sequence ε. Concrete specifications (incorporating the "leverage effect") are to be found in [5, 9].

4.2 Multivariate ECOGARCH(1,1) processes

In the EGARCH(1,1) processes the log volatility process H is an autoregressive process of order one in the symmetric $d \times d$ matrices. Likewise, the ECOGARCH$(1,1)$ process is defined by specifying the log-volatility process as an autoregressive process of order one in continuous time (i.e. an Ornstein-Uhlenbeck type process) taking values in $\mathbb{S}_d$. Moreover, the i.i.d. noise sequence ε is replaced by the increments of a Lévy process.

Let $L = (L_t)_{t\geq 0}$ be a d-dimensional zero-mean Lévy process with Lévy measure ν_L such that $\int_{\|x\|\geq 1} \|x\|^2 \nu_L(dx) < \infty$ and associated jump measure N_L. Furthermore, let $h : \mathbb{R}^d \to \mathbb{R}^m$ with $m = \frac{d(d+1)}{2}$ be a measurable function satisfying

$$\int_{\mathbb{R}^d} \|h(x)\|^2 \nu_L(dx) < \infty, \tag{4.1}$$

and $A \in GL_d(\mathbb{R})$, $\beta \in M_m(\mathbb{R})\backslash\{0\}$ such that all eigenvalues of A have strictly negative real part.

Then we define the d-dimensional ECOGARCH$(1,1)$ process G as the stochastic process satisfying,

$$dG_t := \exp((\mu + H_{t-})/2)dL_t, \quad t > 0, \quad G_0 = 0,$$

where $\mu \in \mathbb{S}_d$ and the vectorised *log-volatility process* $X = (X_t)_{t\geq 0} = (\text{vech}(H)_t)_{t\geq 0}$ is the process in $\mathbb{S}_d$ satisfying

$$dX_t \quad = \quad AX_{t-}dt + \beta\, dM_t\,, \quad t > 0, \tag{4.2}$$

with the initial value $X_0 = \text{vech}(H_0) \in \mathbb{R}_{qm}$ being independent of the driving Lévy process L and

$$M_t := \int_0^t \int_{\mathbb{R}^d\setminus\{0\}} h(x)\tilde{N}_L(ds,dx)\,, \qquad t \geq 0,$$

being a Lévy process. Note that the mean of M is automatically zero due to the use of the compensated jump measure above. The solution of (4.2) is given by

$$X_t = e^{At}X_0 + \int_0^t e^{A(t-s)}\beta\, dM_s. \tag{4.3}$$

In a financial context G is understood to be the log price process of d stocks with volatility (instantaneous covariance matrix) process $\exp(\mu + H_t)$. Moreover, the log returns over a time interval of length $r > 0$ ending at time t, which are especially relevant in a financial context, are described by the increments of G

$$Y_t^{(r)} := G_t - G_{t-r} = \int_{(t-r,t]} \exp((\mu + H_{s-})/2)\, dL_s\,, \qquad t \geq r > 0\,. \tag{4.4}$$

An equidistant sequence of such non-overlapping returns of length r given by $(Y_{ir}^{(r)})_{i\in\mathbb{N}}$ corresponds to a discrete time multivariate EGARCH(1,1) process Y.

4.3 The Approximation

Now we show that for any multivariate ECOGARCH(1,1) process there exists a sequence of piecewise constant processes determined by discrete time multivariate EGARCH(1,1) processes which converges in probability in the Skorokhod topology to the ECOGARCH(1,1) process. This result is also new in the univariate case and should be especially useful for statistical purposes (cp. [15]). Our main theorem of this section is the following where again $m := d(d+1)/2$.

Theorem 4.1. *Let (G,X) in $\mathbb{R}^d \times \mathbb{R}^m$ be a d-dimensional ECOGARCH(1,1) process G and its associated vectorised log-volatility process $X = \text{vech}(H)$ with initial value (G_0,X_0). Let $(t_i^{(n)})_{i\in\mathbb{N}_0}$ for each $n \in \mathbb{N}$ be a strictly increasing sequence in $\mathbb{R}^+$ with $t_0^{(n)} = 0$ and $\lim_{i\to\infty} t_i^{(n)} = \infty$. Defining $\delta^{(n)} = \sup_{i\in\mathbb{N}}\left\{t_i^{(n)} - t_{i-1}^{(n)}\right\}$ assume that $\lim_{n\to\infty}\delta^{(n)} = 0$.*

Then there exists for each $n \in \mathbb{N}$ a function $h_n : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^m$ and a sequence of independent random variables $(\varepsilon_i^{(n)})_{i\in\mathbb{N}}$ in $\mathbb{R}^d$ with finite variance and $E(\varepsilon_i^{(n)}) = 0\,\forall\, i,n \in \mathbb{N}$ such that $h_n\left(\varepsilon_i^{(n)}, t_i^{(n)} - t_{i-1}^{(n)}\right)$ has finite variance, $E\left(h_n\left(\varepsilon_i^{(n)}, t_i^{(n)} - t_{i-1}^{(n)}\right)\right) = 0$ and

$$\text{plim}_{n\to\infty} d_{\mathbb{R}^d\times\mathbb{R}^m}\left((G^{(n)},X^{(n)}),(G,X)\right) = 0\,,$$

where for each $n \in \mathbb{N}$ the process $(G^{(n)},X^{(n)})$ in $\mathbb{R}^d \times \mathbb{R}^m$ is defined by

$$\begin{aligned}
&(G_0^{(n)},X_0^{(n)}) = (G_0,X_0),\\
&\qquad G_{t_i^{(n)}}^{(n)} = G_{t_{i-1}^{(n)}}^{(n)} + \exp\left(\left(\mu + \text{vech}^{-1}\left(X_{t_{i-1}^{(n)}}^{(n)}\right)\right)/2\right)\varepsilon_i^{(n)},\\
&\qquad X_{t_i^{(n)}}^{(n)} = e^{A\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}X_{t_{i-1}^{(n)}}^{(n)} + \beta\, h_n\left(\varepsilon_i^{(n)}, t_i^{(n)} - t_{i-1}^{(n)}\right) \text{ for all } i \in \mathbb{N} \text{ and}
\end{aligned} \tag{4.5}$$

$$(G_t^{(n)}, X_t^{(n)}) = \left(G_{t_{i-1}^{(n)}}^{(n)}, X_{t_{i-1}^{(n)}}^{(n)}\right) \textit{ for } t \in (t_{i-1}^{(n)}, t_i^{(n)}),\ i \in \mathbb{N}.$$

The sequence $(\varepsilon_i^{(n)})_{i\in\mathbb{N}}$ *can be chosen to be i.i.d. provided* $t_i^{(n)} - t_{i-1}^{(n)} = \delta^{(n)}$ *for all* $i \in \mathbb{N}$.
If h *is continuous* h_n *can be chosen such that the sequence of functions* $h_n : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^m$ *satisfies*

$$\lim_{n\to\infty} \left(\sup_{z\in K} \sup_{i\in\mathbb{N}} \left\{ \left\| h_n\left(z, t_i^{(n)} - t_{i-1}^{(n)}\right) - h(z) \right\| \right\} \right) = 0 \tag{4.6}$$

for all compact $K \subset \mathbb{R}^d$. *If* h *is uniformly continuous,* h_n *can be chosen such that* (4.6) *holds with* $\mathbb{R}^d$ *instead of* K.

When the time grids are equidistant, i.e. $t_i^{(n)} - t_{i-1}^{(n)} = \delta^{(n)}$ for all $i \in \mathbb{N}$, and $(\varepsilon_i^{(n)})_{i\in\mathbb{N}}$ is chosen i.i.d., then the increments $\left(G_{t_i^{(n)}}^{(n)} - G_{t_{i-1}^{(n)}}^{(n)}\right)_{i\in\mathbb{N}}$ of $G^{(n)}$ are a discrete time multivariate EGARCH(1,1) process with associated vectorised log-volatility process $\left(X_{t_{i-1}^{(n)}}^{(n)}\right)_{i\in\mathbb{N}} =: \left(\mathrm{vech}(H_i^{(n)})\right)_{i\in\mathbb{N}}$ (apart from the fact that $\mathrm{var}(\varepsilon_i^{(n)}) = I_d$ will usually not be satisfied). We allow for a non-equidistant grid as this may be useful when having irregularly spaced data (cf. [15]). It should be noted that there is an immediate extension to ECOGARCH(p,q) processes with orders $p,q \in \mathbb{N}, p \le q$, (see [4, 5]) where $q > 1$. However, the approximating piecewise constant processes are no longer essentially discrete time EGARCH processes. Furthermore, it is clear that the result remains valid when considering the processes only on a finite time interval $[0,T]$ and looking at partitions $(t_i^{(n)})_{i\in\{1,2,\dots,N^{(n)}\}}$ of this interval with $N^{(n)} \in \mathbb{N}$.

Proof of Theorem 4.1: Let $\|\cdot\|$ be a norm on $\mathbb{R}^{d+m}$. We have that the joint process $\mathbf{L} = (L_t^*, M_t^*)^*_{t\in\mathbb{R}^+}$ is a Lévy process in $\mathbb{R}^{d+m}$ with

$$\mathbf{L}_t = \begin{pmatrix} \gamma_L \\ \gamma_M \end{pmatrix} t + \int_0^t \int_{\|(x^*,h(x)^*)^*\|\le 1} \begin{pmatrix} x \\ h(x) \end{pmatrix} \tilde{N}_L(ds,dx) + \int_0^t \int_{\|(x^*,h(x)^*)^*\|>1} \begin{pmatrix} x \\ h(x) \end{pmatrix} N_L(ds,dx)$$

for all $t \in \mathbb{R}^+$ with $\gamma_L = -\int_{\|(x^*,h(x)^*)^*\|>1} x\nu_L(dx)$ and $\gamma_M = -\int_{\|(x^*,h(x)^*)^*\|>1} h(x)\nu_L(dx)$. Here we used that $\nu_{\mathbf{L}}(W) = \nu_L(f^{-1}(W))$ and $N_{\mathbf{L}}(ds,W) = N_L(ds, f^{-1}(W))$ for all Borel sets $W \subset \mathbb{R}^{d+m}$ where $f : \mathbb{R}^d \to \mathbb{R}^{d+m}$, $x \mapsto (x^*, h(x)^*)^*$. It is clear that $\mathbf{L}$ has finite expectation and $E(\mathbf{L}_1) = 0$.

First Step: Choice of noise sequences $\varepsilon^{(n)}$ *and functions* h_n

Choose a sequence $(m^{(n)})_{n\in\mathbb{N}}$ such that (3.1) is satisfied for $\nu_{\mathbf{L}}$ noting that the existence is ensured by Lemma 3.2. Let $(\mathbf{L}^{(n)})_{n\in\mathbb{N}} = \left((L^{(n)})^*, (M^{(n)})^*\right)^*_{n\in\mathbb{N}}$ be the first jump approximation to $\mathbf{L}$ as given in Theorem 3.1 (b). Hence, $\mathrm{plim}_{n\to\infty} d_{\mathbb{R}^{d+m}}(\mathbf{L}^{(n)}, \mathbf{L}) = 0$ and due to Theorem 3.5 $(\mathbf{L}^{(n)})_{n\in\mathbb{N}}$ has UCV.

Set

$$\varepsilon_i^{(n)} = \Delta L_{t_i^{(n)}}^{(n)} = 1_{(0,\infty)}(\tau_i^{(n)}) \Delta L_{\tau_i^{(n)}} + \gamma_{L,i}^{(n)} \text{ for all } i,n \in \mathbb{N}$$

with

$$\gamma_{L,i}^{(n)} = -\frac{1 - e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)} - t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)} \int_{J^{(n)}} x\nu_L(dx),$$

where $J^{(n)} = \left\{x \in \mathbb{R}^d : \|(x^*, h(x)^*)^*\| > m^{(n)}\right\}$. Then by construction $(\varepsilon_i^{(n)})_{i\in\mathbb{N}}$ is for each $n \in \mathbb{N}$ a sequence of independent random variables having finite variance and zero expectation (cf. Theorem 3.1 (b), (c)). If $t_i^{(n)} - t_{i-1}^{(n)} = \delta^{(n)}$ for all $i \in \mathbb{N}$ then $\gamma_{L,i}^{(n)}$ does not depend on $i \in \mathbb{N}$ and $(\varepsilon_i^{(n)})_{i\in\mathbb{N}}$ is i.i.d.

Moreover,

$$\Delta M^{(n)}_{t_i^{(n)}} = h\left(1_{(0,\infty)}(\tau_i^{(n)})\Delta L_{\tau_i^{(n)}}\right) + \gamma^{(n)}_{M,i} = h_n\left(\varepsilon_i^{(n)}, t_i^{(n)} - t_{i-1}^{(n)}\right)$$

for all $i,n \in \mathbb{N}$, with $\gamma^{(n)}_{M,i} = -\frac{1-e^{-\nu_L\left(J^{(n)}\right)\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}}{\nu_L\left(J^{(n)}\right)}\int_{J^{(n)}} h(x)\nu_L(dx)$ and $h_n : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^m$,

$$(z,t) \mapsto h\left(z + \frac{1-e^{-\nu_L\left(J^{(n)}\right)t}}{\nu_L\left(J^{(n)}\right)}\int_{J^{(n)}} x\nu_L(dx)\right) - \frac{1-e^{-\nu_L\left(J^{(n)}\right)t}}{\nu_L\left(J^{(n)}\right)}\int_{J^{(n)}} h(x)\nu_L(dx).$$

Theorem 3.1 (b), (c) ensures that $h_n\left(\varepsilon_i^{(n)}, t_i^{(n)} - t_{i-1}^{(n)}\right)$ has a finite variance and zero expectation for all $i,n \in \mathbb{N}$.

We have that $\lim_{n\to\infty} \sup_{i\in\mathbb{N}}\left\{\|\gamma^{(n)}_{M,i}\|\right\} = 0$ and $\lim_{n\to\infty} \sup_{i\in\mathbb{N}}\left\{\|\gamma^{(n)}_{L,i}\|\right\} = 0$. From this it is easy to see that (4.6) holds with $\mathbb{R}^d$ instead of K if h is uniformly continuous. If h is only continuous (4.6) follows along the same lines noting that any continuous function is uniformly continuous on compacts.

Second Step: Convergence to the ECOGARCH

Define the processes S in $\mathbb{R}^m$ by $S_t = (1,1,\ldots,1)^* t$ for all $t \in \mathbb{R}^+$ and $(\tilde{X}_t^{(n)})_{t\in\mathbb{R}^+}$ for all $n \in \mathbb{N}$ by

$$\begin{aligned}
\tilde{X}_0^{(n)} &= X_0,\\
\tilde{X}^{(n)}_{t_i^{(n)}} &= e^{A\left(t_i^{(n)}-t_{i-1}^{(n)}\right)}\tilde{X}^{(n)}_{t_{i-1}^{(n)}} + \beta h_n\left(\varepsilon_i^{(n)}, t_i^{(n)} - t_{i-1}^{(n)}\right) \qquad \text{for all } i \in \mathbb{N} \text{ and}\\
\tilde{X}_t &= e^{A\left(t-t_{i-1}^{(n)}\right)}\tilde{X}^{(n)}_{t_{i-1}^{(n)}} \text{ for } t \in (t_{i-1}^{(n)}, t_i^{(n)}),\ i \in \mathbb{N}.
\end{aligned}$$

Below $0_{\mathbb{R}^d}$ denotes the zero in $\mathbb{R}^d$.

Then the joint process $(\tilde{X}^{(n)}, L^{(n)}, M^{(n)}, S)$ satisfies the stochastic integral equation

$$\begin{pmatrix}\tilde{X}_t^{(n)}\\ L_t^{(n)}\\ M_t^{(n)}\\ S_t\end{pmatrix} = \begin{pmatrix}X_0\\ 0_{\mathbb{R}^d}\\ 0_{\mathbb{R}^m}\\ 0_{\mathbb{R}^m}\end{pmatrix} + \int_0^t F\left(\tilde{X}^{(n)}_{s-}\right) d\begin{pmatrix}L_s^{(n)}\\ M_s^{(n)}\\ S_s\end{pmatrix}. \tag{4.7}$$

with

$$F : \mathbb{R}^m \to M_{d+3m,d+2m}(\mathbb{R}), x \mapsto \begin{pmatrix}0_{M_{m,d}(\mathbb{R})} & \beta & Ax\\ I_d & 0_{M_{d,m}(\mathbb{R})} & 0_{M_{d,m}(\mathbb{R})}\\ 0_{M_{m,d}(\mathbb{R})} & I_m & 0_{M_m(\mathbb{R})}\\ 0_{M_{m,d}(\mathbb{R})} & 0_{M_m(\mathbb{R})} & I_m\end{pmatrix}.$$

Obviously F is globally Lipschitz and hence the stochastic integral equation (4.7) has a unique global strong solution. Here we are implicitly using the filtration $(\mathscr{F}_t^{(n)})_{t\in\mathbb{R}}$ for each $n \in \mathbb{N}$ as defined in Theorem 3.5.

Likewise (X,L,M,S) is the unique solution to the stochastic integral equation

$$\begin{pmatrix}X_t\\ L_t\\ M_t\\ S_t\end{pmatrix} = \begin{pmatrix}X_0\\ 0_{\mathbb{R}^d}\\ 0_{\mathbb{R}^m}\\ 0_{\mathbb{R}^m}\end{pmatrix} + \int_0^t F\left(X_{s-}\right) d\begin{pmatrix}L_s\\ M_s\\ S_s\end{pmatrix}.$$

We have that

$$\operatorname{plim}_{n\to\infty} d_{\mathbb{R}^{d+2m}}\left(\left((L^{(n)})^*,(M^{(n)})^*,S^*\right)^*,(L^*,M^*,S^*)^*\right)=0$$

follows from $\operatorname{plim}_{n\to\infty} d_{\mathbb{R}^{d+m}}\left(\mathbf{L}^{(n)},\mathbf{L}\right)=0$. Moreover, the fact that $\mathbf{L}^{(n)}$ has UCV implies that the joint process $\left((L^{(n)})^*,(M^{(n)})^*,S^*\right)^*$ has UCV. Hence, [11, Corollary 5.6] shows that

$$\operatorname{plim}_{n\to\infty} d_{\mathbb{R}^{d+3m}}\left(\left((\tilde{X}^{(n)})^*,(L^{(n)})^*,(M^{(n)})^*,S^*\right)^*,(X^*,L^*,M^*,S^*)^*\right)=0\,,$$

noting that [13, Example 8.2] ensures that F satisfies the necessary technical conditions (alternatively one can use [8, Theorem IX.6.9]). Setting $F(\tilde{X}^{(n)})=\left(F(\tilde{X}_t^{(n)})\right)_{t\in\mathbb{R}^+}$ a continuity argument gives that

$$\operatorname{plim}_{n\to\infty} d_{\mathbb{M}}\left(\left[F(\tilde{X}_t^{(n)}),\left((L^{(n)})^*,(M^{(n)})^*,S^*\right)^*\right],\left[F(X),(L^*,M^*,S^*)^*\right]\right)=0\,,$$

where $\mathbb{M}:=M_{d+3m,d+2m}(\mathbb{R})\times\mathbb{R}^{d+2m}$ and therefore a combination of Theorems 7.7, 7.10, 7.11 of [13] implies that $\left((\tilde{X}^{(n)})^*,(L^{(n)})^*,(M^{(n)})^*,S^*\right)^*_{n\in\mathbb{N}}$ has UCV. Next we observe that for all $T\in\mathbb{R}^+$

$$\sup_{t\le T}\left\|X_t^{(n)}-\tilde{X}_t^{(n)}\right\|\le\sup_{s\in[0,\delta^{(n)}]}\left\|e^{-As}-I_m\right\|_*\sup_{t\le T}\left\|\tilde{X}_t^{(n)}\right\|\,,\tag{4.8}$$

where $\|\cdot\|_*$ is the operator norm induced by $\|\cdot\|$. Since $\left((\tilde{X}^{(n)})^*,(L^{(n)})^*,\ (M^{(n)})^*,S^*\right)^*_{n\in\mathbb{N}}$ has UCV and thus $(\tilde{X}^{(n)})_{n\in\mathbb{N}}$ is uniformly tight (use [13, Theorem 7.6] and [8, Property 6.3, p. 377]), we have from the definition of uniform tightness (cf. [13, Definition 7.4]) that $\sup_{t\le T}\left\|\tilde{X}_t^{(n)}\right\|$ is stochastically bounded in $n\in\mathbb{N}$ for all $T\in\mathbb{R}^+$. Combining this with

$$\lim_{n\to\infty}\sup_{s\in[0,\delta^{(n)}]}\left\|e^{-As}-I_m\right\|_*=0$$

and (4.8) establishes that

$$X^{(n)}-\tilde{X}^{(n)}\to 0 \text{ in ucp as } n\to\infty.\tag{4.9}$$

This implies $\operatorname{plim}_{n\to\infty} d_{\mathbb{R}^{d+2m}}\left(\left((X^{(n)})^*,(\tilde{X}^{(n)})^*,(L^{(n)})^*\right)^*,(X^*,X^*,L^*)^*\right)=0$ and by a continuity argument $\operatorname{plim}_{n\to\infty} d_{\mathbb{S}_d\times\mathbb{R}^{d+m}}\left(\left[Z^{(n)},\left((\tilde{X}^{(n)})^*,(L^{(n)})^*\right)^*\right],\left[Z,(X^*,L^*)^*\right]\right)=0$ with Z and $(Z^{(n)})_{n\in\mathbb{N}}$ defined by $Z_t^{(n)}=\exp\left(\left(\mu+\operatorname{vech}^{-1}\left(X_t^{(n)}\right)\right)/2\right),\quad Z_t=\exp\left(\left(\mu+\operatorname{vech}^{-1}(X_t)\right)/2\right)$ for all $t\in\mathbb{R}^+$.

Finally, we observe that

$$G_t^{(n)}=G_0+\int_0^t Z_{s-}^{(n)}dL_s^{(n)},\qquad \tilde{X}_t^{(n)}=X_0+\int_0^t I_m d\tilde{X}_s^{(n)} \text{ and}$$
$$G_t=G_0+\int_0^t Z_{s-}dL_s,\qquad X_t=X_0+\int_0^t I_m dX_s \text{ for all } t\in\mathbb{R}^+.$$

Therefore [11, Theorem 2.2] (or alternatively [8, Theorem VI.6.22]) shows that

$$\operatorname{plim}_{n\to\infty} d_{\mathbb{R}^d\times\mathbb{R}^m}\left((G^{(n)},\tilde{X}^{(n)}),(G,X)\right)=0,$$

since $((\tilde{X}^{(n)})^*,(L^{(n)})^*)^*$ has UCV. Hence, (4.9) establishes

$$\operatorname{plim}_{n\to\infty} d_{\mathbb{R}^d\times\mathbb{R}^m}\left((G^{(n)},X^{(n)}),(G,X)\right)=0.$$

$\square$

Remark 4.2. *a) We have not discretised time together with (L,M) at the beginning, because we are thus able to recover the exponential decay of* (4.3) *in* (4.5). *If we immediately discretise t as well, we would get $A(t_i - t_{i-1})$ instead of $\exp(A(t_i - t_{i-1}))$ in* (4.5).

b) When using the standard Euler scheme for simulations of ECOGARCH processes, one has to simulate $(L_{t_i} - L_{t_{i-1}}, M_{t_i} - M_{t_{i-1}})$ jointly, as one cannot simply calculate $M_{t_i} - M_{t_{i-1}}$ from $L_{t_i} - L_{t_{i-1}}$. When using the first jump approximation as in Theorem 4.1, one just needs to simulate the "first jumps" for L (i.e. one basically has to simulate a compound Poisson process) and can then calculate the "first jumps" of M by a deterministic transformation. Hence, it will often be easier to simulate the first jump approximation.

c) Using Remark 3.3 and Lemma 3.2 we can, for instance, also use $J^{(n)} = \{\|x\| > m^{(n)}\}$ with $m^{(n)} = (\delta^{(n)})^{1/5}$, if $C\|x\| \geq \|h(x)\| \geq c\|x\|$ for some $C, c > 0$ and all x. The latter condition is e.g. obviously satisfied for the univariate "standard choice" $h(x) = \theta x + \gamma|x|$ with $\theta, \gamma \in \mathbb{R}, \theta + \gamma \neq 0$ and $\theta - \gamma \neq 0$. This is particularly helpful in simulations, because we then just need to simulate the first jump approximation of L and can obtain the one of (L,M) by a simple deterministic transformation.

In the following we consider one of the special choices for h introduced in [5], which is able to model the leverage effect, and thereafter give some variants of our main Theorem 4.1.

Proposition 4.3. *Assume that h is given by $h(\eta) = \Theta\eta + \Gamma\mathrm{vech}\left((\eta\eta^*)^{1/2}\right)$ with $\Theta \in M_{m,d}(\mathbb{R})$ and $\Gamma \in M_m(\mathbb{R})$. Then h is Lipschitz and thereby uniformly continuous.*

Proof. To show that h as given is Lipschitz it suffices to show that $\mathbb{R}^d \to \mathbb{S}_d$, $\eta \mapsto (\eta\eta^*)^{1/2} = \eta\eta^*/\|\eta\|_2$ is Lipschitz. But this is immediate from

$$\left|\frac{\eta_i\eta_j}{\|\eta\|_2} - \frac{\tilde{\eta}_i\tilde{\eta}_j}{\|\tilde{\eta}\|_2}\right| \leq |\eta_i - \tilde{\eta}_i| + |\eta_j - \tilde{\eta}_j| + \|\eta - \tilde{\eta}\|_2$$

for all $\eta, \tilde{\eta} \in \mathbb{R}^d$, $i, j \in \{1, 2, \ldots, d\}$. □

Theorem 4.4. *Let the set-up of Theorem 4.1 be given and assume that $\mathrm{var}(L_1) = I_d$. Then in Theorem 4.1 the function h_n and the sequence of independent random variables $(\varepsilon_i^{(n)})_{i\in\mathbb{N}}$ can for all $n \in \mathbb{N}$ be chosen such that $\varepsilon_i^{(n)} = \left(\sqrt{t_i^{(n)} - t_{i-1}^{(n)}}\right)\zeta_i^{(n)}$ for all $i, n \in \mathbb{N}$ where $(\zeta_i^{(n)})_{i\in\mathbb{N}}$ is for all $n \in \mathbb{N}$ a sequence of independent d-dimensional random variables with $\mathrm{var}(\zeta_i^{(n)}) = I_d$ for all $i, n \in \mathbb{N}$.*

Moreover, if h is continuous h_n can be chosen such that the sequence of functions $h_n : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^m$ satisfies

$$\lim_{n\to\infty}\left(\sup_{z\in K}\sup_{i\in\mathbb{N}}\left\{\left\|h_n\left(z, t_i^{(n)} - t_{i-1}^{(n)}\right) - h(z)\right\|\right\}\right) = 0 \tag{4.10}$$

for all compact $K \subset \mathbb{R}^d$.

Proof. Let $\mathbf{L}, L^{(n)}, M^{(n)}, \gamma_{L,i}^{(n)}, \gamma_{M,i}^{(n)}$ be as in the proof of Theorem 4.1. We have that

$$\begin{aligned} V_i^{(n)} := \mathrm{var}\left(1_{(0,\infty)}(\tau_i^{(n)})\Delta L_{\tau_i^{(n)}}\right) &= \frac{1 - e^{-\nu_L(J^{(n)})(t_i^{(n)} - t_{i-1}^{(n)})}}{\nu_L(J^{(n)})}\int_{J^{(n)}} xx^*\nu_L(dx) \\ &- \left(\frac{1 - e^{-\nu_L(J^{(n)})(t_i^{(n)} - t_{i-1}^{(n)})}}{\nu_L(J^{(n)})}\right)^2 \int_{J^{(n)}} x\nu_L(dx)\int_{J^{(n)}} x^*\nu_L(dx) \end{aligned}$$

for all $i,n \in \mathbb{N}$. A series expansion shows

$$\lim_{n\to\infty}\sup_{i\in\mathbb{N}}\left\|\left(\frac{1-e^{-\nu_L(J^{(n)})(t_i^{(n)}-t_{i-1}^{(n)})}}{\left(\sqrt{t_i^{(n)}-t_{i-1}^{(n)}}\right)\nu_L(J^{(n)})}\right)^2\int_{J^{(n)}}x\nu_L(dx)\int_{J^{(n)}}x^*\nu_L(dx)\right\|_2=0. \tag{4.11}$$

Combining this with

$$\begin{aligned}&\lim_{n\to\infty}\sup_{i\in\mathbb{N}}\left\|\frac{1-e^{-\nu_L(J^{(n)})(t_i^{(n)}-t_{i-1}^{(n)})}}{\nu_L(J^{(n)})(t_i^{(n)}-t_{i-1}^{(n)})}\int_{J^{(n)}}xx^*\nu_L(dx)-I_d\right\|_2\\&\quad\le\lim_{n\to\infty}\left\|\int_{J^{(n)}}xx^*\nu_L(dx)-I_d\right\|_2+\lim_{n\to\infty}\left(\sum_{k=2}^{\infty}\frac{\left(\nu_L(J^{(n)})\delta^{(n)}\right)^{k-1}}{k!}\int_{J^{(n)}}\|x\|_2^2\nu_L(dx)\right)=0\end{aligned}$$

establishes

$$\lim_{n\to\infty}\sup_{i\in\mathbb{N}}\left\|\frac{V_i^{(n)}}{t_i^{(n)}-t_{i-1}^{(n)}}-I_d\right\|=0. \tag{4.12}$$

Hence, there exists a $N\in\mathbb{N}$ such that $V_i^{(n)}\in GL_d(\mathbb{R})$ for all $i\in\mathbb{N}$ and $n\ge N$. Without loss of generality we assume $N=1$. Then a continuity and compactness argument gives

$$\lim_{n\to\infty}\sup_{i\in\mathbb{N}}\left\|\left(V_i^{(n)}\right)^{-1/2}\sqrt{t_i^{(n)}-t_{i-1}^{(n)}}-I_d\right\|=0. \tag{4.13}$$

Define for each $n\in\mathbb{N}$ the process $(\Sigma_t^{(n)})_{t\in\mathbb{R}^+}$ by

$$\Sigma_t^{(n)}=I_d \text{ for } t\in[0,t_1^{(n)}/2),\quad \Sigma_t^{(n)}=\left(V_i^{(n)}\right)^{-1/2}\sqrt{t_i^{(n)}-t_{i-1}^{(n)}}$$

for $t\in\left[\left(t_i^{(n)}+t_{i-1}^{(n)}\right)/2,\left(t_i^{(n)}+t_{i+1}^{(n)}\right)/2\right)$ and $i\in\mathbb{N}$. Since (4.13) implies that $(\Sigma_t^{(n)})_{t\in\mathbb{R}^+}$ converges uniformly to the identity, it follows that

$$\operatorname{plim}_{n\to\infty}d_{M_{d+m}\times\mathbb{R}^{d+m}}\left(\left(\begin{pmatrix}\Sigma^{(n)}&0\\0&I_m\end{pmatrix},\mathbf{L}^{(n)}\right),(I_{d+m},\mathbf{L})\right)=0.$$

Setting $\tilde{L}_t^{(n)}=\int_0^t\Sigma_{s-}^{(n)}dL_s^{(n)}$ and recalling that $\mathbf{L}^{(n)}$ has UCV, we thus have from [11, Theorem 2.2] or [13, Theorem 7.10] that

$$\operatorname{plim}_{n\to\infty}d_{\mathbb{R}^{d+m}}\left(\left((\tilde{L}^{(n)})^*,(M^{(n)})^*\right),(L^*,M^*)^*\right)=0$$

and from [13, Theorem 7.11] that $\left((\tilde{L}^{(n)})^*,(M^{(n)})^*\right)_{n\in\mathbb{N}}$ has UCV.

Setting

$$\zeta_i^{(n)}=\left(V_i^{(n)}\right)^{-1/2}\left(1_{(0,\infty)}(\tau_i^{(n)})\Delta L_{\tau_i^{(n)}}+\gamma_{L,i}^{(n)}\right),\quad \varepsilon_i^{(n)}=\left(\sqrt{t_i^{(n)}-t_{i-1}^{(n)}}\right)\zeta_i^{(n)},$$

$$V_n(t)=\frac{1-e^{-\nu_L(J^{(n)})t}}{\nu_L(J^{(n)})}\int_{J^{(n)}}xx^*\nu_L(dx)-\left(\frac{1-e^{-\nu_L(J^{(n)})t}}{\nu_L(J^{(n)})}\right)^2\int_{J^{(n)}}x\nu_L(dx)\int_{J^{(n)}}x^*\nu_L(dx),$$

$$h_n : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^m,$$
$$(z,t) \mapsto h\left(\frac{(V_n(t))^{1/2}}{\sqrt{t}}z + \frac{1-e^{-\nu_L\left(J^{(n)}\right)t}}{\nu_L\left(J^{(n)}\right)}\int\limits_{J^{(n)}} x\nu_L(dx)\right) - \frac{1-e^{-\nu_L\left(J^{(n)}\right)t}}{\nu_L\left(J^{(n)}\right)}\int\limits_{J^{(n)}} h(x)\nu_L(dx)$$

for all $i,n \in \mathbb{N}$ it is easy to see that $(\varepsilon_i^{(n)})_{i\in\mathbb{N}}$, $(\zeta_i^{(n)})_{i\in\mathbb{N}}$ and h_n have for all $n \in \mathbb{N}$ the claimed properties. Finally, noting that

$$\tilde{L}_t^{(n)} = \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} \varepsilon_i^{(n)} \text{ and } M_t^{(n)} = \sum_{i\in\mathbb{N}:t_i^{(n)}\le t} h_n\left(\varepsilon_i^{(n)}, t_i^{(n)} - t_{i-1}^{(n)}\right) \text{ for all } i,n \in \mathbb{N},\, t \in \mathbb{R}^+$$

the proof continues now as the proof of Theorem 4.1 with $\tilde{L}^{(n)}$ in the place of $L^{(n)}$. □

Note that in the proof presented above we have given a variant of the first jump approximation of a Lévy process fixing not only the mean to zero as in Theorem 3.1 (b) but also the variance to a multiple of the identity.

Proposition 4.5. *Let the set-up of Theorem 4.1 be given and assume that h is linear. Then $h_n(z,t) = h(z)$ can be chosen for all $n \in \mathbb{N}$.*

Proof. Obviously $h_n = h$ for h_n as defined in the proof of Theorem 4.1. □

Acknowledgements

The author is very grateful to the anonymous referee and the associate editor for helpful remarks and suggestions. Furthermore, he especially thanks Stephan Haug for many inspiring discussions and remarks and is grateful to Claudia Klüppelberg and Jean Jacod for helpful remarks on a draft of this paper. Finally, financial support from the Deutsche Forschungsgemeinschaft through the graduate programme "Angewandte Algorithmische Mathematik" at the Technische Universität München is gratefully acknowledged.

References

[1] D. Applebaum. *Lévy Processes and Stochastic Calculus*, volume 93 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2004.

[2] S. Asmussen and J. Rosinski. Approximations of small jumps of Lévy processes with a view towards stochastic simulation. *J. Appl. Probab.*, 38:482–493, 2001.

[3] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1999.

[4] S. Haug and C. Czado. An exponential continuous time GARCH process. *J. Appl. Probab.*, 44:960–976, 2007.

[5] S. Haug and R. Stelzer. Multivariate ECOGARCH processes. In Preparation, 2008.

[6] J. Jacod, T. G. Kurtz, S. Méléard, and P. Protter. The approximate Euler method for Lévy driven stochastic differential equations. *Ann. Inst. H. Poincaré Probab. Statist.*, 41:523–558, 2005.

[7] J. Jacod and P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26:267–307, 1998.

[8] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, 2nd edition, 2003.

[9] H. Kawakatsu. Matrix exponential GARCH. *J. Econometrics*, 134:95–128, 2006.

[10] C. Klüppelberg, A. Lindner, and R. Maller. A continuous-time GARCH process driven by a Lévy process: Stationarity and second-order behaviour. *J. Appl. Probab.*, 41:601–622, 2004.

[11] T. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19:1035–1070, 1991.

[12] T. Kurtz and P. Protter. Wong-Zakai corrections, random evolutions and simulation schemes for SDE's. In E. Mayer-Wolf, E. Merzbach, and A. Shwartz, editors, *Stochastic Analysis – Liber Amicorum for Moshe Zakai*, pages 331–346. Academic Press, 1991.

[13] T. Kurtz and P. Protter. Weak convergence of stochastic integrals and differential equations. In D. Talay and L. Tularo, editors, *Probabilistic Models for Nonlinear Partial Differential Equations*, volume 1627 of *Lecture Notes in Mathematics*, pages 1–41. Springer, 1996.

[14] M. Loéve. *Probability Theory I*, volume 45 of *Graduate Texts in Mathematics*. Springer, New York, 4th edition, 1977.

[15] R. A. Maller, G. Müller, and A. Szimayer. GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli*, 14:519–542, 2008.

[16] P. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, New York, 2nd edition, 2004.

[17] P. Protter and D. Talay. The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.*, 25:393–423, 1997.

[18] J. Rosiński. Series representations of Lévy processes from the perspective of point processes. In O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick, editors, *Lévy Processes Theory and Applications*, pages 401–415, Boston, 2001. Birkhäuser.

[19] S. Rubenthaler. Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. *Stochastic Process. Appl.*, 103:311–349, 2003.

[20] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1999.

[21] M. Signahl. On error rates in normal approximations and simulation schemes for Lévy processes. *Stoch. Models*, 19:287–298, 2003.

[22] A. Szimayer and R. A. Maller. Finite approximation schemes for Lévy processes, and their application to optimal stopping problems. *Stochastic Process. Appl.*, 117:1422–1447, 2007.