

COGARCH as a continuous-time limit of GARCH(1,1)

Jan Kallsen* Bernhard Vesenmayer†

Abstract

COGARCH is an extension of the GARCH time series concept to continuous time, which has been suggested by Klüppelberg, Lindner, Maller (2004). We show that any COGARCH process can be represented as the limit in law of a sequence of GARCH(1,1) processes. As a by-product we derive the infinitesimal generator of the bivariate Markov process representation of COGARCH. Moreover, we argue heuristically that COGARCH and the classical bivariate diffusion limit of Nelson (1990) are probably the only continuous-time limits of GARCH.

Key words: GARCH, continuous time, limit theorem, Markov process, generator

Mathematics Subject Classification (2000): 60F17, 60J25, 91B70

1 Introduction

ARCH (autoregressive conditionally heteroscedastic) and GARCH (generalised ARCH) time series models are very popular in financial econometrics because they capture some of the distinctive features of asset price and other series. They are inherently discrete-time models which raises the natural question of continuous-time extensions, limits, or analogues. [Nel90] shows that properly rescaled GARCH(1,1) models converge in law to a bivariate diffusion process. Interestingly, some peculiar GARCH features are lost in the limit. Firstly, the single source of innovation for both volatility and return series splits into two series in the limit. Secondly, jumps are no longer present in the bivariate diffusion. This second phenomenon typically occurs if innovations are rescaled, as e.g. in Donsker's invariance principle which shows weak convergence of random walks to Brownian motion. The transition from one to two sources of randomness, however, is peculiar to GARCH-type models. This

*Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Christian-Albrechts-Platz 4, 24098 Kiel, Germany, (e-mail: kallsen@math.uni-kiel.de).

†HVB-Stiftungsinstitut für Finanzmathematik, Zentrum Mathematik, TU München, Boltzmannstraße 3, 85747 Garching bei München, Germany, (e-mail: vesenmay@ma.tum.de).

dissimilarity of GARCH and its diffusion limit is underlined by the fact that these models behave differently from the point of view of statistical equivalence (cf. [Wan02, BWZ03]).

Recently, [KLM04] suggested a continuous-time analogue of GARCH(1,1). The definition of these COGARCH (continuous-time GARCH) processes is inspired by some intuitive limit considerations. In contrast to Nelson's diffusion limit, jumps and a single source of randomness as distinctive features of GARCH appear in COGARCH as well. In this paper we show that COGARCH processes can indeed be obtained as limit in law of a sequence of GARCH(1,1) models. The apparent contradiction to Nelson's result is explained by a different limiting procedure. Whereas Nelson rescales the size of innovations, we apply some sort of random thinning, i.e. we decrease the probability of non-trivial innovations.

The existence of two entirely different continuous-time limits naturally leads to the question — which was in fact raised by a referee to an earlier version — whether further processes can be obtained as limit in law of a properly constructed sequence of GARCH models. It may be quite hard to give a complete answer or to even make the idea of a “proper construction” precise. Nevertheless, we argue — on an admittedly very informal level — that Nelson's and the COGARCH models are probably the only continuous-time limits of a “reasonable” sequence of GARCH time series. The argument is based on the limit theory of [JS03]. Roughly speaking, this theory states that convergence of semimartingale characteristics means convergence in law of the corresponding processes. We use the same machinery later to rigorously prove convergence in law of GARCH to COGARCH and bivariate diffusion, respectively.

[KLM04] observed that COGARCH allows for a bivariate Markov process representation. Since semimartingale characteristics are naturally linked to infinitesimal generators, we derive the generator of this Markov process as a by-product.

Very recently, results related to the present paper have been derived independently. [Pan05] calculates the generator of COGARCH on an informal level without giving exact proofs. The convergence of discrete-time GARCH to COGARCH is also derived in [MMS07]. Rather than applying general theory, the construction of [MMS07] is specifically tailored to COGARCH. This approach does not indicate whether other limits could be obtained as well. On the other hand, it leads to convergence in probability as opposed to weak convergence in the present paper. Further references on continuous-time limits of GARCH and in general include [AL05, Zhe06, Kal75, KP91].

Altogether, the aim of this paper is threefold. Firstly, we show that COGARCH can be obtained as continuous-time limit of a sequence of GARCH models. Secondly, we argue heuristically that COGARCH and a slight extension of Nelson's bivariate diffusion are probably the only possible limits. And finally, this paper illustrates the use of the limit theory of [JS03] for deriving possible limit models and for rigorously proving convergence in law.

But let us also stress what we do not attempt to do here. As noted above, [Wan02] shows that Nelson's limit is not statistically equivalent to GARCH(1,1). Whether or not the situation is different for COGARCH and the question of statistical inference as a whole is left to future research (but cf. [MMS07] in this respect). Moreover, we do not discuss

whether COGARCH is a reasonable or even recommendable model e.g. for financial data. But together with the structural similarity of GARCH and COGARCH, our limit theorem suggests that COGARCH may indeed deserve to be considered *the* continuous-time analogue of GARCH. On the other hand, this continuous-time limit no longer allows for the simple statistical inference that contributed decisively to the ubiquitous use of GARCH in the first place.

The paper is structured as follows. In Section 2 we informally derive the conceivable continuous-time limits of a sequence of GARCH models. Subsequently, we give a rigorous proof for convergence in law to COGARCH. Moreover, we derive the generator of COGARCH in Section 3. For sake of completeness we reconsider convergence to Nelson's diffusion limit in Section 4. The appendix contains background material on semimartingale characteristics and their relation to weak convergence and infinitesimal generators.

Unexplained notation is used as in [JS03]. We write $|x|$ for the Euclidean norm of a vector x . The Dirac measure in x is denoted by ε_x .

2 Informal derivation of possible limit processes

To any \mathbb{R}^d -valued semimartingale X there is associated a triplet (B, C, ν) of *characteristics* or (B, \tilde{C}, ν) of *modified characteristics*, where B resp. C, \tilde{C} denote \mathbb{R}^d - resp. $\mathbb{R}^{d \times d}$ -valued predictable processes and ν a random measure on $\mathbb{R}_+ \times \mathbb{R}^d$. The first characteristic B depends on a *truncation function* as e.g. $h(x) = |x|1_{\{|x| \leq 1\}}$, which is chosen a priori. For ease of exposition we choose the identity $h(x) = x$ in this informal section. This is not a proper "truncation" function but this slightly inaccurate choice leads to simpler formulas and avoids burying key ideas by technicalities.

For this choice of h , the first characteristic B corresponds to the predictable trend or *compensator* of X . The matrix-valued process \tilde{C} consists of aggregate covariances of the instantaneous increments of X . Finally, the predictable random measure of jumps ν contains information on the intensity of jumps. The triplet can be easily expressed as

$$B_t = \sum_{s=1}^{[t]} E(\Delta X_s | \mathcal{F}_{s-1}), \quad (2.1)$$

$$\tilde{C}_t = \sum_{s=1}^{[t]} (E(\Delta X_s \Delta X_s^\top | \mathcal{F}_{s-1}) - E(\Delta X_s | \mathcal{F}_{s-1}) E(\Delta X_s | \mathcal{F}_{s-1})^\top), \quad (2.2)$$

$$\nu([0, t] \times A) = \sum_{s=1}^{[t]} E(1_{A \setminus \{0\}}(\Delta X_s) | \mathcal{F}_{s-1}), \quad A \in \mathcal{B}^d \quad (2.3)$$

for discrete-time processes, i.e. if X changes only at integer times. For continuous-time processes, the triplet of characteristics is obtained through a number of rules, which are summarized in the appendix. For more background and precise definitions we refer to [JS03].

The key message of the limit theory in [JS03] is that convergence in law is intimately related with convergence of semimartingale characteristics. Let X^n denote a whole sequence

of processes. We expect convergence in law $X^n \rightarrow X$ in the Skorohod topology if we have

$$B_t^n \rightarrow B_t, \quad \tilde{C}_t^n \rightarrow \tilde{C}_t, \quad \nu^n([0, t] \times \cdot) \rightarrow \nu([0, t] \times \cdot) \quad (2.4)$$

for the corresponding sequences of modified characteristics. Here we remain unspecific about the proper kind of convergence for such result to be true. Rigorous statements require additional technical conditions related to e.g. tightness of the sequence X^n . An instance of a sufficient condition is stated in Theorem A.5 in the appendix. We neglect technical issues and adopt a very informal point of view in this section. By considering the characteristics of discrete-time GARCH models, we wonder what kind of rescaling may naturally lead to a continuous-time limit.

As a side remark, semimartingale characteristics are closely related to the generator of a Markov process, a fact which is discussed in detail in Section A.3 of the appendix. As with characteristics, proper convergence of a sequence of generators implies convergence in law of the corresponding processes. Results along these lines can be found in [EK86].

Recall that a GARCH(1,1) process is defined recursively by

$$Y_k = Z_k \sigma_k, \quad (2.5)$$

$$\sigma_k^2 = \beta + \lambda Y_{k-1}^2 + \delta \sigma_{k-1}^2, \quad (2.6)$$

where $Z_k, k = 1, 2, \dots$ are i.i.d. random variables and $\beta > 0, \lambda > 0, 0 < \delta < 1$ are constants. Moreover, Z_0, σ_0 are supposed to be independent random variables and $Y_0 := Z_0 \sigma_0$. In order to derive convergence results we consider a piecewise constant continuous-time extension of (2.5), (2.6), namely $(\sum_{k=0}^{\lfloor nt \rfloor} Y_k, \sigma_{\lfloor nt \rfloor + 1}^2)_{t \in \mathbb{R}_+}$. We denote by $GARCH_n(\eta, \beta, \lambda, \delta, Q)$ the set of such processes with $Z_0 := 0, \mathcal{L}(\sigma_0^2) = \eta, \mathcal{L}(Z_k) = Q$ for $k \geq 1$. Observe that the mesh size $1/n$ tends to 0 for $n \rightarrow \infty$.

The remainder of this section is dedicated to determining possible weak limits of sequences $GARCH_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n)$ as n tends to ∞ . To this end let

$$(G^n, (\sigma^n)^2) \in GARCH_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n).$$

Observe that superscripts n do not refer to powers or components.

The modified characteristics $(B^n, \tilde{C}^n, \nu^n)$ of $(G^n, (\sigma^n)^2)$ can basically be obtained by using rules (2.1)-(2.3), adjusted for the fact that $(G^n, (\sigma^n)^2)$ changes at multiples of $1/n$ rather than 1 (cf. [JS03, II.3.11 and II.3.18] for details). For $h(x) = x$ we have

$$\begin{aligned} B_t^{n,1} &= \sum_{k=0}^{\lfloor nt \rfloor} \sigma_{k/n}^n \int_{\mathbb{R}} x Q_n(dx), \\ B_t^{n,2} &= \sum_{k=0}^{\lfloor nt \rfloor} \left(\beta_n + (\sigma_{k/n}^n)^2 \left(\delta_n - 1 + \lambda_n \int_{\mathbb{R}} x^2 Q_n(dx) \right) \right), \\ \tilde{C}_t^{n,11} &= \sum_{k=0}^{\lfloor nt \rfloor} (\sigma_{k/n}^n)^2 \left(\int_{\mathbb{R}} x^2 Q_n(dx) - \left(\int_{\mathbb{R}} x Q_n(dx) \right)^2 \right), \\ \tilde{C}_t^{n,12} &= \sum_{k=0}^{\lfloor nt \rfloor} \lambda_n (\sigma_{k/n}^n)^3 \left(\int_{\mathbb{R}} x^3 Q_n(dx) - \int_{\mathbb{R}} x Q_n(dx) \int_{\mathbb{R}} x^2 Q_n(dx) \right), \end{aligned}$$

$$\begin{aligned}
\tilde{C}_t^{n,21} &= \tilde{C}_t^{n,12}, \\
\tilde{C}_t^{n,22} &= \sum_{k=0}^{\lfloor nt \rfloor} \lambda_n^2 (\sigma_{k/n}^n)^4 \left(\int_{\mathbb{R}} x^4 Q_n(dx) - \left(\int_{\mathbb{R}} x^2 Q_n(dx) \right)^2 \right), \\
\nu^n([0, t] \times A) &= \sum_{k=0}^{\lfloor nt \rfloor} \int_{\mathbb{R}} 1_{A \setminus \{0\}} \left(\beta_n + (\sigma_{k/n}^n)^2 (\delta_n - 1 + \lambda_n x^2) \right) Q_n(dx) \quad \forall A \in \mathcal{B}^2.
\end{aligned}$$

The sums can be converted to integrals as e.g.

$$B_t^{n,1} = \int_0^{\lfloor nt \rfloor} \sigma_s^n n \int_{\mathbb{R}} x Q_n(dx) ds.$$

If we approximate $\frac{\lfloor nt \rfloor}{n} \approx t$ and denote by Z^n a random variable with law Q_n , we obtain

$$B_t^{n,1} \approx \int_0^t \sigma_s^n n E(Z^n) ds, \quad (2.7)$$

$$B_t^{n,2} \approx \int_0^t (n\beta_n + (\sigma_s^n)^2 n (\delta_n - 1 + \lambda_n E((Z^n)^2))) ds, \quad (2.8)$$

$$\tilde{C}_t^{n,11} \approx \int_0^t (\sigma_s^n)^2 n \text{Var}(Z^n) ds, \quad (2.9)$$

$$\tilde{C}_t^{n,12} \approx \int_0^t (\sigma_s^n)^3 n \lambda_n \text{Cov}(Z^n, (Z^n)^2) ds, \quad (2.10)$$

$$\tilde{C}_t^{n,22} \approx \int_0^t (\sigma_s^n)^4 n \lambda_n^2 \text{Var}((Z^n)^2) ds, \quad (2.11)$$

$$\begin{aligned}
\nu^n([0, t] \times A) &\approx \int_0^t \int_{\mathbb{R}} 1_{A \setminus \{0\}} \left(\beta_n + (\sigma_s^n)^2 (\delta_n - 1 + \lambda_n x^2) \right) n Q_n(dx) ds \quad \forall A \in \mathcal{B}^2. \\
&\quad (2.12)
\end{aligned}$$

What kind of triplet (B, \tilde{C}, ν) can reasonably occur in the limit? Since $(G^n, (\sigma^n)^2)$ is of Markovian type, we expect a similar structure for the limiting process (G, σ^2) . Specifically, (2.7)-(2.12) suggest limiting characteristics of the form

$$B_t^1 = \int_0^t \sigma_s b_1 ds, \quad (2.13)$$

$$B_t^2 = \int_0^t (b_2 + \sigma_s^2 \tilde{b}_2) ds, \quad (2.14)$$

$$\tilde{C}_t^{11} = \int_0^t \sigma_s^2 c_{11} ds, \quad (2.15)$$

$$\tilde{C}_t^{12} = \int_0^t \sigma_s^3 c_{12} ds, \quad (2.16)$$

$$\tilde{C}_t^{22} = \int_0^t \sigma_s^4 c_{22} ds, \quad (2.17)$$

$$\nu([0, t] \times A) = \int_0^t \int_{\mathbb{R}} 1_{A \setminus \{0\}} \left(\beta + \sigma_s^2 (\delta + \lambda x^2) \right) Q(dx) ds \quad \forall A \in \mathcal{B}^2. \quad (2.18)$$

with parameters $b_1, b_2, \tilde{b}_2, c_{11}, c_{12}, c_{22}, \beta, \delta, \lambda$ and a measure Q on \mathbb{R} . Recall that we look for sequences $\eta_n, \beta_n, \lambda_n, \delta_n, Q_n$ such that convergence (2.4) for the triplets holds. In view of (2.7)-(2.12) and (2.13)-(2.18) this suggests that

$$nE(Z^n) = b_1 + o(1), \quad (2.19)$$

$$n\beta_n = b_2 + o(1), \quad (2.20)$$

$$n(\delta_n - 1 + \lambda_n E((Z^n)^2)) = \tilde{b}_2 + o(1), \quad (2.21)$$

$$n\text{Var}(Z^n) = c_{11} + o(1), \quad (2.22)$$

$$n\lambda_n \text{Cov}(Z^n, (Z^n)^2) = c_{12} + o(1), \quad (2.23)$$

$$n\lambda_n^2 \text{Var}((Z^n)^2) = c_{22} + o(1), \quad (2.24)$$

$$\int_{\mathbb{R}} f(\sigma x) nQ_n(dx) = \int_{\mathbb{R}} f(\sigma x) Q(dx) + o(1), \quad (2.25)$$

$$\int_{\mathbb{R}} f(\beta_n + \sigma^2(\delta_n - 1 + \lambda_n x^2)) nQ_n(dx) = \int_{\mathbb{R}} f(\beta + \sigma^2(\delta + \lambda x^2)) Q(dx) + o(1) \quad (2.26)$$

holds for arbitrary $\sigma^2 > 0$ and sufficiently regular functions $f : \mathbb{R} \rightarrow \mathbb{R}$ that vanish in a neighbourhood of zero. The o - and O -notation refers to $n \rightarrow \infty$.

At this point it is not clear what combinations of variables do really occur in the limit. If the jump measure ν vanishes, we are left with at most six parameters, namely $b_1, b_2, \tilde{b}_2, c_{11}, c_{12}, c_{22}$. Up to c_{12} , all of them occur in Nelson's limit. It is in fact possible to have nonzero c_{12} as well if one allows for a skewed law Q_n . Indeed, one may e.g. choose

$$\begin{aligned} \beta_n &:= \frac{b_2}{n}, \\ \delta_n &:= 1 + \frac{\tilde{b}_2}{n} - \frac{\lambda E(Z^2)}{\sqrt{n}}, \\ \lambda_n &:= \sqrt{n}\lambda, \\ Z^n &\sim \frac{b_1}{n} + \frac{Z}{\sqrt{n}}, \end{aligned}$$

where $\lambda \in \mathbb{R}_+$ and the random variable Z are chosen such that $E(Z) = 0$ and

$$\text{Cov}(Z, \lambda Z^2) = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}.$$

We discuss this limit in Section 4 using a slightly different notation.

The case with jumps is more involved. (2.25) means that the law Q_n of Z^n resembles Q/n away from zero. It may therefore consist of two parts: "large" values of Z^n occur with probability $O(n^{-1})$ and "small" ones with probability close to 1. *Small* here means that their size tends to 0 as $n \rightarrow \infty$. Since $\text{Var}(Z_n) = O(n^{-1})$ by (2.22), the small part contributes only $o(n^{-1})$ to $\text{Var}((Z^n)^2)$. In order to obtain a nontrivial measure Q in the limit, Z_n must have large values with probability exactly of order n^{-1} . Consequently, the contribution of

large jumps to $\text{Var}((Z^n)^2)$ is of order n^{-1} as well. Since the left-hand side of (2.24) must not explode, the sequence (λ_n) should be bounded as $n \rightarrow \infty$.

From (2.26) we expect $\beta_n \rightarrow \beta$. In view of Equation (2.20), this implies $\beta = 0$. Since $E((Z^n)^2) = \text{Var}(Z^n) + (E(Z^n))^2$ is of order n^{-1} , we have $\delta_n = 1 + O(n^{-1})$ by (2.21). Equation (2.26) suggests $\lambda_n \rightarrow \lambda$ and $\delta_n - 1 \rightarrow \delta$, which in turn yields $\delta = 0$.

Recall that “small” values of Z^n contribute only $o(n^{-1})$ to the variance of $(Z^n)^2$. In view of boundedness of the sequence (λ_n) , this implies that small values vanish in the limit c_{22} of (2.24). For any real-valued pure jump process without fixed times of discontinuity, the modified second characteristic \tilde{C} in its triplet (B, \tilde{C}, ν) is entirely determined by the jump measure via

$$\tilde{C}_t = \int_{[0,t] \times \mathbb{R}} x^2 \nu(ds, x).$$

In view of the limiting jump measure (2.18) and $\beta = 0, \delta = 0$, we therefore expect

$$c_{22} = \int \lambda^2 x^4 Q(dx). \quad (2.27)$$

A similar consideration yields

$$c_{12} = \int \lambda x^3 Q(dx). \quad (2.28)$$

This choice $\beta = 0, \delta = 0$, (2.27), (2.28) of parameters in (2.13)-(2.18) corresponds to the COGARCH process, which is discussed in detail in the following section. As in the continuous case above, (2.7)-(2.12) indicate how to obtain these parameters in the limit. One may e.g. choose

$$\begin{aligned} \beta_n &:= \frac{b_2}{n}, \\ \lambda_n &:= \lambda, \\ \delta_n &:= (e^{\tilde{b}_2 - \lambda c_{11}})^{\frac{1}{n}} = 1 + \frac{\tilde{b}_2 - \lambda c_{11}}{n} + o(n^{-1}). \end{aligned}$$

The construction of Q_n is more involved. Away from the origin, we want it to resemble $n^{-1}Q$. Close to the origin, we have a part contributing to the instantaneous variance c_{11} of the limit and another one which takes care of the drift b_1 . We refer to Section 3.3 for a precise construction and for the proof of convergence. The preceding informal considerations are of course far from a rigorous derivation. But they motivate our conjecture that no further limit processes exist.

3 COGARCH

COGARCH processes are defined in terms of a driving Lévy process L . More specifically, we set

$$\begin{aligned} X_t &:= -t \log \delta - \sum_{s \leq t} \log \left(1 + \frac{\lambda}{\delta} (\Delta L_s)^2 \right), \\ \sigma_t^2 &:= \left(\beta \int_0^t e^{X_{s-}} ds + \sigma_0^2 \right) e^{-X_t}, \\ G_t &:= \int_0^t \sigma_{s-} dL_s, \end{aligned}$$

where $\beta > 0, \lambda \geq 0, 0 < \delta < 1$ and σ_0^2 is a positive \mathcal{F}_0 -measurable random variable. In contrast to [KLM04] we choose the right-continuous version of σ^2 in order to stay within the semimartingale setting. Alternatively, one can express the COGARCH volatility process σ^2 in integral form (cf. [KLM04, Proposition 3.2]):

$$\sigma_t^2 = \sigma_0^2 + \int_0^t \left(\beta + \sigma_{s-}^2 \left(\log \delta - \frac{\lambda}{\delta} \tau_L^2 \right) \right) ds + \frac{\lambda}{\delta} \int_0^t \sigma_{s-}^2 d[L, L]_s, \quad (3.1)$$

$$G_t = \int_0^t \sigma_{s-} dL_s. \quad (3.2)$$

In the sequel we write $COGARCH(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$ for the set of all such processes $(G, \sigma^2)_t$. Here, $(\gamma_L, \tau_L^2, \Pi_L)$ denotes the Lévy-Khintchine triplet of L relative to some truncation function h_L on \mathbb{R} and η stands for the law of σ_0^2 .

3.1 Characteristics of COGARCH

As is apparent from Section 2, the approach in this paper relies crucially on semimartingale characteristics. We refer to the appendix for definitions, notation, and properties. We start by determining the semimartingale characteristics of COGARCH.

Theorem 3.1 *Let $(G, \sigma^2) \in COGARCH(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$. The differential characteristics $(b^{(G, \sigma^2)}, c^{(G, \sigma^2)}, F^{(G, \sigma^2)})$ of (G, σ^2) with respect to $h(x_1, x_2) = (h_L(x_1), h_L(x_2))$ are given by*

$$\begin{aligned} b_t^{(G, \sigma^2)} &= \begin{pmatrix} \sigma_{t-} \gamma_L + \int (h_L(\sigma_{t-x}) - \sigma_{t-} h_L(x)) \Pi_L(dx) \\ \beta + \sigma_{t-}^2 \log \delta + \int h_L(\sigma_{t-}^2 \frac{\lambda}{\delta} x^2) \Pi_L(dx) \end{pmatrix}, \\ c_t^{(G, \sigma^2)} &= \begin{pmatrix} \sigma_{t-}^2 \tau_L^2 & 0 \\ 0 & 0 \end{pmatrix}, \\ F_t^{(G, \sigma^2)}(A) &= \int 1_A \begin{pmatrix} \sigma_{t-x} \\ \sigma_{t-}^2 \frac{\lambda}{\delta} x^2 \end{pmatrix} \Pi_L(dx) \quad \forall A \in \mathcal{B}^2 \text{ with } 0 \notin A. \end{aligned}$$

Proof. The characteristics can be calculated following the construction of the COGARCH process. We use the notation $I(t) = t$ for the identity process. The differential characteristics of the process $(L, [L, L], I)$ relative to the truncation function $h_3(x_1, x_2, x_3) = (h_L(x_1), h_L(x_2), h_L(x_3))$ are given by

$$\begin{aligned} b^{(L, [L, L], I)} &= \begin{pmatrix} \gamma_L \\ \tau_L^2 + \int h_L(x^2) \Pi_L(dx) \\ 1 \end{pmatrix}, \\ c^{(L, [L, L], I)} &= \begin{pmatrix} \tau_L^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ F^{(L, [L, L], I)}(A) &= \int 1_A \begin{pmatrix} x \\ x^2 \\ 0 \end{pmatrix} \Pi_L(dx) \quad \forall A \in \mathcal{B}^3. \end{aligned}$$

Applying Proposition A.3 to (3.1), (3.2) yields the differential characteristics of (G, σ^2) as stated in the assertion. \square

Some useful theorems are only stated for processes whose state space is the whole real line. For this technical reason it is more convenient to work with $\log \sigma^2$ instead of the positive process σ^2 . Put differently, we study processes under the following transformation:

$$g(x, y) := (x, \log y). \quad (3.3)$$

The characteristics of $(G, \log \sigma^2) = g(G, \sigma^2)$ are immediately obtained from Theorem 3.1 and Proposition A.4.

Corollary 3.2 *Let $(G, \sigma^2) \in \text{COGARCH}(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$. The differential characteristics (b, c, F) of $g(G, \sigma^2) = (G, \log \sigma^2)$ with respect to $h(x_1, x_2) = (h_L(x_1), h_L(x_2))$ are given by*

$$b_t = \begin{pmatrix} \sigma_{t-} \gamma_L + \int (h_L(\sigma_{t-} x) - \sigma_{t-} h_L(x)) \Pi_L(dx) \\ \log \delta + \frac{\beta}{\sigma_{t-}^2} + \int h_L(\log(1 + \frac{\lambda}{\delta} x^2)) \Pi_L(dx) \end{pmatrix}, \quad (3.4)$$

$$c_t = \begin{pmatrix} \sigma_{t-}^2 \tau_L^2 & 0 \\ 0 & 0 \end{pmatrix}, \quad (3.5)$$

$$F_t(A) = \int 1_A \begin{pmatrix} \sigma_{t-} x \\ \log(1 + \frac{\lambda}{\delta} x^2) \end{pmatrix} \Pi_L(dx) \quad \forall A \in \mathcal{B}^2 \text{ with } 0 \notin A. \quad (3.6)$$

For the application of Theorem A.5 we need that the law P of $g(G, \sigma^2)$ is uniquely determined by its semimartingale characteristics. To be more precise, we must consider the canonical process $X = (X^{(1)}, X^{(2)})$ on the path space $(\mathbb{D}(\mathbb{R}^2), \mathcal{D}(\mathbb{R}^2), \mathbf{D}(\mathbb{R}^2), P)$, i.e. under the law of $g(G, \sigma^2)$. From [Jac79, 12.66] and simple arguments as e.g. in [Kal98, Proposition 2.34] it follows that the characteristics of X have the same form as (3.4)-(3.6) if $\sigma_{t-}^2 = \exp(\log \sigma_{t-}^2)$ is replaced by $\exp(X_{t-}^{(2)})$.

By [JS03, III.2.26] the set of solutions to this martingale problem coincides with the set of weak solutions to some related stochastic differential equation (SDE). Consequently, it suffices to verify pathwise uniqueness for this equation. In our case the SDE is of the following form:

$$\begin{aligned}
d \begin{pmatrix} G_t \\ \log \sigma_t^2 \end{pmatrix} &= \begin{pmatrix} \sigma_{t-\gamma_L} + \int (h_L(\sigma_{t-x}) - \sigma_{t-h_L(x)}) \Pi_L(dx) \\ \log \delta + \frac{\beta}{\sigma_{t-}^2} + \int h_L(\log(1 + \frac{\lambda}{\delta} x^2)) \Pi_L(dx) \end{pmatrix} dt + \begin{pmatrix} \sigma_{t-\tau_L} \\ 0 \end{pmatrix} dW_t \\
&+ h \begin{pmatrix} \sigma_{t-x} \\ \log(1 + \frac{\lambda}{\delta} x^2) \end{pmatrix} (p(dt, dx) - dt \Pi_L(dx)) \\
&+ h' \begin{pmatrix} \sigma_{t-x} \\ \log(1 + \frac{\lambda}{\delta} x^2) \end{pmatrix} p(dt, dx), \tag{3.7}
\end{aligned}$$

where $h'(x) = x - h(x)$, $G_0 = 0$ and σ_0^2 is distributed according to the law η . Moreover, W is a real-valued standard Wiener process and p denotes a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measure $dt \otimes \Pi_L(dx)$, see [JS03, II.1.20]. [Jac79, 14.18] shows that pathwise uniqueness for such an SDE holds under local Lipschitz conditions. This leads to the desired result:

Lemma 3.3 *The law of $g(G, \sigma^2)$ is uniquely determined by its differential characteristics (3.4)-(3.6) and the initial condition $\mathcal{L}(g(G_0, \sigma_0^2)) = g(\varepsilon_0 \otimes \eta)$.*

Proof. We start by showing that uniqueness in law holds for the SDE (3.7). To this end, it suffices to show that the local Lipschitz conditions stated in [Jac79, 14.14] are met. In our case the latter can be written as follows.

For every $n \in \mathbb{N}$ there exist finite increasing processes F^n, G^n , with F^n predictable, and such that the following holds: for any two càdlàg functions $f_1, f_2 : \mathbb{R}_+ \rightarrow \mathbb{R}^2$ with $|f_1(t)| \leq n, |f_2(t)| \leq n$ for all t and $Z_t := \sup_{s \leq t} |f_1(s) - f_2(s)|$ we have

1.

$$\int_0^t \left(e^{\frac{1}{2} f_1^{(2)}(s-)} - e^{\frac{1}{2} f_2^{(2)}(s-)} \right)^2 \tau_L^2 ds \leq \int_0^t Z_{s-}^2 dF_s^n,$$

2.

$$\int_0^t \int_{\mathbb{R}} \left(h_L(e^{\frac{1}{2} f_1^{(2)}(s-)} x) - h_L(e^{\frac{1}{2} f_2^{(2)}(s-)} x) \right)^2 \Pi_L(dx) ds \leq \int_0^t Z_{s-}^2 dF_s^n,$$

3.

$$\begin{aligned}
&\int_0^t \int_{\mathbb{R}} \left| h'_L(e^{\frac{1}{2} f_1^{(2)}(s-)} x) - h'_L(e^{\frac{1}{2} f_2^{(2)}(s-)} x) \right| p(ds, dx) \\
&+ \int_0^t \left(\left| e^{\frac{1}{2} f_1^{(2)}(s-)} - e^{\frac{1}{2} f_2^{(2)}(s-)} \right| \gamma_L + \int_{\mathbb{R}} \left| h_L(e^{\frac{1}{2} f_1^{(2)}(s-)} x) - h_L(e^{\frac{1}{2} f_2^{(2)}(s-)} x) \right. \right. \\
&\quad \left. \left. - (e^{\frac{1}{2} f_1^{(2)}(s-)} - e^{\frac{1}{2} f_2^{(2)}(s-)}) h_L(x) \right| \Pi_L(dx) \right) ds \\
&+ \left| \int_0^t \beta (e^{-f_1^{(2)}(s-)} - e^{-f_2^{(2)}(s-)}) ds \right| \leq \int_0^t Z_{s-} dG_s^n,
\end{aligned}$$

where $h'_L(x) := x - h_L(x)$.

Uniqueness of the solution does not depend on the truncation function. We consider here some h_L, h'_L with Lipschitz constant 1 and $h_L(x) = 0$ for $|x| \geq 2$, $h'_L(x) = 0$ for $|x| \leq 1$. Using the mean value theorem we get

$$\begin{aligned} \left| e^{\frac{1}{2}f_1^{(2)}(s-)} - e^{\frac{1}{2}f_2^{(2)}(s-)} \right| &\leq \frac{1}{2}e^{n/2}Z_{s-}, \\ \left| h_L(e^{\frac{1}{2}f_1^{(2)}(s-)}x) - h_L(e^{\frac{1}{2}f_2^{(2)}(s-)}x) \right| &\leq \frac{1}{2}e^{n/2}|x|1_{[-2e^{n/2}, 2e^{n/2}]}(x)Z_{s-}. \end{aligned}$$

With these properties it is straightforward to verify that the conditions hold for

$$\begin{aligned} F_t^n &:= \left(\tau_L^2 \vee \int_{[-2e^{n/2}, 2e^{n/2}]} |x|^2 \Pi_L(dx) \right) e^{nt}, \\ G_t^n &:= \int_0^t \int_{\mathbb{R}} |x| 1_{[-e^{-n/2}, e^{-n/2}]}(x) p(ds, dx) \\ &\quad + \left(\gamma_L + \int |x| 1_{[e^{-n/2}, 2e^{n/2}]}(|x|) \Pi_L(dx) + e^{n/2} \beta \right) e^{n/2}t. \end{aligned}$$

Applying [Jac79, 14.18] yields pathwise uniqueness, which implies uniqueness in law by [Jac79, 14.94]. By [JS03, III.2.26], the set of weak solutions to (3.7) coincides with the set of solution measures to the martingale problem defined by (b, c, F) . This shows the assertion. \square

With this result we can now show the stronger condition of local uniqueness, see [JS03, III.2.37].

Lemma 3.4 *Local uniqueness holds for the martingale problem corresponding to characteristics (3.4)-(3.6). (Strictly speaking, we refer here to the induced martingale problem on the canonical path space, cf. the discussion following Corollary 3.2.)*

Proof. The differential characteristics in (3.4)-(3.6) do not depend specifically on t . Therefore the "Markovian" type of situation of [JS03, III.2.40] is given. Lemma 3.3 yields that the required uniqueness holds. Thus theorem [JS03, III.2.40] can be applied, yielding the assertion. \square

3.2 Infinitesimal generator of COGARCH

The processes in $COGARCH(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$ are Markovian by [KLM04, Cor.3.1]. The same argument as in [KLM04] yields that they are in fact strong Markov processes. In this section we determine their infinitesimal generator by applying the results of Section A.3.

Theorem 3.5 Let $(G, \sigma^2) \in \text{COGARCH}(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$. On the set $C_c^2(\mathbb{R} \times (0, \infty))$ its infinitesimal generator is defined and satisfies

$$\begin{aligned} \mathcal{K}f(x_1, x_2) &= D_1f(x_1, x_2) \left(\sqrt{x_2}\gamma_L + \int (h_L(\sqrt{x_2}y) - \sqrt{x_2}h_L(y))\Pi_L(dy) \right) \\ &\quad + D_2f(x_1, x_2) (\beta + x_2 \log \delta) + D_{11}f(x_1, x_2)x_2\tau_L^2 \\ &\quad + \int \left(f\left(x_1 + \sqrt{x_2}y, x_2\left(1 + \frac{\lambda}{\delta}y^2\right)\right) - f(x_1, x_2) \right. \\ &\quad \left. - D_1f(x_1, x_2)h_L(\sqrt{x_2}y) \right) \Pi_L(dy). \end{aligned} \quad (3.8)$$

Proof. The right-hand side of (3.8) does not depend on the choice of h_L . We assume h_L to be continuous. Continuity in the sense of Section A.3 holds for the characteristics of the transformed Markov process $g(G, \sigma^2)$, which are computed in Theorem 3.1. Some elementary calculations yield

$$K(x; \{y \in \mathbb{R}^2 : |x + y| \leq n\}) \leq \Pi_L(\{y \in \mathbb{R} : |y| \geq (\sqrt{\delta/\lambda} \wedge e^{-n})\}) < \infty$$

for $|x| \geq 3n$, which implies condition (A.2). Theorem A.7 yields that the generator of $g(G, \sigma^2)$ on the set $C_c^2(\mathbb{R}^2)$ is given by

$$\begin{aligned} \mathcal{K}^g f(x_1, x_2) &= D_1f(x_1, x_2) \left(e^{\frac{1}{2}x_2}\gamma_L + \int (h_L(e^{\frac{1}{2}x_2}y) - e^{\frac{1}{2}x_2}h_L(y))\Pi_L(dy) \right) \\ &\quad + D_2f(x_1, x_2) \left(\log \delta + \beta e^{-x_2} + \int h_L\left(\log\left(1 + \frac{\lambda}{\delta}y^2\right)\right)\Pi_L(dy) \right) \\ &\quad + D_{11}f(x_1, x_2)e^{x_2}\tau_L^2 \\ &\quad + \int \left(f\left(x_1 + e^{\frac{1}{2}x_2}y, x_2 + \log\left(1 + \frac{\lambda}{\delta}y^2\right)\right) - f(x_1, x_2) \right. \\ &\quad \left. - D_1f(x_1, x_2)h_L(e^{\frac{1}{2}x_2}y) - D_2f(x_1, x_2)h_L\left(\log\left(1 + \frac{\lambda}{\delta}y^2\right)\right) \right) \Pi_L(dy). \end{aligned} \quad (3.9)$$

Since $g^{-1} \in C^\infty(\mathbb{R}^2)$, we have $\mathcal{K}f(x_1, x_2) = \mathcal{K}^g(f \circ g^{-1})(g(x_1, x_2))$. Elementary calculus yields (3.8).

In Section A (and hence for Theorem A.7) we assume to work on the canonical path space. However, it is easy to see that the generator of $g(G, \sigma^2)$ on the original space coincides with the generator of the canonical process under the induced law on the path space $\mathbb{D}(\mathbb{R}^2)$. \square

Similarly we can apply Proposition A.8 to show that the generator determines the distribution uniquely.

Lemma 3.6 For fixed (G_0, σ_0^2) the infinitesimal generator (3.8) on $C_c^2(\mathbb{R} \times (0, \infty))$ determines the law of a corresponding strong Markov process uniquely.

Proof. Let $(G, \sigma^2), (\tilde{G}, \tilde{\sigma}^2)$ be $\mathbb{R} \times (0, \infty)$ -valued strong Markov processes with infinitesimal generator (3.8) on $C_c^2(\mathbb{R} \times (0, \infty))$. Then it can be shown similarly as in the previous proof that (3.9) is the generator of $g(G, \sigma^2), g(\tilde{G}, \tilde{\sigma}^2)$ on $C_c^2(\mathbb{R}^2)$. By Lemma 3.3 and Proposition A.8 we get that $g(G, \sigma^2)$ has the same law as $g(\tilde{G}, \tilde{\sigma}^2)$. Since g^{-1} is measurable, this holds for (G, σ^2) and $(\tilde{G}, \tilde{\sigma}^2)$ as well. \square

3.3 Convergence of GARCH(1,1) to COGARCH

In this section we show that any COGARCH process can be obtained as limit in law of a properly chosen sequence of GARCH models. We start with a process

$$(G, \sigma^2) \in \text{COGARCH}(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$$

with parameters η, β , etc. as in the beginning of this section. For ease of notation we assume the truncation function h_L to be Lipschitz and symmetric, and we set $h(x_1, x_2) := (h_L(x_1), h_L(x_2))$. Let

$$\eta_n := \eta, \tag{3.10}$$

$$\begin{aligned} Q_n &:= \frac{1}{n} \Pi_L^{A_n} + \frac{\tau_L^2}{\sqrt{n}} \left(\frac{1}{2} \varepsilon_{n^{-1/4}} + \frac{1}{2} \varepsilon_{-n^{-1/4}} \right) \\ &\quad + \left(1 - \frac{\tau_L^2}{\sqrt{n}} - \frac{1}{n} \Pi_L(A_n) \right) \varepsilon_{\frac{\gamma_n}{n}}, \end{aligned} \tag{3.11}$$

$$\lambda_n := \frac{\lambda}{\delta}, \tag{3.12}$$

$$\delta_n := \left(\delta e^{-\frac{\lambda}{\delta} \tau_L^2} \right)^{\frac{1}{n}}, \tag{3.13}$$

$$\beta_n := \frac{\beta}{n} \tag{3.14}$$

with

$$\begin{aligned} \Pi_L^{A_n}(A) &:= \Pi_L(A \cap A_n) \quad \forall A \in \mathcal{B}, \\ A_n &:= \{y \in \mathbb{R} : |y| \geq m_n\}, \\ \gamma_n &:= \gamma_L - \int_{A_n} h_L(x) \Pi_L(dx), \end{aligned}$$

where m_n is a decreasing sequence with $m_n \rightarrow 0$ and $0 \leq \Pi_L(A_n) \leq n^{-1/4}, n \in \mathbb{N}$. Such a sequence obviously exists. Clearly Q_n is a probability measure, at least for sufficiently large n . The first term in (3.11) generates the jumps of the limiting COGARCH process. The second takes care of the Brownian motion part. It vanishes if $\tau_L^2 = 0$. The third term provides the drift and partially compensates the jumps.

As in Section 2 we consider continuous-time embeddings of discrete GARCH models. The remainder of this section is dedicated to convergence of $\text{GARCH}_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n)$ to $\text{COGARCH}(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$. We start by showing the convergence of the transformed processes (cf. (3.3)). To this end let

$$(G^n, (\sigma^n)^2) \in \text{GARCH}_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n), \tag{3.15}$$

with parameters as in (3.10)-(3.14). As before superscripts n do not refer to powers or components.

The modified characteristics $(B^n, \tilde{C}^n, \nu^n)$ (cf. [JS03, III.3.6]) of $g(G^n, (\sigma^n)^2)$ with respect to the truncation function h can be computed easily, following the approach of [JS03, II.3.11 and II.3.18]:

$$\begin{aligned}
B_t^{n,1} &= \int_0^{\lfloor nt \rfloor} n \int_{\mathbb{R}} h_L(\sigma_s^n x) Q_n(dx) ds, \\
B_t^{n,2} &= \int_0^{\lfloor nt \rfloor} n \int_{\mathbb{R}} h_L \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \lambda_n x^2 \right) \right) Q_n(dx) ds, \\
\tilde{C}_t^{n,11} &= \int_0^{\lfloor nt \rfloor} n \left(\int_{\mathbb{R}} h_L(\sigma_s^n x)^2 Q_n(dx) - \left(\int_{\mathbb{R}} h_L(\sigma_s^n x) Q_n(dx) \right)^2 \right) ds, \\
\tilde{C}_t^{n,12} &= \int_0^{\lfloor nt \rfloor} n \left(\int_{\mathbb{R}} h_L(\sigma_s^n x) h_L \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \lambda_n x^2 \right) \right) Q_n(dx) \right. \\
&\quad \left. - \int_{\mathbb{R}} h_L(\sigma_s^n x) Q_n(dx) \int_{\mathbb{R}} h_L \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \lambda_n x^2 \right) \right) Q_n(dx) \right) ds, \\
\tilde{C}_t^{n,21} &= \tilde{C}_t^{n,12}, \\
\tilde{C}_t^{n,22} &= \int_0^{\lfloor nt \rfloor} n \left(\int_{\mathbb{R}} h_L \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \lambda_n x^2 \right) \right)^2 Q_n(dx) \right. \\
&\quad \left. - \left(\int_{\mathbb{R}} h_L \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \lambda_n x^2 \right) \right) Q_n(dx) \right)^2 \right) ds, \\
\nu^n([0, t] \times A) &= \int_0^{\lfloor nt \rfloor} n \int_{\mathbb{R}} 1_{A \setminus \{0\}} \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \lambda_n x^2 \right) \right) Q_n(dx) ds \quad \forall A \in \mathcal{B}^2.
\end{aligned}$$

For applying Theorem A.5 we must work with the canonical process $X = (X^{(1)}, X^{(2)})$ on the path space $(\mathbb{D}(\mathbb{R}^2), \mathcal{D}(\mathbb{R}^2), \mathbf{D}(\mathbb{R}^2), \mathcal{L}(g(G, \sigma^2)))$. We denote by (B, \tilde{C}, ν) the modified characteristics of X under $\mathcal{L}(g(G, \sigma^2))$. According to Corollary 3.2, they are given by

$$\begin{aligned}
B_t &= \int_0^t \left(\begin{array}{c} e^{\frac{1}{2}X_{s^-}^{(2)}} \gamma_L + \int (h_L(e^{\frac{1}{2}X_{s^-}^{(2)}} x) - e^{\frac{1}{2}X_{s^-}^{(2)}} h_L(x)) \Pi_L(dx) \\ \log \delta + \beta e^{-X_{s^-}^{(2)}} + \int h_L(\log(1 + \frac{\lambda}{\delta} x^2)) \Pi_L(dx) \end{array} \right) ds, \\
\tilde{C}_t &= \int_0^t \left(\begin{array}{cc} e^{X_{s^-}^{(2)}} \tau_L^2 & 0 \\ 0 & 0 \end{array} \right) \\
&\quad + \int_{\mathbb{R}} \left(\begin{array}{cc} (h_L(e^{\frac{1}{2}X_{s^-}^{(2)}} x))^2 & h_L(e^{\frac{1}{2}X_{s^-}^{(2)}} x) h_L(\log(1 + \frac{\lambda}{\delta} x^2)) \\ h_L(e^{\frac{1}{2}X_{s^-}^{(2)}} x) h_L(\log(1 + \frac{\lambda}{\delta} x^2)) & (h_L(\log(1 + \frac{\lambda}{\delta} x^2)))^2 \end{array} \right) \Pi_L(dx) \Big) ds, \\
\nu([0, t] \times A) &= \int_0^t 1_A \left(\begin{array}{c} e^{\frac{1}{2}X_{s^-}^{(2)}} \\ \log(1 + \frac{\lambda}{\delta} x^2) \end{array} \right) \Pi_L(dx) ds.
\end{aligned}$$

Moreover, let

$$\begin{aligned} S_a &:= \inf \{t \in \mathbb{R}_+ : |X_t| \geq a \text{ or } |X_{t-}| \geq a\}, \\ S_a^n &:= \inf \{t \in \mathbb{R}_+ : |g(G_t^n, (\sigma_t^n)^2)| \geq a \text{ or } |g(G_{t-}^n, (\sigma_{t-}^n)^2)| \geq a\}. \end{aligned}$$

As intermediary result we show convergence of these characteristics.

Lemma 3.7 *The modified semimartingale characteristics $(B^n, \tilde{C}^n, \nu^n)$ of $g(G^n, (\sigma^n)^2)$ as given above converge to (B, \tilde{C}, ν) in the following sense:*

1. $\sup_{s \leq t} |B_{s \wedge S_a^n}^n - (B_{s \wedge S_a}) \circ g(G^n, (\sigma^n)^2)| \xrightarrow{P} 0$ for all $t > 0, a > 0$,
2. $\tilde{C}_{t \wedge S_a^n} - (\tilde{C}_{t \wedge S_a}) \circ g(G^n, (\sigma^n)^2) \xrightarrow{P} 0$ for all $t > 0, a > 0$,
3. $f * \nu_{t \wedge S_a^n}^n - (f * \nu_{t \wedge S_a}) \circ g(G^n, (\sigma^n)^2) \xrightarrow{P} 0$ for all $t > 0, a > 0, f \in C(\mathbb{R}^d)$ (cf. Theorem A.5).

Proof. Obviously we have

$$A_n \nearrow \mathbb{R}, \quad (3.16)$$

$$\Pi_L(A_n) = o(\sqrt{n}). \quad (3.17)$$

In the sequel we write $U_t^n \sim V_t^n$ if $\sup_{s \leq t \wedge S_a^n} |U_s^n - V_s^n| \xrightarrow{P} 0$ for $n \rightarrow \infty$.

1. Using (3.16), (3.17), symmetry of h_L , and dominated convergence, we can approximate $B^{n,1}$ as follows:

$$\begin{aligned} B_t^{n,1} &= \int_0^{\lfloor \frac{nt}{n} \rfloor} \left(\int_{A_n} h_L(\sigma_s^n x) \Pi_L(dx) + 0 \right. \\ &\quad \left. + n \left(1 - \frac{\tau_L^2}{\sqrt{n}} - \frac{1}{n} \Pi_L(A_n) \right) h_L \left(\sigma_s^n \frac{\gamma_n}{n} \right) \right) ds \\ &\sim \int_0^{\lfloor \frac{nt}{n} \rfloor} \left(\int_{A_n} h_L(\sigma_s^n x) \Pi_L(dx) + \sigma_s^n \gamma_n \right) ds \\ &= \int_0^{\lfloor \frac{nt}{n} \rfloor} \left(\int_{A_n} (h_L(\sigma_s^n x) - \sigma_s^n h_L(x)) \Pi_L(dx) + \sigma_s^n \gamma_L \right) ds \\ &\sim \int_0^t \left(\int_{\mathbb{R}} (h_L(\sigma_s^n x) - \sigma_s^n h_L(x)) \Pi_L(dx) + \sigma_s^n \gamma_L \right) ds \\ &= B_t^1 \circ g(G^n, (\sigma^n)^2). \end{aligned}$$

By (3.16), (3.17), dominated convergence, and $(\frac{a}{n} + b^{1/n})^n \xrightarrow{n \rightarrow \infty} be^a$, we have

$$\begin{aligned}
B_t^{n,2} &= \int_0^{\lfloor nt \rfloor} \left(\int_{A_n} h_L \left(\log \left(\frac{\beta}{n(\sigma_s^n)^2} + \delta_n + \frac{\lambda}{\delta} x^2 \right) \right) \Pi_L(dx) \right. \\
&\quad + \tau_L^2 \sqrt{n} h_L \left(\log \left(\frac{\beta}{n(\sigma_s^n)^2} + \delta_n + \frac{\lambda}{\delta} \frac{1}{\sqrt{n}} \right) \right) \\
&\quad \left. + (n - \tau_L^2 \sqrt{n} - \Pi_L(A_n)) h_L \left(\log \left(\frac{\beta}{n(\sigma_s^n)^2} + \delta_n + \frac{\lambda (\gamma_n)^2}{\delta n^2} \right) \right) \right) ds \\
&\sim \int_0^{\lfloor nt \rfloor} \left(\int_{A_n} h_L \left(\log \left(1 + \frac{\lambda}{\delta} x^2 \right) \right) \Pi_L(dx) \right. \\
&\quad \left. + \frac{\lambda}{\delta} \tau_L^2 + n h_L \left(\log \left(\frac{\beta}{n(\sigma_s^n)^2} + \delta_n \right) \right) \right) ds \\
&\sim \int_0^{\lfloor nt \rfloor} \left(\int_{\mathbb{R}} h_L \left(\log \left(1 + \frac{\lambda}{\delta} x^2 \right) \right) \Pi_L(dx) + \log \delta + \frac{\beta}{(\sigma_s^n)^2} \right) ds \\
&\sim \int_0^t \left(\int_{\mathbb{R}} h_L \left(\log \left(1 + \frac{\lambda}{\delta} x^2 \right) \right) \Pi_L(dx) + \log \delta + \frac{\beta}{(\sigma_s^n)^2} \right) ds \\
&= B_t^2 \circ g(G^n, (\sigma^n)^2).
\end{aligned}$$

2. From

$$\begin{aligned}
h_L \left(\sigma_s^n \frac{\gamma_n}{n} \right) &= o \left(\frac{1}{\sqrt{n}} \right), \\
h_L \left(\frac{\sigma_s^n}{n^{1/4}} \right) &= o(1), \\
h_L \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \lambda_n \frac{\gamma_n^2}{n^2} \right) \right) &= o \left(\frac{1}{\sqrt{n}} \right), \\
h_L \left(\log \left(\frac{\beta_n}{(\sigma_s^n)^2} + \delta_n + \frac{\lambda_n}{\sqrt{n}} \right) \right) &= o(1)
\end{aligned}$$

for $s \leq S_a^n$, dominated convergence, and (3.16), (3.17) it follows that

$$\begin{aligned}
\tilde{C}_t^{n,11} &\sim \int_0^{\lfloor nt \rfloor} \left(\int_{A_n} h_L(\sigma_s^n x)^2 \Pi_L(dx) + \tau_L^2 \sqrt{n} h_L \left(\frac{\sigma_s^n}{n^{1/4}} \right)^2 \right. \\
&\quad \left. + n \left(1 - \frac{\tau_L^2}{\sqrt{n}} - \frac{1}{n} \Pi_L(A_n) \right) h_L \left(\sigma_s^n \frac{\gamma_n}{n} \right)^2 + o(1) \right) ds \\
&\sim \int_0^t \left(\int_{\mathbb{R}} h_L(\sigma_s^n x)^2 \Pi_L(dx) + (\sigma_s^n)^2 \tau_L^2 \right) ds \\
&= \tilde{C}_t^{11} \circ g(G^n, (\sigma^n)^2)
\end{aligned}$$

and similarly

$$\begin{aligned}\tilde{C}_t^{n,12} &\sim \int_0^t \int_{\mathbb{R}} h_L(\sigma_s^n x) h_L\left(\log\left(1 + \frac{\lambda}{\delta} x^2\right)\right) \Pi_L(dx) ds = \tilde{C}_t^{12} \circ g(G^n, (\sigma^n)^2), \\ \tilde{C}_t^{n,22} &\sim \int_0^t \int_{\mathbb{R}} h_L\left(\log\left(1 + \frac{\lambda}{\delta} x^2\right)\right)^2 \Pi_L(dx) ds = \tilde{C}_t^{22} \circ g(G^n, (\sigma^n)^2).\end{aligned}$$

3. Let $f \in C(\mathbb{R}^d)$. Since σ_-^n and $1/\sigma_-^n$ are bounded on $\llbracket 0, S_a^n \rrbracket$ and f vanishes in a neighbourhood of 0, we can ignore the integrals related to $\varepsilon_{\gamma_n/n}, \varepsilon_{n^{-1/4}}, \varepsilon_{-n^{-1/4}}$ for n large enough. Dominated convergence yields

$$\begin{aligned}f * \nu_t^n &\sim \int_0^{\lfloor nt \rfloor} \int_{A_n} f\left(\log\left(\frac{\beta}{n(\sigma_s^n)^2} + \delta_n + \frac{\lambda}{\delta} x^2\right)\right) \Pi_L(dx) ds \\ &\sim \int_0^t \int_{\mathbb{R}} f\left(\log\left(1 + \frac{\lambda}{\delta} x^2\right)\right) \Pi_L(dx) ds \\ &= (f * \nu_t) \circ g(G^n, (\sigma^n)^2).\end{aligned}$$

This shows the claim. \square

Provided with the previous results we can finally prove convergence in law of $(G^n, (\sigma^n)^2)$ to (G, σ^2) .

Theorem 3.8 *Let $(G, \sigma^2) \in \text{COGARCH}(\eta, \beta, \lambda, \delta, (\gamma_L, \tau_L^2, \Pi_L))$ with parameters η, β , etc. as in the beginning of this section. Then the sequence $(G^n, (\sigma^n)^2) \in \text{GARCH}_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n)$ defined as in (3.10)-(3.15) converges in law to (G, σ^2) .*

Proof. We show convergence in law of $g(G^n, (\sigma^n)^2)$ to $g(G, \sigma^2)$. This implies convergence in law of $(G^n, (\sigma^n)^2)$ to (G, σ^2) because g^{-1} is continuous. We proceed by verifying the conditions of Theorem A.5. Since convergence does not depend on the truncation function, we may choose h_L as in the proof of Lemma 3.3. We have $\text{Var}(\int_0^t f(s) ds)_t = \int_0^t |f(s)| ds$. Therefore it is easy to verify that the following definition provides a majorizing function of $\sum_{i=1}^2 \text{Var}(B^i)^{S_a}$:

$$\begin{aligned}F_t(a) &:= te^{a/2} \left(\gamma_L + 2\Pi_L([-e^{-a/2}, e^{-a/2}]^C) + e^{a/2} \beta \right) \\ &\quad + t \left(\frac{1}{2} |\log \delta| + \int h_L\left(\log\left(1 + \frac{\lambda}{\delta} x^2\right)\right) \Pi_L(dx) \right).\end{aligned}$$

Along the same lines, one takes care of $(\sum_{i=1}^2 C^{ii} + (|x|^2 \wedge 1) * \nu)^{S_a}$. The local condition on big jumps is satisfied because

$$\begin{aligned}&\lim_{m \rightarrow \infty} \sup_{\alpha \in \mathbb{D}(\mathbb{R}^2)} \int_0^{t \wedge S_a} \int_{\mathbb{R}} 1_{\{|y| > m\}} \left(\frac{e^{\frac{1}{2} \alpha^{(2)}(s^-) x}}{\log\left(1 + \frac{\lambda}{\delta} x^2\right)} \right) \Pi_L(dx) ds \\ &\leq \lim_{m \rightarrow \infty} t \Pi_L \left\{ x \in \mathbb{R} : e^{a/2} x^2 + \left(\log\left(1 + \frac{\lambda}{\delta} x^2\right) \right)^2 > m^2 \right\} \\ &= 0.\end{aligned}$$

Local uniqueness for the martingale problem is shown in Lemma 3.4.

In order to show Skorokhod-continuity let $\alpha_n \rightarrow \alpha$ in $\mathbb{D}(\mathbb{R}^2)$, i.e. there exist strictly increasing functions $\lambda_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\lambda_n(0) = 0$ such that $\lim_{n \rightarrow \infty} \sup_{t \geq 0} |\lambda_n(t) - t| = 0$ and $\lim_{n \rightarrow \infty} \sup_{s \leq t} |\alpha_n \circ \lambda_n(s) - \alpha(s)| = 0$ for any $t \geq 0$ (cf. [JS03, VI.1.14]). If we set

$$\Delta_n(s) := e^{\frac{1}{2}\alpha_n^{(2)}(\lambda_n(s-))} - e^{\frac{1}{2}\alpha^{(2)}(s-)},$$

then $\lim_{n \rightarrow \infty} \sup_{s \leq t} |\Delta_n(s)| = 0$. Moreover, we have

$$\begin{aligned} & \sup_{s \leq t} |B_s^1(\alpha_n \circ \lambda_n) - B_s^1(\alpha)| \\ & \leq |\gamma_L| |\Delta_n(s)| t + 2 \sup_{s \leq t} |\Delta_n(s)| \int |x| 1_{[e^{-a/2}, 2e^{a/2}]}(|x|) \Pi_L(dx) t \\ & \rightarrow 0 \end{aligned}$$

for $n \rightarrow \infty$, where a is chosen such that $\sup_{s \leq t} |\alpha_n \circ \lambda_n(s)| \leq a$ and $\sup_{s \leq t} |\alpha(s)| \leq a$ for any n . Consequently, $B^1(\alpha_n \circ \lambda_n) \rightarrow B^1(\alpha)$ in the Skorohod topology, which implies that $B_t^1(\alpha_n) \rightarrow B_t^1(\alpha)$ for all t where α does not jump (cf. [JS03, VI.2.1]). Skorohod-continuity of B^2, \tilde{C}, ν is shown similarly.

From

$$\mathcal{L}(\log((\sigma_0^n)^2)) = \mathcal{L}\left(\log\left(\frac{\beta}{n} + (\delta_n + 0)(\sigma_0)^2\right)\right)$$

with $\mathcal{L}(\sigma_0^2) = \eta$ it follows that $g(G_0^n, (\sigma_0^n)^2) \rightarrow g(G_0, (\sigma_0)^2)$ in law. Convergence of the characteristics is shown in Lemma 3.7. Thus Theorem A.5 can be applied and the assertion follows. \square

We have shown that any COGARCH process is obtained as limit in law of GARCH(1,1) models. But we can also turn things around. Given a concrete GARCH(1,1) model, proper rescaling of the parameters naturally leads to a converging sequence. Indeed, consider a particular set $\eta, \beta, \lambda, \delta, Q$ of GARCH(1,1) parameters such that $\beta > 0, \lambda \geq 0, 0 < \delta < 1$. Let

$$\begin{aligned} \eta_n &:= \eta, \\ Q_n &:= \frac{1}{n}Q + \left(1 - \frac{1}{n}\right)\varepsilon_0, \\ \lambda_n &:= \lambda, \\ \delta_n &:= \delta^{\frac{1}{n}}, \\ \beta_n &:= \frac{\beta}{n}. \end{aligned}$$

Obviously we have $(\eta_1, \beta_1, \lambda_1, \delta_1, Q_1) = (\eta, \beta, \lambda, \delta, Q)$. The probability measure Q_n corresponds to a randomized experiment. In each step a coin is tossed showing *heads* with probability $\frac{1}{n}$. Only if *heads* turn up, the innovation Z_k is drawn according to the law Q , otherwise it is chosen to be 0. As noted in the introduction, this means that we decrease the *probability* of innovations rather than their *size*. Theorem 3.8 shows that $GARCH_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n)$

converges in law to $COGARCH(\eta, \beta, \lambda\delta, \delta, (\int h_L(x)Q(dx), 0, Q))$. From the form of the triplet it can be seen that the driving Lévy process is a compound Poisson process, i.e. piecewise constant between discrete jumps times. The jumps occur with intensity 1 and they are distributed according to Q .

4 Nelson's limit revisited

In this section we reconsider Nelson's classical bivariate diffusion limit of GARCH. To this end denote by (G, σ^2) the solution to

$$dG_t = \alpha\sigma_t dt + \gamma_1\sigma_t dW_t \quad (4.1)$$

$$d\sigma_t^2 = (\beta + \delta\sigma_t^2)dt + \gamma_2\sigma_t^2 d\widetilde{W}_t \quad (4.2)$$

with $\mathcal{L}(G_0, \sigma_0^2) = \eta$. Here, $\alpha, \beta, \delta, \gamma_1, \gamma_2$ are parameters with $\beta = 0$ and W, \widetilde{W} denote standard Wiener processes with constant correlation $\varrho \in [-1, 1]$, i.e. $[W, \widetilde{W}]_t = \varrho t$. The literature on continuous-time limits of GARCH considers mostly Gaussian innovations Z^n in (2.5), (2.6), which leads to $\varrho = 0$ in the limit. By allowing for skewed laws Q_n of Z^n , non-zero correlation can be obtained as well. In order to show convergence to (G, σ^2) above, we consider a sequence $GARCH_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n)$ with parameters of the form

$$\eta_n := \eta, \quad (4.3)$$

$$\beta_n := \frac{\beta}{n}, \quad (4.4)$$

$$Q^n = \mathcal{L}\left(\frac{\alpha}{n} + \frac{Z}{\sqrt{n}}\right), \quad (4.5)$$

$$\lambda_n := \sqrt{n}\lambda, \quad (4.6)$$

$$\delta_n := 1 + \frac{\delta}{n} - \frac{\lambda E(Z^2)}{\sqrt{n}}, \quad (4.7)$$

where $\lambda \in \mathbb{R}_+$ and the random variable Z are chosen such that $E(Z^4) < \infty$, $E(Z) = 0$ and

$$\text{Cov}(Z, \lambda Z^2) = \begin{pmatrix} \gamma_1^2 & \varrho\gamma_1\gamma_2 \\ \varrho\gamma_1\gamma_2 & \gamma_2^2 \end{pmatrix}.$$

It is not hard to see that any matrix on the right can be expressed as such covariance matrix. The differential characteristics $(b^{(G, \sigma^2)}, c^{(G, \sigma^2)}, F^{(G, \sigma^2)})$ of (G, σ^2) are obtained from (4.1)-(4.2) as

$$\begin{aligned} b_t^{(G, \sigma^2)} &= \begin{pmatrix} \sigma_t \alpha \\ \beta + \sigma_t^2 \delta \end{pmatrix}, \\ c_t^{(G, \sigma^2)} &= \begin{pmatrix} \sigma_t^2 E(Z^2) & \sigma_t^3 \lambda E(Z^3) \\ \sigma_t^3 \lambda E(Z^3) & \sigma_t^4 \lambda^2 \text{Var}(Z^2) \end{pmatrix}, \\ F_t^{(G, \sigma^2)} &= 0. \end{aligned}$$

As in the previous section, we consider $(G, \log \sigma^2) = g(G, \sigma^2)$ for technical reasons. Applying Proposition A.4 yields its differential characteristics (b, c, F) as follows:

$$\begin{aligned} b_t &= \begin{pmatrix} \sigma_t \alpha \\ \beta \sigma_t^{-2} + \delta - \frac{1}{2} \lambda^2 \text{Var}(Z^2) \end{pmatrix}, \\ c_t &= \begin{pmatrix} \sigma_t^2 E(Z^2) & \sigma_t \lambda E(Z^3) \\ \sigma_t \lambda E(Z^3) & \lambda^2 \text{Var}(Z^2) \end{pmatrix}, \\ F_t &= 0. \end{aligned}$$

We are now ready to state the convergence result which parallels Theorem 3.8.

Theorem 4.1 *A sequence $(G^n, (\sigma^n)^2) \in \text{GARCh}_n(\eta_n, \beta_n, \lambda_n, \delta_n, Q_n)$ with parameters as given in (4.4)-(4.4) converges in law to (G, σ^2) as in (4.1)-(4.2) with initial law η .*

Proof. This is shown along the same lines as Theorem 3.8, some steps actually being simpler. \square

A Semimartingale characteristics and convergence

We generally use the notation of [JS03]. By $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ we denote a fixed *truncation function*, i.e. h is bounded by some constant M_h and $h(x) = x$ holds in some open neighbourhood U_h of 0. Unless otherwise stated, we assume that the characteristics are given with respect to h .

A.1 Differential characteristics

This paper relies heavily on the *calculus* of semimartingale characteristics. For the convenience of the reader we summarize a few basic properties which can be found in [JS03] or [Kal06], respectively.

Definition A.1 Let (B, C, ν) be the characteristics of a semimartingale X . If there are predictable processes b, c and a transition kernel F from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(\mathbb{R}^d, \mathcal{B}^d)$ such that

$$\begin{aligned} B_t &= \int_0^t b_s ds, \\ C_t &= \int_0^t c_s ds, \\ \nu([0, t] \times A) &= \int_0^t F_s(A) ds \quad \forall A \in \mathcal{B}^d, \end{aligned}$$

we call (b, c, F) the *differential characteristics* of X and denote them by ∂X . We implicitly assume that (b, c, F) is a good version in the sense that the values of c are non-negative symmetric matrices, $F_s(\{0\}) = 0$ and $\int (1 \wedge |x|^2) F_s(dx) < \infty$.

From an intuitive viewpoint one can interpret differential characteristics as a local Lévy-Khintchine triplet. Very loosely speaking, a semimartingale with differential characteristics (b, c, F) resembles locally after t a Lévy process with triplet $(b, c, F)(\omega, t)$, i.e. with drift rate b , diffusion matrix c , and jump measure F . For discrete or more precisely piecewise constant processes differential characteristics do not exist. Nevertheless, the triplet (B, C, ν) quantifies the local dynamics of the process in some sense (cf. [JS03] for details).

Proposition A.2 *An \mathbb{R}^d -valued semimartingale X with $X_0 = 0$ is a Lévy process if and only if it has a version (b, c, F) of the differential characteristics which does not depend on (ω, t) . In this case (b, c, F) is equal to the Lévy-Khintchine triplet.*

Proposition A.3 *Let X be an \mathbb{R}^d -valued semimartingale and H an $\mathbb{R}^{n \times d}$ -valued predictable process with $H^{j \cdot} \in L(X), j = 1, \dots, n$ (i.e. integrable with respect to X). If $\partial X = (b, c, F)$, then the differential characteristics of the \mathbb{R}^n -valued integral process*

$$\int_0^\cdot H_t dX_t := \left(\int_0^\cdot H_t^{j \cdot} dX_t \right)_{j=1, \dots, n}$$

are given by $\partial(\int_0^\cdot H_t dX_t) = (\tilde{b}, \tilde{c}, \tilde{F})$, where

$$\begin{aligned} \tilde{b}_t &= H_t b_t + \int (\tilde{h}(H_t x) - H_t h(x)) F_t(dx), \\ \tilde{c}_t &= H_t c_t H_t^\top, \\ \tilde{F}_t(A) &= \int 1_A(H_t x) F_t(dx) \quad \forall A \in \mathcal{B}^n \text{ with } 0 \notin A. \end{aligned}$$

Here $\tilde{h} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ denotes the truncation function which is used on \mathbb{R}^n .

Proposition A.4 *Let X be an \mathbb{R}^d -valued semimartingale with differential characteristics $\partial X = (b, c, F)$. Suppose that $f : U \rightarrow \mathbb{R}^n$ is twice continuously differentiable on some open subset $U \subset \mathbb{R}^d$ such that X, X_- are U -valued. Then the \mathbb{R}^n -valued semimartingale $f(X)$ has differential characteristics $\partial(f(X)) = (\tilde{b}, \tilde{c}, \tilde{F})$, where*

$$\begin{aligned} \tilde{b}_t^i &= \sum_{k=1}^d D_k f^i(X_{t-}) b_t^k + \frac{1}{2} \sum_{k,l=1}^d D_{kl} f^i(X_{t-}) c_t^{kl} \\ &\quad + \int \left(\tilde{h}^i(f(X_{t-} + x) - f(X_{t-})) - \sum_{k=1}^d D_k f^i(X_{t-}) h^k(x) \right) F_t(dx), \\ \tilde{c}_t^{ij} &= \sum_{k,l=1}^d D_k f^i(X_{t-}) c_t^{kl} D_l f^j(X_{t-}), \\ \tilde{F}_t(A) &= \int 1_A(f(X_{t-} + x) - f(X_{t-})) F_t(dx) \quad \forall A \in \mathcal{B}^n \text{ with } 0 \notin A. \end{aligned}$$

A.2 Convergence in law

Our main results are based on the following limit theorem in [JS03]. It states convergence of a sequence of processes, given that the characteristics converge, some majoration and continuity conditions hold, and the distribution of the limit is uniquely determined by its characteristics. For unexplained notation cf. [JS03].

Theorem A.5 ([JS03, IX.3.39]) *Let X be the canonical process $X_t(\alpha) = \alpha(t)$ on the path space $(\mathbb{D}(\mathbb{R}^d), \mathcal{D}(\mathbb{R}^d), \mathbf{D}(\mathbb{R}^d))$. We assume that X is a semimartingale relative to P with initial distribution $\eta = \mathcal{L}(X_0)$ and characteristics (B, C, ν) . Let X^n denote a sequence of \mathbb{R}^d -valued semimartingales (not necessarily on the same space) with characteristics (B^n, C^n, ν^n) . The truncation function h is supposed to be continuous and the same for all processes. Moreover, let D be a dense subset of \mathbb{R}_+ and set*

$$\begin{aligned} S_a &:= \inf\{t \in \mathbb{R}_+ : |X_t| \geq a \text{ or } |X_{t-}| \geq a\}, \\ S_a^n &:= \inf\{t \in \mathbb{R}_+ : |X_t^n| \geq a \text{ or } |X_{t-}^n| \geq a\}. \end{aligned}$$

Assume:

1. *The local strong majoration hypothesis: for all $a > 0$ there is an increasing continuous and deterministic function $F(a)$ such that the stopped processes $\sum_{i \leq d} \text{Var}(B^i)^{S_a}$ and $(\sum_{i \leq d} C^{ii} + (|x|^2 \wedge 1) * \nu)^{S_a}$ are strongly majorized by $F(a)$. Here, $\text{Var}(B^i)$ denotes the total variation process of the i th component of B .*

2. *The local condition on big jumps: for all $a > 0, t > 0$,*

$$\lim_{m \rightarrow \infty} \sup_{\alpha \in \mathbb{D}(\mathbb{R}^d)} \nu\left(\alpha; [0, t \wedge S_a(\alpha)] \times \{x : |x| > m\}\right) = 0.$$

3. *Local uniqueness for the martingale problem given by (B, C, ν) (more precisely for the martingale problem $\mathcal{M}(\sigma(X_0), X | \eta; B, C, \nu)$ in the language of [JS03]).*

4. *Continuity condition: for all $t \in D$ and all*

$$\begin{aligned} f &\in C(\mathbb{R}^d) \\ &:= \{f : \mathbb{R}^d \rightarrow \mathbb{R} : f \text{ is bounded, continuous, and } 0 \text{ in a neighbourhood of } 0\} \end{aligned}$$

*the functions $\alpha \mapsto B_t(\alpha), \tilde{C}_t(\alpha), f * \nu_t(\alpha)$ are Skorokhod-continuous on $\mathbb{D}(\mathbb{R}^d)$.*

5. $X_0^n \rightarrow X_0$ in law.

6. *The following three conditions hold:*

$$(a) \sup_{s \leq t} |B_{s \wedge S_a^n}^n - (B_{s \wedge S_a}) \circ X^n| \xrightarrow{P} 0 \text{ for all } t > 0, a > 0,$$

$$(b) \tilde{C}_{t \wedge S_a^n}^n - (\tilde{C}_{t \wedge S_a}) \circ X^n \xrightarrow{P} 0 \text{ for all } t \in D, a > 0,$$

$$(c) f * \nu_{t \wedge S_a^n}^n - (f * \nu_{t \wedge S_a}) \circ X^n \xrightarrow{P} 0 \text{ for all } t \in D, a > 0, f \in C(\mathbb{R}^d).$$

Then the laws $\mathcal{L}(X^n)$ converge weakly to $P = \mathcal{L}(X)$.

A.3 Characteristics and generator

[Jac79] investigates the link between the generator and the semimartingale characteristics of a Markov process. We are particularly interested in the following problem: Given a Markovian semimartingale X with differential characteristics (b, c, F) , how does its generator act on functions in

$$C_c^2(\mathbb{R}^d) := \{f : \mathbb{R}^d \rightarrow \mathbb{R} : f \text{ twice continuously differentiable with compact support}\}?$$

Up to a small gap, the answer follows directly from the results in [Jac79].

To make things more precise, let X denote the canonical process on the path space $(\mathbb{D}(\mathbb{R}^d), \mathcal{D}(\mathbb{R}^d), \mathbf{D}(\mathbb{R}^d))$. We assume that X is a Markov process which is a semimartingale relative to $P_x, x \in \mathbb{R}^d$ with $X_0 = x$ P_x -a.s. and differential characteristics (b, c, F) of the form

$$\begin{aligned} b_t^i(\omega) &:= \beta^i(X_{t-}(\omega)), \\ c_t^{ij}(\omega) &:= \gamma^{ij}(X_{t-}(\omega)), \\ F_t(\omega; A) &:= K(X_{t-}(\omega); A) \quad \forall A \in \mathcal{B}^d, \end{aligned} \tag{A.1}$$

where $\beta : \mathbb{R}^d \rightarrow \mathbb{R}^d, \gamma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ are measurable functions such that $\gamma(x)$ is symmetric and non-negative definite and K is a non-negative transition kernel from \mathbb{R}^d into \mathbb{R}^d satisfying $K(x, \{0\}) = 0$ and $\int (1 \wedge |y|^2) K(x; dy) < \infty$. Moreover, we assume that β, γ are continuous and $x \mapsto \int f(y) K(x; dy)$ is continuous for any bounded, continuous function f satisfying $f(y) \leq C(1 \wedge |y|^2)$ for some $C < \infty$, i.e. $x \mapsto (1 \wedge |y|^2) K(x; dy)$ is supposed to be weakly continuous. Finally, we assume

$$\sup \left\{ K(x; \{y \in \mathbb{R}^d : |x + y| \leq n\}) : |x| > 3n \right\} < \infty \tag{A.2}$$

for any $n \in \mathbb{N}$.

[Jac79, 13.55] states that X or rather $\mathcal{L}(X)$ solves a martingale problem related to the following linear operator on $C^2(\mathbb{R}^d)$, which here denotes the set of twice continuously differentiable functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ such that f and its first and second partial derivatives are bounded:

$$\begin{aligned} \mathcal{H}f(x) &= \sum_{i \leq d} \beta^i(x) D_i f(x) + \frac{1}{2} \sum_{i, j \leq d} \gamma^{ij}(x) D_{ij} f(x) \\ &\quad + \int \left(f(x + y) - f(x) - \sum_{i \leq d} h^i(y) D_i f(x) \right) K(x; dy). \end{aligned} \tag{A.3}$$

By [Jac79, 13.40] this operator \mathcal{H} coincides with the generator of X on the set

$$D := \{f \in C^2(\mathbb{R}^d) : t \mapsto E_x(\mathcal{H}f(X_t)) \text{ is right-continuous for all } x \in \mathbb{R}^d\}.$$

It remains to be shown that $C_c^2(\mathbb{R}^d) \subset D$. We begin with the following lemma, which is inspired by [DFS03].

Lemma A.6 For any $n \in \mathbb{N}$ there exists some $C_n < \infty$ such that

$$\sup_{|x| \leq n} |\mathcal{K} f(x)| \leq C_n \|f\|$$

for any $f \in C^2(\mathbb{R}^d)$, where

$$\|f\| := \sup_{x \in \mathbb{R}^d} \left\{ |f(x)| + \sum_{i \leq d} |D_i f(x)| + \sum_{i, j \leq d} |D_{ij} f(x)| \right\}. \quad (\text{A.4})$$

Moreover, $\sup_{x \in \mathbb{R}^d} |\mathcal{K} f(x)| < \infty$ for $f \in C_c^2(\mathbb{R}^d)$.

Proof. Let $n \in \mathbb{N}$. The first two terms of $\mathcal{K} f$ are bounded by a multiple of $\|f\|$ because β, γ are bounded. The integral term satisfies

$$\begin{aligned} & \int_{\mathbb{R}^d \setminus U_h} \left| f(x+y) - f(x) - \sum_{i \leq d} h^i(y) D_i f(x) \right| K(x; dy) \\ & \leq \int_{\mathbb{R}^d \setminus U_h} \left(2 \|f\| + \left| \sum_{i \leq d} D_i f(x) |h(y)| \right| \right) K(x; dy) \\ & \leq (2 + M_h) \|f\| K(x; \mathbb{R}^d \setminus U_h) \\ & = C_{1,n} \|f\| \end{aligned}$$

for some constant $C_{1,n} < \infty$ which depends on x only through n . In the last step we use the continuity of K . With Taylor's formula we can now estimate the remaining part. There exists some $z(x, y) \in \{x + \tau y : \tau \in (0, 1)\}$ such that

$$\begin{aligned} & \int_{U_h} \left| f(x+y) - f(x) - \sum_{i \leq d} h^i(y) D_i f(x) \right| K(x; dy) \\ & = \int_{U_h} \left| f(x+y) - f(x) - \sum_{i \leq d} y_i D_i f(x) \right| K(x; dy) \\ & = \int_{U_h} \left| \frac{1}{2} \sum_{i, j \leq d} D_{ij} f(z(x, y)) y_i y_j \right| K(x; dy) \\ & \leq \frac{1}{2} \|f\| \int_{U_h} \max\{|y_i y_j| : i, j \leq d\} K(x; dy) \\ & \leq \frac{1}{2} \|f\| \int_{U_h} |y|^2 K(x; dy) \\ & \leq C_{2,n} \|f\| \end{aligned}$$

for some constant $C_{2,n} < \infty$ which depends on x only through n (again by continuity of K). Together the first assertion follows.

In particular, we have $\sup_{|x| \leq 3n} |\mathcal{K} f(x)| < \infty$ for $f \in C_c^2(\mathbb{R}^d)$ with $\text{supp}(f) \subset K_n :=$

$\{x \in \mathbb{R}^d : |x| \leq n\}$. By (A.2) we have

$$\begin{aligned} \sup_{|x|>3n} |\mathcal{K}f(x)| &= \sup_{|x|>3n} \left| \int_{\mathbb{R}^d} f(x+y)K(x;dy) \right| \\ &\leq \|f\| \sup_{|x|>3n} K(x; \{y \in \mathbb{R}^d : |x+y| \leq n\}) \\ &< \infty, \end{aligned}$$

which yields the second claim. \square

Theorem A.7 *The infinitesimal generator of X coincides with \mathcal{K} on the set $C_c^2(\mathbb{R}^d)$.*

Proof. [Jac79, 13.55] states that X is a solution to the martingale problem corresponding to $(\mathcal{K}, C^2(\mathbb{R}^d))$. Applying [Jac79, 13.40] yields that \mathcal{K} is the infinitesimal generator on the set

$$D := \{f \in C^2(\mathbb{R}^d) : t \mapsto E_x(\mathcal{K}f(X_t)) \text{ is right-continuous for all } x\}.$$

$\mathcal{K}f(X_t)$ is bounded for every $f \in C_c^2(\mathbb{R}^d)$ by Lemma A.6. In view of dominated convergence, it remains to verify that $x \mapsto \mathcal{K}f(x)$ is continuous. The first two terms of $\mathcal{K}f(x)$ are continuous in x by continuity of β, γ . For the integral part let $x_n \rightarrow x$ and $\varepsilon > 0$. Define

$$g(x, y) := f(x+y) - f(x) - \sum_{i \leq d} h^i(y) D_i f(x).$$

Since f and its first two partial derivatives are uniformly continuous, h is bounded and K is weakly continuous, there is some $N \in \mathbb{N}$ such that

$$\begin{aligned} \left| \int g(x, y) (K(x_n; dy) - K(x; dy)) \right| &\leq \frac{\varepsilon}{2}, \\ \left| \int (1 \wedge |y|^2) (K(x_n; dy) - K(x; dy)) \right| &\leq 1, \\ |g(x_n, y) - g(x, y)| &\leq \delta \quad \forall y \in \mathbb{R}^d, \\ \sup_{\xi \in \mathbb{R}^d} |D_{ij}f(x_n + \xi) - D_{ij}f(x + \xi)| &\leq \delta d^{-2} \quad \text{for } i, j \leq d \end{aligned}$$

for all $n \geq N$, where

$$\delta := \frac{\varepsilon}{2} \left(\int (1 \wedge |y|^2) K(x; dy) + 1 \right)^{-1}.$$

W.l.o.g. $1 \leq |y|$ for $y \notin U_h$. A Taylor expansion yields

$$|g(x_n, y) - g(x, y)| \leq \sum_{i, j \leq d} \sup_{\xi \in \mathbb{R}^d} |D_{ij}f(x_n + \xi) - D_{ij}f(x + \xi)| |y_i| |y_j| \leq \delta |y|^2$$

for $y \in U_h$. Altogether, we have

$$\begin{aligned}
& \left| \int g(x_n, y) K(x_n; dy) - \int g(x, y) K(x; dy) \right| \\
& \leq \left| \int (g(x_n, y) - g(x, y)) K(x_n; dy) \right| + \left| \int g(x, y) (K(x_n; dy) - K(x; dy)) \right| \\
& \leq \int |g(x_n, y) - g(x, y)| K(x_n; dy) + \frac{\varepsilon}{2} \\
& \leq \int (\delta \wedge \delta |y|^2) K(x_n; dy) + \frac{\varepsilon}{2} \\
& \leq \delta \left(\int (1 \wedge |y|^2) K(x; dy) + 1 \right) + \frac{\varepsilon}{2} \\
& = \varepsilon.
\end{aligned}$$

for $n \geq N$. This yields the claim. \square

If the law of X is uniquely determined by b, c, F , it is determined by \mathcal{K} as well.

Proposition A.8 *Suppose that β, γ, K determine the law of X uniquely. More precisely, for any $x \in \mathbb{R}^d$ and any semimartingale \tilde{X} on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{F}}, \tilde{P})$ with $\tilde{X}_0 = x$ and differential characteristics of the form*

$$(\tilde{b}, \tilde{c}, \tilde{F})(t, \omega) = (\beta(\tilde{X}_{t-}(\omega)), \gamma(\tilde{X}_{t-}(\omega)), K(\tilde{X}_{t-}(\omega)))$$

we assume $\mathcal{L}(\tilde{X}) = P_x$ (which is also the law of X under P_x).

Moreover, let Y be any strong Markov process whose generator coincides with \mathcal{K} on $C_c^2(\mathbb{R}^d)$. Then Y has the same law as X for any starting value x .

Proof. For $g \in C^2(\mathbb{R}^d)$ define

$$C(g) := g(Y) - g(Y_0) - \int_0^\cdot \mathcal{K} g(Y_{s-}) ds.$$

Let $f \in C^2(\mathbb{R}^d)$. W.l.o.g. we assume $f(0) = 0$. There exists some function $k \in C_c^2(\mathbb{R})$ with $k(x) = x$ if $|x| \leq 1$. Define the sequence $f_n(x) := f(nk(x/n))$. Then $f_n \in C_c^2(\mathbb{R}^d)$ and $f_n(x) = f(x)$ for $|x| \leq n$. Moreover, the norm $\|f_n\|$ in the sense of (A.4) is bounded uniformly in n by some multiple of $\|f\|$. [Jac79, 13.38] yields that $C(f_n)$ is a local martingale for any n . Set $T_m := \inf\{t : |Y_t| \leq m\}$. By Lemma A.6 we have

$$|C(f_n)_{t^{T_m}}^{T_m}| \leq \|f\|(2 + C_m t)$$

for some finite constant C_m which does not depend on n . Consequently, $C(f_n)^{T_m}$ is a martingale for any m, n . Dominated convergence yields that $C(f)^{T_m}$ is a martingale as well, i.e. $C(f)$ is a local martingale. From [Jac79, 13.55] it follows that Y has differential characteristics

$$(\tilde{b}, \tilde{c}, \tilde{F})(t, \omega) = (\beta(\tilde{Y}_{t-}(\omega)), \gamma(\tilde{Y}_{t-}(\omega)), K(\tilde{Y}_{t-}(\omega))),$$

which by assumption implies $\mathcal{L}(Y) = \mathcal{L}(X)$ for any starting value x . \square

Acknowledgements

The authors want to thank the referees for valuable suggestions.

References

- [AL05] C. Alexander and E. Lazar. On the continuous limit of GARCH. ICMA Center Discussion Papers in Finance 2005-13, 2005.
- [BWZ03] L. Brown, Y. Wang, and L. Zhao. On the statistical equivalence at suitable frequencies of GARCH and stochastic volatility models with the corresponding diffusion models. *Statistica Sinica*, 13:993–1013, 2003.
- [DFS03] D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13:984–1053, 2003.
- [EK86] S. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. Wiley, New York, 1986.
- [Jac79] J. Jacod. *Calcul Stochastique et Problèmes de Martingales*, volume 714 of *Lecture Notes in Mathematics*. Springer, Berlin, 1979.
- [JS03] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, Berlin, second edition, 2003.
- [Kal75] O. Kallenberg. Limits of compound and thinned point processes. *Journal of Applied Probability*, 12:269–278, 1975.
- [Kal98] J. Kallsen. *Semimartingale Modelling in Finance*. Dissertation Universität Freiburg i. Br., 1998.
- [Kal06] J. Kallsen. A didactic note on affine stochastic volatility models. In Yu. Kabanov, R. Liptser, and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance*, pages 343–368. Springer, Berlin, 2006.
- [KLM04] C. Klüppelberg, A. Lindner, and R. Maller. A continuous-time GARCH process driven by a Lévy process: Stationarity and second order behaviour. *Journal of Applied Probability*, 41:601–622, 2004.
- [KP91] T. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability*, 19:1035–1070, 1991.
- [MMS07] R. Maller, G. Müller, and A. Szimayer. GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli*, 2007. To appear.

- [Nel90] D. Nelson. ARCH models as diffusion approximations. *Journal of Econometrics*, 45:7–38, 1990.
- [Pan05] G. Panayotov. *Three Essays on Volatility Issues in Financial Markets*. Ph.D. dissertation (University of Maryland), College Park, 2005.
- [Wan02] Y. Wang. Asymptotic nonequivalence of GARCH models and diffusions. *The Annals of Statistics*, 30:754–783, 2002.
- [Zhe06] Z. Zheng. Recrossing the bridge from discrete time to continuous time towards a complete model with stochastic volatility I. Preprint University of Wisconsin-Milwaukee, 2006.