AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

Weak diffusion limits of dynamic conditional correlation models

Christian M. Hafner, Sebastien Laurent and Francesco Violante

CREATES Research Paper 2015-2

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Weak diffusion limits of dynamic conditional correlation models

Christian M. Hafner [*] Sebastien Laurent [†] Francesco Violante [‡]

January 14, 2015

Abstract

The properties of dynamic conditional correlation (DCC) models are still not entirely understood. This paper fills one of the gaps by deriving weak diffusion limits of a modified version of the classical DCC model. The limiting system of stochastic differential equations is characterized by a diffusion matrix of reduced rank. The degeneracy is due to perfect collinearity between the innovations of the volatility and correlation dynamics. For the special case of constant conditional correlations, a non-degenerate diffusion limit can be obtained. Alternative sets of conditions are considered for the rate of convergence of the parameters, obtaining time-varying but deterministic variances and/or correlations. A Monte Carlo experiment confirms that the quasi approximate maximum likelihood (QAML) method to estimate the diffusion parameters is inconsistent for any fixed frequency, but that it may provide reasonable approximations for sufficiently large frequencies and sample sizes.

Keywords: cDCC, Weak diffusion limits, QAML, CCC, GARCH diffusion.

JEL Classification: C13, C22, C51

[*]Université catholique de Louvain, ISBA & CORE, B-1348 Louvain-la-Neuve, Belgium. `Christian.hafner@uclouvain.be`.

[†]Aix-Marseille University (Aix-Marseille School of Economics), 13089 Aix-en-Provence, CNRS & EHESS, IAE. `sebastien.laurent@univ-amu.fr`

[‡]Aarhus University & CREATES, 8210 Aarhus V, Denmark. `fviolante@econ.au.dk`.
Francesco Violante acknowledges financial support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation, and from The Danish Council for Independent Research, Social Sciences, Individual Research grant and Sapere Aude Research Talent grant.

1 Introduction

Continuous and discrete time volatility models are often considered as two competitive views to modeling financial time series. Thanks to the analytical tractability ensured by Itô calculus, continuous time models have played a central role in theoretical finance. The continuous time setting allows to have a deeper understanding of the properties of the corresponding discrete time model and to assess probabilistic and statistic properties of discrete time sequences such as stationarity, existence of moments or distributional results which are otherwise intractable in discrete time, see e.g. Nelson (1990), Nelson and Foster (1994) and Nelson (1994).

From an applied viewpoint, inference on continuous time parameters of stochastic volatility models represents an important issue. The intractable likelihood functions and the unobservable volatility process require sophisticated estimation procedures. Several estimation methods have been proposed, such as the simulation based method of moments, Duffie and Singleton (1993), the quasi-indirect inference of Broze, Scaillet, and Zakoian (1998) or Bayesian Markov chain Monte Carlo methods, Jones (2003). Bollerslev, Engle, and Nelson (1994) and Ghysels, Harvey, and Renault (1996) provide exhaustive surveys on stochastic volatility models. Therefore, discrete time volatility models have been most often preferred by the applied econometrician. Rather than estimating and forecasting with a diffusion model observed at discrete points in time, it is in fact often easier to use a discrete model directly.

The theory of convergence of discrete time Markov sequences towards continuous time diffusion processes, see e.g. Stroock and Varadhan (1979), Kushner (1984) and Ethier and Kurtz (1986), provides the theoretical foundation to establish mutual complementarities, possible inter-changeability and connections between the two approaches. Nelson (1990) provides conditions ensuring the weak convergence of a discrete time Markov chain, defined by a system of stochastic difference equations, towards a diffusion, defined by a system of stochastic differential equations. The proposed approach requires the convergence, as the interval between observations shrinks to zero, of a number of conditional moments to well defined limits at an appropriate rate. In the context of GARCH-type models, Nelson (1990), illustrates the convergence through various GARCH specifications. This approach has been used by Duan (1997) to derive the diffusion limit of the Augmented GARCH model, by Fornari and Mele (1997) to study the continuous time behavior of the class of non linear ARCH models proposed by Ding, Granger, and Engle (1993), by Alexander and Lazar (2005) to derive the diffusion limit of a weak GARCH process and in a related setting by Trifi (2006) to illustrate the convergence results for the CEV-ARCH model of Fornari and Mele (2006) and the CMSV model of Jeantheau (2004) and Hobson and Rogers (1998). In the multivariate case, apart from Nelson (1994) in the context of asymptotic filtering theory, to our knowledge, the relationship between discrete and continuous time volatility and correlation models has not been addressed yet.

The potential advantage of the Nelson approximation approach lies essentially in estimation and forecasting. Considering the discrete time model as a diffusion approximation suggests to infer the parameters of the diffusion model by the parameter estimates of a discrete time GARCH-type model. Hence, a natural alternative to the direct estimation of the diffusion parameters consists in inferring the diffusion parameters by means of a tractable likelihood function of an approximating discrete time multivariate GARCH process. Following Fornari and Mele (2006), this approach is called estimation by quasi-approximated maximum likelihood (QAML). Requiring a feasible computational effort, this approach has been advocated e.g. by Engle and Lee (1996), Lewis (2000), Barone-Adesi, Rasmussen, and Ravanelli (2005) and Stentoft (2011) among others. Its computational advantage becomes important in the multivariate case, where volatility models within the conditional correlation class can be estimated in a straightforward two-step procedure, estimating first the conditional variances, then conditional correlations. However, the main drawback of estimation by QAML is the difficulty of proving its consistency even if the discrete time approximation is closed under temporal aggregation, see Drost and Nijman (1993) and Drost and Werker (1996). In the univariate GARCH case, Wang (2002) has shown that the statistical experiments resulting from the estimation of the diffusion model and its approximating discrete time model are not equivalent, which would imply inconsistency of the QAML estimator also in the multivariate case.

In this paper we focus on conditional correlation models with GARCH dynamics for the variances of the marginal processes. We recover the diffusion limit of a modified version of the classical Dynamic Conditional Correlation (DCC) model of Engle (2002), called consistent DCC (cDCC), proposed by Aielli (2006). Unlike DCC, the cDCC model has a martingale difference property of the correlation dynamics and is therefore easier to treat from a theoretical viewpoint. For this specification, we derive the existence of a degenerate weak diffusion limit. The degeneracy is due to the particular structure of the discrete time model in which the noise propagation system of the variances and the one of the correlation driving process are perfectly correlated. This structure is preserved in the diffusion limit which is characterized by a diffusion matrix of reduced rank. More precisely, the diffusion of the variances and of the diagonal elements of the correlation driving process are pairwise governed by the same Brownian motion.

As a special case, we consider the Constant Conditional Correlation model (CCC) of Bollerslev (1990), which can be obtained from the cDCC under suitable parameter restrictions. The CCC-GARCH model is particularly interesting because, unlike the cDCC-GARCH process, it admits a non-degenerate diffusion and, in the bivariate specification, a closed form solution for the diffusion limit.

We then propose and discuss alternative sets of conditions regarding the speed of convergence of parameters of the cDCC-GARCH model. In this way, other types of degenerated diffusions can be obtained which are characterized by a stochastic price process while variances and/or correlations remain time varying but deterministic. In the same spirit of Corradi (2000), we then discuss what kind of processes can be obtained as Euler approximation of the alternative diffusions processes.

The paper is completed by a simulation study to investigate the performance of the QAML estimator of the diffusion parameters in our model framework. For the parameters characterizing the innovation in variances and in correlations, we find a negative bias in all cases, irrespective of the sample size, which only decreases as the time intervals shrinks to zero. This confirms the results of Wang (2002) that care needs to be taken in inferring diffusion parameters from a discrete type approximation when there is no statistical equivalence of the likelihood estimators. For the remaining model parameters, however, no substantial biases are found and the mean square error converges to zero as the sample size increases for a given time interval.

The paper is organized as follows. In Section 2 we present the theorem of weak convergence of discrete time Markov chains. In Section 3 we study the continuous time behavior of the cDCC and CCC models. We also present the degenerate diffusions induced by a reparameterization of the convergence conditions. In Section 4, we illustrate through a Monte Carlo simulation our convergence results. In Section 5 we conclude and discuss directions for further research. All proofs are provided in the Appendix.

2 Weak convergence of stochastic systems

In this section we introduce a set of conditions for the convergence of a system of discrete time stochastic difference equations towards system of stochastic differential equations based on the work of Stroock and Varadhan (1979), Kushner (1984), Ethier and Kurtz (1986) and Nelson (1990).

Let us define by $D\left([0,\infty), \mathbb{R}^N\right)$ the space of cadlag mappings from $[0,\infty)$ into $\mathbb{R}^N$ and $\mathcal{B}\left(\mathbb{R}^N\right)$ the Borel sets on $\mathbb{R}^N$. Pr_h is the probability measure on $D\left([0,\infty), \mathbb{R}^N\right)$ for each $h > 0$. Let $\mathcal{F}_{kh}$ be the σ-field generated by (kh, $X_0^{(h)}$, $X_h^{(h)}$, $X_{2h}^{(h)}$, ..., $X_{kh}^{(h)}$), where $X_{kh}^{(h)}$ is an N-dimensional discrete time Markov chain indexed by $h > 0$, $k \in \mathbb{N}$, with ν_h a probability measure on $\left(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N)\right)$, such that $Pr_h[X_0^{(h)} \in \Gamma] = \nu_h(\Gamma)$ for any $\Gamma \in \mathcal{B}\left(\mathbb{R}^N\right)$ defines the distribution of the starting point $X_0^{(h)}$, and with transition probabilities $Pr_h[X_{(k+1)h}^{(h)} \in \Gamma | \mathcal{F}_{kh}] = \Pi_{h,kh}(X_{kh}^{(h)}, \Gamma)$, $\forall k \in \mathbb{N}$, $\Gamma \in \mathcal{B}\left(\mathbb{R}^N\right)$. Let us now define $X_t^{(h)}$ a continuous time process, formed from the discrete time process $X_{kh}^{(h)}$ as a cadlag step function with jumps at $h, 2h, 3h, ...$, such that $Pr_h[X_t^{(h)} = X_{kh}^{(h)}, kh < t < (k+1)h] = 1$. Finally, let X_t be a continuous time process obtained from $X_t^{(h)}$ by shrinking h towards zero. X_t represents the limiting diffusion process to which, under Assumptions 1 to 4 given below, the discrete time process $X_t^{(h)}$ weakly converges as $h \to 0$. Let Ω denote the space of $N \times N$ matrices, and $\Omega' \subset \Omega$ the set of symmetric positive semi-definite $N \times N$ matrices.

For the convergence results we need the following assumptions.

Assumption 1. *There exist a continuous function* $a(x,t) : \mathbb{R}^N \times [0,\infty) \to \Omega'$ *and a continuous measurable function* $b(x,t) : \mathbb{R}^N \times [0,\infty) \to \mathbb{R}^N$ *such that for all* $r > 0$

and $(k-1)h < t < kh$

$$a)\ \lim_{h\to 0}\sup_{\|x\|\leqslant r}\left\| h^{-1}E\left[X^{(h)}_{(k+1)h} - X^{(h)}_{kh}\middle| X^{(h)}_{kh}=x\right] - b(x,t)\right\| = 0, \tag{1}$$

$$b)\ \lim_{h\to 0}\sup_{\|x\|\leqslant r}\left\| h^{-1}E\left[(X^{(h)}_{(k+1)h} - X^{(h)}_{kh})(X^{(h)}_{(k+1)h} - X_{kh})^{(h)\prime}\middle| X^{(h)}_{kh}=x\right] - a(x,t)\right\| = 0, \tag{2}$$

$$c)\ \exists\delta>0 : \lim_{h\to 0}\sup_{\|x\|\leqslant r}\left\| h^{-1}E\left[\left|(X^{(h)}_{(k+1)h} - X^{(h)}_{kh})_i\right|^{2+\delta}\middle| X^{(h)}_{kh}=x\right]\right\| = 0,$$ *where* $(.)_i$ *is the* i^{th} *element of the vector* $(X^{(h)}_{(k+1)h} - X^{(h)}_{kh})$.

Assumption 2. *There exists a continuous function* $\sigma(x,t) : I\!R^N \times [0,\infty) \to \Omega$ *such that for all* $x \in I\!R^N$ *and* $t \geq 0$, $a(x,t) = \sigma(x,t)\sigma(x,t)'$.

Assumption 3. $X^{(h)}_0$ *converges in distribution, as* $h \to 0$, *to a random variable* X_0 *with probability measure* ν_0 *on* $(I\!R^N, \mathcal{B}(I\!R^N))$.

Assumption 4. $\nu_0, b(x,t), a(x,t)$ *uniquely specify the distribution of a diffusion process* X_t *with initial distribution* ν_0, *drift vector* $b(x,t)$ *and diffusion matrix* $a(x,t)$.

We can now state the following theorem for the weak convergence of discrete time stochastic sequences.

Theorem of weak convergence (Nelson, 1990). *Under Assumptions 1 to 4, the sequence of discrete time process* $X^{(h)}_{kh}$ *indexed by* $h > 0$ $k \in I\!N$, *converges in distribution, as* $h \to 0$, *to the diffusion process* X_t, *i.e. the solution of the system of stochastic differential equations*

$$dX_t = b(X_t,t)dt + \sigma(X_t,t)dW_t, \tag{3}$$

where W_t *is an N-dimensional vector of mutually independent standard Brownian motions, independent from* X_0, *and with initial distribution* ν_0. *The process* X_t *exists, it is finite in finite time intervals almost surely, it is distributionally unique and its distribution does not depend on the choice of* $\sigma(x,t)$.

For the proof we refer to Nelson (1990). Conditions under which ν_0, $b(x,t)$ and $a(x,t)$ ensure finiteness of the process in finite time intervals and uniqueness of the limiting diffusion are extensively discussed in Stroock and Varadhan (1979), Ethier and Kurtz (1986), and Nelson (1990). To ensure weak existence, uniqueness and non-explosion of the diffusion process X_t on compact sets we rely on 'Condition A' of Nelson (1990), i.e.,

Condition 1 (weak existence and uniqueness)**.** *Let $a(x,t)$ and $b(x,t)$ be continuous in both x and t with two partial derivatives with respect to x.*

Following Theorem 10.2.2 of Stroock and Varadhan (1979), we impose the following conditions of non-explosiveness of X_t.

Condition 2 (non-explosiveness)**.** *For each $T > 0$, there is a $C_T < \infty$ such that*

$$\sup_{0 \leq t \leq T} \|a(x,t)\| \leq C_T(1 + |x|^2), \quad x \in \mathbb{R}^N$$

and

$$\sup_{0 \leq t \leq T} \langle x, b(x,t) \rangle \leq C_T(1 + |x|^2), \quad x \in \mathbb{R}^N.$$

These are not the weakest possible conditions, but they are easy to check and will suffice in our model framework.

3 Main theoretical results

Let $Y_{kh}^{(h)}$ be an N-dimensional vector of logarithmic prices indexed by kh, $k \in \mathbb{N}$, $h > 0$. The superscript (h) represents the time interval between two consecutive observations, i.e. for given h, prices are observed at times $h, 2h, 3h, ...$ We let the parameters depend on the sampling frequency. Furthermore, the variance of the innovations is made proportional to h. In this paper we focus on the covariance stationary case, hence usual suitable positivity and stationarity constraints on the parameters of the variances and correlation driving process apply, see Bollerslev (1986), Engle (2002), Aielli (2006) and Aielli (2013).

In the remainder of the paper we use the following operators: vec() stacks the columns of a matrix into a vector, vech() stacks the lower triangular portion of a square matrix into a vector, vechl() stacks the strictly lower triangular portion of a square matrix into a vector (i.e., excluding the diagonal elements), diag() stacks the diagonal of a square matrix into a vector. Furthermore, 1_N is a $(N \times 1)$ vector of ones and I_N is the $(N \times N)$ identity matrix. We also make use of the following elementary matrices: D_N denotes the $(N^2 \times N(N+1)/2)$ duplication matrix, which for any symmetric matrix A transforms vech(A) into vec(A), D_N^+ its generalized inverse, see e.g. Lütkepohl (1996) for details, I^* is defined such that

$\mathrm{diag}(A) = I^{*}\mathrm{vech}(A)$ with $I^{*} = I^{+\prime}D_N$ and $I^{+} = (1_N \otimes I_N) \odot [1_N' \otimes \mathrm{vec}(I_N)]$ transforms $\mathrm{vec}(A)$ into $\mathrm{diag}(A)$. Finally, I^{-} is defined such that $\mathrm{vechl}(A)\mathrm{vechl}(A)' = I^{-}(\mathrm{vech}(A)\mathrm{vech}(A)')I^{-\prime} = I^{-}D_N^{+}(A \otimes A)D_N^{+\prime}I^{-\prime}$.[1]

3.1 The cDCC-GARCH process

The discrete time cDCC-GARCH process of Aielli (2006) for the log price vector of an N-dimensional portfolio of assets Y_t is specified as follows:

$$
\begin{aligned}
Y_t &= Y_{t-1} + S_t\,\eta_t, \quad \eta_t|\mathcal{F}_{t-1} \sim \mathrm{N}(0, R_t), \quad t \in \mathbb{N} && (4)\\
V_{t+1} &= c + A\,S_t^2(\eta_t \odot \eta_t) + B\,V_t, && (5)\\
Q_{t+1} &= \bar{Q} + \vartheta P_t\eta_t\eta_t{}' P_t + \gamma Q_t, && (6)\\
P_t &= (Q_t \odot I_N)^{1/2}, && (7)\\
R_t &= P_t^{-1}\,Q_t\,P_t^{-1}, && (8)
\end{aligned}
$$

where $\mathcal{F}_t = \sigma(Y_\tau, \tau \leq t)$, S_t is a diagonal matrix with positive diagonal elements, $V_t = \mathrm{diag}(S_t^2)$. The parameters of the model are: $\bar{Q}$ $(N \times N)$ positive definite; A, B $(N \times N)$ diagonal with positive diagonal elements, c $(N \times 1)$ elementwise positive, and ϑ, γ: positive scalars.

The standard DCC model of Engle (2002) is very similar but instead of (6) specifies $Q_{t+1} = \bar{Q} + \vartheta\eta_t\eta_t' + \gamma Q_t$. The advantage of the cDCC model is that the recursion for Q_t preserves the martingale difference property, i.e., $\mathrm{E}[P_t\eta_t\eta_t'P_t - Q_t|\mathcal{F}_t] = 0$. Hence, the process $\{P_t\eta_t, \mathrm{vech}(Q_t)\}$ is a multivariate semi-strong GARCH process in the sense of Drost and Nijman (1993) and Hafner (2008).

Consider now the properties of the system of stochastic difference equations (4)-(8) as the time is partitioned more and more finely as in Section 2, letting the parameters and the covariance matrix of innovations depend on the length h of time intervals. We begin by partitioning time in (4)-(8), for $h > 0$ and $k \in \mathbb{N}$, according

[1] Examples: $I^{+}_{N=3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}'$ and $I^{-}_{N=3} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$.

to the following scheme

$$
\begin{aligned}
Y_{kh}^{(h)} &= Y_{(k-1)h}^{(h)} + S_{kh}^{(h)}\,\eta_{kh}^{(h)}, && (9)\\
V_{(k+1)h}^{(h)} &= c_h + A_h h^{-1}\, S_{kh}^{(h)2}(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)}) + B_h\, V_{kh}^{(h)}, && (10)\\
Q_{(k+1)h}^{(h)} &= \bar{Q}_h + \vartheta_h h^{-1} P_{kh}^{(h)}\,\eta_{kh}^{(h)}\,\eta_{kh}^{(h)\prime}\,P_{kh}^{(h)} + \gamma_h\, Q_{kh}^{(h)}, && (11)\\
P_{kh}^{(h)} &= (Q_{kh}^{(h)} \odot I_N)^{1/2}, && (12)\\
R_{kh}^{(h)} &= P_{kh}^{(h)-1}\, Q_{kh}^{(h)*}\, P_{kh}^{(h)-1}, && (13)
\end{aligned}
$$

and

$$
Pr_h[(Y_0, V_0, Q_0) \in \Gamma] = \nu_h(\Gamma) \text{ for any } \Gamma \in \mathcal{B}\left(\mathrm{I\!R}^{N(N+5)/2}\right) \qquad (14)
$$

where $\eta_{kh}^{(h)}$ is an $(N \times 1)$ vector of standardized but conditionally correlated innovations such that $R_{kh}^{(h)-1/2}\eta_{kh}^{(h)} \sim$ i.i.d. $\mathrm{N}(0, hI_N)$. Further, $S_{kh}^{(h)}$ is an $(N \times N)$ diagonal matrix of conditional standard deviations with the $(N \times 1)$ vector of conditional variances denoted by $V_{kh}^{(h)} = \mathrm{diag}(S_{kh}^{(h)2})$. For the correlation driving process $Q_{kh}^{(h)}$ we have, under the restrictions $\bar{Q}_h \in \Omega', \vartheta_h, \gamma_h \geq 0$, that $Q_{kh}^{(h)} \in \Omega'$. We will denote the non-redundant elements of $Q_{kh}^{(h)}$ by $q_{kh}^{(h)} = \mathrm{vech}(Q_{kh}^{(h)})$. Finally, (14) defines the distribution of the initial value of the random vector (Y_0, V_0, Q_0). Note that, for a given h, the vector $(Y_{kh}^{(h)}, V_{(k+1)h}^{(h)}, q_{(k+1)h}^{(h)})'$ is a discrete time Markov process, so that the theory of Section 2 applies. Again, note also that $h^{-1}\mathrm{E}[(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)}) - Q_{kh}^{(h)}|\mathcal{F}_{kh}] = 0$, so that, for a given h, the process $\{h^{-1/2}P_{kh}^{(h)}\eta_{kh}^{(h)}, q_{kh}^{(h)}\}$ is a multivariate semi-strong GARCH process.

Without loss of generality, we reparameterize the drift in the recursion $Q_{kh}^{(h)}$ as a combination of a frequency invariant component and frequency dependent parameters. The drift $\bar{Q}_h$ can be expressed as $\bar{Q}_h = (1 - \vartheta_h - \gamma_h)\bar{Q}$.[2] This transformation will be particularly useful when deriving the diffusion limit of the Constant Conditional Correlation (CCC) model of Bollerslev (1990). In fact, under the parameter restriction $\vartheta_h = \gamma_h = 0$, $Q_{kh}^{(h)} = \bar{Q}_h = \bar{Q}$, and therefore $R_{kh}^{(h)} = \bar{Q}$ for all h.[3] We denote $\bar{q} = \mathrm{vech}(\bar{Q})$ the non redundant elements of $\bar{Q}$.

[2]The same transformation can be carried out for the intercept of the $V_{(k+1)h}^{(h)}$ process, i.e., $c_h = (I_N - A_h - B_h)\bar{c}$. The vector $\bar{c}$ is frequency invariant and contains the (rescaled) unconditional variances of the return process $(Y_{(k+1)h}^{(h)} - Y_{kh}^{(h)})$, i.e., $\bar{c} = \mathrm{E}[(Y_{(k+1)h}^{(h)} - Y_{kh}^{(h)}) \odot (Y_{(k+1)h}^{(h)} - Y_{kh}^{(h)})]/h = \mathrm{E}[V_{(k+1)h}^{(h)}]$, $\forall h$.

[3]Note that even though in general $\bar{Q}$ does not need to be a correlation matrix, i.e. $\mathrm{diag}(\bar{Q}) = 1_N$, under the CCC parameter restrictions the diagonal elements of $\bar{Q}$ are not identifiable together

Before deriving the diffusion limit of cDCC-GARCH process we determine the convergence rates of the discrete time parameters for the moment conditions to converge as the sampling frequency increases, as required by Assumption 1.

Proposition 1. *Assumption 1 holds under the following convergence rates for the parameters of the discrete time cDCC-GARCH process (9)-(11),*

$$c_h = h\,c + o(h) \tag{15}$$

$$(A_h + B_h - I_N) = -h\,\Lambda + o(h) \tag{16}$$

$$A_h = \sqrt{h}\,A + o(\sqrt{h}) \tag{17}$$

$$(1 - \vartheta_h - \gamma_h) = h\,\phi + o(h) \tag{18}$$

$$\vartheta_h = \sqrt{h}\,\vartheta + o(\sqrt{h}), \tag{19}$$

for some $(N \times 1)$ vector c, A and Λ are diagonal $N \times N$ matrices with positive diagonal elements, and scalars $\phi, \vartheta > 0$.

The convergence rates in Proposition 1 ensure that the first and the second conditional moments per unit of time converge, as $h \to 0$, to well-behaved limits and that the first difference of the process $[Y_{kh}^{(h)\prime}, V_{kh}^{(h)\prime}, q_{kh}^{(h)\prime}]'$ satisfies Assumption 1.

Note that $c > 0$ (elementwise) ensures positivity of the variance process, $A > 0$ and $\vartheta > 0$ ensure that the rescaled second conditional moment does not vanish as $h \to 0$,[4] while $\Lambda > 0$ and $\phi > 0$ ensure covariance stationarity of the return process.[5]

Under our assumptions, we have the following result of the diffusion limit of the cDCC-GARCH process.

Theorem 1 (Diffusion limit of the cDCC-GARCH model). *Under (15) to (19), the discrete time cDCC-GARCH process (9)-(11) weakly converges to the diffusion process $X_t = [Y_t', V_t', q_t']'$ which is the solution to the system of stochastic*

with the intercept of the $V_{(k+1)h}^{(h)}$ process. Fixing $\mathrm{diag}(\bar{Q}) = 1_N$ ensures that: (i) $\mathrm{E}[V_{(k+1)h}^{(h)}] = (I_N - A_h - B_h)^{-1} c_h$ is the rescaled unconditional variance of the return process ($\mathrm{E}[(Y_{(k+1)h}^{(h)} - Y_{kh}^{(h)}) \odot (Y_{(k+1)h}^{(h)} - Y_{kh}^{(h)})]/h$), (ii) $\bar{Q}$ can be directly interpreted as the (un)conditional correlation of $(Y_{(k+1)h}^{(h)} - Y_{kh}^{(h)})$.

[4] These conditions imply that the diffusion limit of the cDCC-GARCH process converges to a continuous time stochastic volatility process.

[5] In the univariate setting, two special cases, $\Lambda = 0$ (integrated variance) and $\Lambda < 0$ (strictly stationary but not covariance stationary GARCH process) are also discussed in Nelson (1990). In this paper we restrict the analysis to the covariance stationary case.

differential equations

$$dX_t = b(X_t,t)dt + \sigma(X_t,t)dW_t, \tag{20}$$

where the drift, $b(X_t,t)$, is given by

$$b(X_t,t) = \begin{bmatrix} 0 \\ c - \Lambda V_t \\ \phi(\bar{q} - q_t) \end{bmatrix}, \tag{21}$$

while the scale, $\sigma(X_t,t)$, is a continuous mapping such that, for all $X_t \in I\!R^{N(N+5)/2}$ and $t \geq 0$, $a(X_t,t) = \sigma(X_t,t)\sigma(X_t,t)'$ where $a(X_t,t)$ is given by

$$a(X_t,t) = \begin{bmatrix} a_{YY} & 0 & 0 \\ 0 & a_{VV} & a_{Vq} \\ 0 & a'_{Vq} & a_{qq} \end{bmatrix}, \tag{22}$$

with

$$\begin{aligned} a_{YY} &= S_t R_t S_t \\ a_{VV} &= 2AS_t^2(R_t \odot R_t)S_t^2 A \\ a_{Vq} &= \vartheta\left[I^* K_t D_N^+(P_t \otimes P_t)D_N^{+\prime} - 1_N q_t'\right]' S_t^2 A \\ a_{qq} &= \vartheta^2[D_N^+(P_t \otimes P_t)D_N^{+\prime} K_t D_N^+(P_t \otimes P_t)D_N^{+\prime} - q_t q_t'] \end{aligned}$$

where $K_t = 2D_N^+(R_t \otimes R_t)D_N^{+\prime} + vech(R_t)vech(R_t)'$. The matrix $a(X_t,t)$ is singular and its rank is equal to $N(N+3)/2 < \dim(a(X_t,t)) = N(N+5)/2$. The conditional correlation, R_t, is given at each point in time by (13).

Note first that the drift term $b(X_t,t)$ is linear in X_t, which is due to the fact that the cDCC-GARCH process satisfies a semi-strong GARCH structure, meaning that increments to the state variables have a conditional mean that is linear in the state. In particular, as shown in the proof, we can use that in the cDCC model $\mathrm{E}[\Delta q^{(h)}_{(k+1)h}|\mathcal{F}_{kh}] = (1-\vartheta_h-\gamma_h)(\bar{q}-q^{(h)}_{kh})$. This is, however, not the case in the DCC model, where this expectation would be a function of the conditional correlation matrix $R^{(h)}_{kh}$, which is a nonlinear function of the state variable $q^{(h)}_{kh}$. Therefore, it appears very difficult if not impossible to obtain analytical results for the diffusion limit of the DCC model.

The singularity of $a(X_t,t)$ is due to the particular structure of the model in which the noise propagation of the variance processes and the one of the diagonal elements of the correlation driving processes are pairwise perfectly correlated. This is because, although (possibly) different in terms of level and dynamics, (10) and (11) are driven by the same innovations. For example, in the special case where $(I_N - A_h - B_h)^{-1}c_h = \mathrm{diag}(\bar{Q})$, $A_h = \vartheta_h I_N$, $B_h = \gamma_h I_N$, the model reduces to a scalar VEC model with N redundant equations.

To investigate the implications of singularity of the diffusion matrix $a(X_t, t)$, let us rearrange the order of the elements of the diffusion process X_t as $[Y_t', V_t', q_t^{(d)\prime}, q_t^{(l)\prime}]'$, where $q_t^{(d)} = \mathrm{diag}(Q_t)$ and $q_t^{(l)} = \mathrm{vechl}(Q_t)$. The two partial diffusion processes $[Y_t', V_t', q_t^{(l)\prime}]'$ and $[Y_t', q_t^{(d)\prime}, q_t^{(l)\prime}]'$ share the same correlation structure, while $\mathrm{Corr}(\mathrm{d}V_{t,i}, \mathrm{d}Q_{t,ii}) = 1\ \forall i$ implies that the two partial diffusions are driven by the same vector of Brownian innovations. Thus, the relevant part in terms of noise propagation system of the diffusion limit of the cDCC-GARCH process consists of a system of $N(N+3)/2$ stochastic differential equations, either $[Y_t', V_t', q_t^{(l)\prime}]$ or $[Y_t', q_t^{(d)\prime}, q_t^{(l)\prime}]$, while the remaining N diffusion processes, $q_t^{(d)}$ or V_t respectively, are characterized by a specific deterministic part (drift) but a common, though appropriately rescaled, stochastic component. To illustrate this point, let us consider the following partition of the diffusion matrix in (22), whose elements have been appropriately reordered,

$$a(X_t, t) = \begin{bmatrix} a_{YY} & 0 & 0 & 0 \\ 0 & a_{VV} & a_{Vq^{(d)}} & a_{Vq^{(l)}} \\ 0 & a'_{Vq^{(d)}} & a_{q^{(d)}q^{(d)}} & a_{q^{(d)}q^{(l)}} \\ 0 & a'_{Vq^{(l)}} & a'_{q^{(d)}q^{(l)}} & a_{q^{(l)}q^{(l)}} \end{bmatrix} \tag{23}$$

where

$$\begin{aligned}
a_{Vq^{(d)}} &= 2\vartheta A S_t^2 (R_t \odot R_r) P_t^2 = \vartheta a_{VV} (S_t^2 A)^{-1} P_t^2 \\
a_{q^{(d)}q^{(d)}} &= 2\vartheta^2 P_t^2 (R_t \odot R_r) P_t^2 = \vartheta^2 P_t^2 (A S_t^2)^{-1} a_{VV} (S_t^2 A)^{-1} P_t^2 \\
a_{Vq^{(l)}} &= \vartheta A S^2 \left[I^* (D_N^+ K_t D_N^{+\prime}) \left(D_N^+ (P_t \otimes P_t) D_N^{+\prime} \right) I^{-\prime} - 1_N q_t' I^{-\prime} \right] \\
a_{q^{(d)}q^{(l)}} &= \vartheta^2 P_t^2 [I^* (D_N^+ K_t D_N^{+\prime})(D_N^+ (P_t \otimes P_t) D_N^{+\prime}) I^{-\prime} - 1_N q_t' I^{-\prime}] = \vartheta P_t^2 (A S_t^2)^{-1} a_{Vq^{(l)}} \\
a_{q^{(l)}q^{(l)}} &= \vartheta^2 I^- [(D_N^+ (P_t \otimes P_t) D_N^{+\prime}) K_t (D_N^+ (P_t \otimes P_t) D_N^{+\prime}) - q_t q_t'] I^{-\prime}.
\end{aligned}$$

Let us also define $C_t = \vartheta P_t^2 (A S_t^2)^{-1}$. We can rewrite (23) as

$$\begin{bmatrix} a_{YY} & 0 & 0 & 0 \\ 0 & a_{VV} & a_{VV} C_t' & a_{Vq^{(l)}} \\ 0 & C_t a_{VV} & C_t a_{VV} C_t' & C_t a_{Vq^{(l)}} \\ 0 & a'_{Vq^{(l)}} & a'_{Vq^{(l)}} C_t' & a_{q^{(l)}q^{(l)}} \end{bmatrix}. \tag{24}$$

Therefore,

$$a([Y_t', V_t', q_t^{(l)\prime}]', t) = \begin{bmatrix} a_{YY} & 0 & 0 \\ 0 & a_{VV} & a_{Vq^{(l)}} \\ 0 & a_{Vq^{(l)}}' & a_{q^{(l)}q^{(l)}} \end{bmatrix} \tag{25}$$

$$a([Y_t', q_t^{(d)\prime}, q_t^{(l)\prime}]', t) = \begin{bmatrix} a_{YY} & 0 & 0 \\ 0 & C_t a_{VV} C_t' & C_t a_{Vq^{(l)}} \\ 0 & a_{Vq^{(l)}}' C_t' & a_{q^{(l)}q^{(l)}} \end{bmatrix}. \tag{26}$$

The decomposition in (25) and (26) shows that the two partial processes $[V_t', q_t^{(l)\prime}]'$ and $[q_t^{(d)\prime}, q_t^{(l)\prime}]'$, both uncorrelated with Y_t, share the same correlation structure. Furthermore, from Theorem 1, it immediately follows that $[V_t', q_t^{(l)\prime}]'$ and $[q_t^{(d)\prime}, q_t^{(l)\prime}]'$ are elementwise perfectly correlated[6], which implies that the two diffusion processes are driven by the same vector of Brownian motions. However, although either partial diffusion process $[Y_t', V_t', q_t^{(l)\prime}]'$ or $[Y_t', q_t^{(d)\prime}, q_t^{(l)\prime}]'$ is sufficient alone to fully characterize the noise propagation system of the cDCC diffusion limit, they are both necessary to characterize the distributions of Y_t and V_t which depend on both V_t and $q_t = [q_t^{(d)\prime}, q_t^{(l)\prime}]'$ through the correlation process R_t.[7]

3.2 A special case: the CCC-GARCH process

As a special case, consider the Constant Conditional Correlation (CCC) model of Bollerslev (1990). The cDCC process nests the CCC process under the following parameter restrictions

$$\vartheta_h = \gamma_h = 0 \quad \forall h.$$

Thus, the CCC-GARCH process can be written as

$$Y_{kh}^{(h)} = Y_{(k-1)h}^{(h)} + S_{kh}^{(h)} \eta_{kh}^{(h)}, \tag{27}$$

$$V_{(k+1)h}^{(h)} = c_h + A_h h^{-1} S_{kh}^{(h)2} (\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)}) + B_h V_{kh}^{(h)}, \tag{28}$$

where c_h, A_h, B_h are defined as before and $\eta_{kh}^{(h)}$ is an $(N \times 1)$ vector of standardized but correlated innovations, such that $\eta_{kh}^{(h)} \sim \mathrm{N}(0, hR)$, where R represents

[6] More generally, $\mathrm{Corr}(\mathrm{d}V_{t,i}, \mathrm{d}V_{t,j}) = \mathrm{Corr}(\mathrm{d}q_{t,i}^{(d)}, \mathrm{d}q_{t,j}^{(d)}) = \mathrm{Corr}(\mathrm{d}V_{t,i}, \mathrm{d}q_{t,j}^{(d)}) = (R_t \odot R_t)_{ij}\ \forall i, j = 1, ..., N$.

[7] Note that the partial system $[Y_t', q_t^{(d)\prime}, q_t^{(l)\prime}]'$ is however sufficient to characterize the distribution of the correlation driving process Q_t and hence of the correlation R_t.

the (frequency invariant) constant conditional correlation matrix. This model, although rather restrictive in practice, is particularly interesting because, unlike the cDCC-GARCH process, it allows for a non-degenerate diffusion and, in the bivariate specification, a closed form solution for the diffusion limit. The rates of convergence for the parameters and the CCC-GARCH process are stated in Proposition 2 and Theorem 2.

Proposition 2. *Under the following convergence rates for the parameters of the discrete time CCC-GARCH process (27)-(28)*

$$c_h = h\,c + o(h) \tag{29}$$

$$(A_h + B_h - I_N) = -h\,\Lambda + o(h) \tag{30}$$

$$A_h = \sqrt{h}\,A + o(\sqrt{h}), \tag{31}$$

for some $(N \times 1)$ vector c, $(N \times N)$ diagonal matrices A and Λ with positive and finite elements, Assumption 1 holds.

The same considerations on the parameters as in Proposition 1 hold by symmetry with the cDCC-GARCH process.

Theorem 2 (Diffusion limit of the CCC-GARCH model). *Under the convergence conditions in Proposition 2, the CCC-GARCH process (27)-(28) weakly converges to the non-degenerate diffusion process $X_t = [Y_t'\ V_t']'$ solution to a system of stochastic differential equations of the form (20), with drift*

$$b(X_t, t) = \begin{bmatrix} 0 \\ c - \Lambda V_t \end{bmatrix} \tag{32}$$

and diffusion matrix

$$a(X_t, t) = \begin{bmatrix} S_t\,R\,S_t & 0 \\ 0 & 2AS_t^2(R \odot R)S_t^2 A \end{bmatrix} \tag{33}$$

and driven by a vector W_t of $2N$ mutually independent Brownian motions, independent of the initial value $X_0 = [Y_0\ V_0]'$.

The diffusion limit of the CCC model is clearly non-degenerate because it is driven by as many Brownian motions as the number of variables in the system and whose covariance matrix is non-singular.

It is clear that the diffusion limit of the cDCC-GARCH process (as well as of the CCC-GARCH process) is a continuous time stochastic volatility model (i.e., stochastic variances and correlations). We discuss next the case when rates of convergence other than the ones introduced in Proposition 1, but still satisfying Assumption 1, are used.

3.3 Alternative convergence conditions

In this section we reconsider the continuous time approximation of the cDCC-GARCH process (9)-(11). The convergence rate $h^{1/2}$, suggested in Proposition 1, represents the slowest rate of convergence for the parameters A_h and ϑ_h satisfying Assumption 1. More generally, the rate $h^{1/2}$ represents the only rate ensuring that the second conditional moments $\mathrm{Var}(V^{(h)}_{(k+1)h} - V^{(h)}_{kh}|\mathcal{F}_{kh})$, $\mathrm{Var}(q^{(h)}_{(k+1)h} - q^{(h)}_{kh}|\mathcal{F}_{kh})$ and $\mathrm{Cov}[(V^{(h)}_{(k+1)h} - V^{(h)}_{kh}), (q^{(h)}_{(k+1)h} - q^{(h)}_{kh})|\mathcal{F}_{kh}]$ scaled by h^{-1}, do not vanish as $h \to 0$. As shown in Theorem 1, the resulting diffusion limit is characterized by stochastic variances of the marginal processes and stochastic correlation driving process.

However, there are other admissible convergence rates for A_h and ϑ_h which also satisfy Assumption 1. Thus, depending on the continuous time approximation we consider, we can recover different types of diffusion for the process (9)-(11).[8] This alternative set of results is shown in Proposition 3 and Theorem 3.

Proposition 3. *Assumption 1 holds under the following convergence rates for the parameters A_h and ϑ_h*

$$\lim_{h\to 0} h^{-(\frac{1}{2}+\delta_1)} A_h = \tilde{A} < \infty \tag{34}$$

and

$$\lim_{h\to 0} h^{-(\frac{1}{2}+\delta_2)} \vartheta_h = \tilde{\vartheta} < \infty, \tag{35}$$

for some $(N \times N)$ diagonal matrix $\tilde{A} > 0$ (elementwise), $\tilde{\vartheta} > 0$, $\delta_1 \geq 0$ and $\delta_2 \geq 0$.

Note that under (34) and (35), A_h and ϑ_h are of order $h^{1/2+\delta_1}$ and $h^{1/2+\delta_2}$, respectively. Clearly, the special case $\delta_1 = \delta_2 = 0$ is covered by Proposition 1.

Proposition 3 suggests alternative sets of conditions regarding the speed of convergence of the discrete time parameters under which Assumption 1 holds. The

[8] The following arguments can be easily extended to the CCC-GARCH process, although this case is not explicitly treated here.

implications of Proposition 3 are straightforward. If either $\delta_1 > 0$ or $\delta_2 > 0$, then the terms depending on $\eta_{kh}^{(h)}$ on the right hand side of (10) and/or (11) are of order $o(h^{1/2})$. Consequently, the conditional second moments scaled by h^{-1} converge to zero as $h \to 0$. The resulting diffusion limits are degenerate and are characterized by time varying but deterministic variances of the marginal processes and/or a deterministic correlation driving process. We have the following results.

Theorem 3 (**Alternative convergence conditions**). *Under (15), (16), (18) and (34)-(35), the discrete time cDCC-GARCH process (9)-(11) admits a degenerate diffusion limit. The diffusion process $X_t = [Y_t' \ V_t' \ q_t']'$ is the solution to a system of stochastic differential equations of the form (20), with drift given by (21) and diffusion matrix given respectively by*

i) (deterministic variances but stochastic correlation) under (15), (16), (18), (19) and (34)

$$a(X_t, t) = \begin{bmatrix} S_t R_t S_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \vartheta^2 [D_N^+ (P_t \otimes P_t) D_N^{+\prime} K_t \\ & & \quad D_N^+ (P_t \otimes P_t) D_N^{+\prime} - q_t q_t'] \end{bmatrix}. \tag{36}$$

The diffusion process defined by (20), (21) and (36) is driven by $N(N+3)/2$ independent standard Brownian motions;

ii) (stochastic variance but deterministic correlation) under (15), (16), (17), (18) and (35)

$$a(X_t, t) = \begin{bmatrix} S_t R_t S_t & 0 & 0 \\ 0 & 2AS_t^2 (R_t \odot R_r) S_t^2 A & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{37}$$

The diffusion process defined by (20), (21) and (37) is driven by $2N$ independent standard Brownian motions;

iii) (deterministic variances and correlation) under (15), (16), (18) and both (34) and (35)

$$a(X_t, t) = \begin{bmatrix} S_t R_t S_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{38}$$

The diffusion process defined by (20), (21) and (38) is driven by N independent standard Brownian motions.

It is possible to characterize the types of processes that can be obtained as Euler approximation of the different diffusions in Theorem 3. These approximations are not unique. For example, in the unvariate GARCH case, Corradi (2000) has shown that an Euler approximation of a degenerate diffusion process is GARCH, while that of a non-degenerate diffusion is stochastic volatility. In the same spirit, and using stochastic calculus results of Steele (2001) p.123, we can show that the following type of processes are Euler approximations of the three diffusions defined in Theorem 3: *i*) a process with stochastic correlation and GARCH variances, *ii*) a process with stochastic volatility and cDCC correlation, and *iii*) a cDCC-GARCH process as in (9)-(11), respectively.

4 Monte Carlo evidence on estimation by approximation

In this section we investigate the performance of the quasi approximate maximum likelihood (QAML) procedure of Fornari and Mele (2006), discussed in the introduction, in our model framework using a Monte Carlo simulation study. Estimation by QAML essentially involves two types of biases: First, the *approximation bias* arising from the approximation of an exact, but unknown, discrete time representation of the underlying diffusion process. And second, the *finite sample bias* due to the availability of a sample of only a finite number of observations. For the drift parameter of a Cox-Ingersoll-Ross type process, Phillips and Yu (2009) have shown that the approximation bias for alternative approximation schemes is typically negligible compared to the finite sample bias. This motivates the use of QAML for GARCH-type processes, where the exact discrete time model is unknown. Rather than comparing with alternative estimation strategies, e.g. simulated MLE as in Kleppe, Yu, and Skaug (2010), we focus on the properties of the simple QAML procedure and, in particular, the relative importance of approximation and estimation bias.

For univariate GARCH models, Wang (2002) has shown the non-equivalence of the statistical experiments resulting from the estimation of the discrete time model and its weak diffusion limit. Nevertheless, many studies have used QAML, see e.g. Engle and Lee (1996), Broze, Scaillet, and Zakoian (1998), Lewis (2000), Barone-Adesi, Rasmussen, and Ravanelli (2005) and Stentoft (2011), arguing that the approximation bias tends to disappear as the frequency increases. For the related case of estimating temporally aggregated multivariate GARCH models, the bias of QAML has been shown to be negligible, see Hafner and Rombouts (2007). Therefore, it is of interest to see whether this finding extends to the estimation of some or all parameters of the cDCC-GARCH diffusion limit.

We estimate the parameters of a sequence, indexed by h, of discrete time cDCC-GARCH models with i.i.d. innovations. Then, for each h, we use the relationships given in Proposition 1 to obtain the diffusion parameters and we investigate the behavior of the latter as $h \to 0$. To keep the computational burden feasible, we focus on the bivariate case, $N = 2$, but our results should generalize in an obvious way to higher dimensions. Using the representations of Section 3.1, the cDCC-GARCH diffusion limit can be written as

$$\begin{bmatrix} \mathrm{d}Y_{1t} \\ \mathrm{d}Y_{2t} \end{bmatrix} = \begin{bmatrix} \sqrt{V_{1t}} & 0 \\ 0 & \sqrt{V_{2t}} \end{bmatrix} \Upsilon^{(1)}(\rho_t)^{\frac{1}{2}} \mathrm{d}W_t^{(1)} \tag{39}$$

$$\begin{bmatrix} \mathrm{d}V_{1t} \\ \mathrm{d}V_{2t} \\ \mathrm{d}Q_{12t} \end{bmatrix} = \begin{bmatrix} c_1 - \Lambda_{11}V_{1t} \\ c_2 - \Lambda_{22}V_{2t} \\ \phi(\bar{Q}_{12} - Q_{12t}) \end{bmatrix} \mathrm{d}t + \sqrt{2} \begin{bmatrix} A_{11}V_{1t} & 0 & 0 \\ 0 & A_{22}V_{2t} & 0 \\ 0 & 0 & \vartheta Q_{12t}\sqrt{\frac{1+\rho_t^2}{2\rho_t^2}} \end{bmatrix} \Upsilon^{(2)}(\rho_t)^{\frac{1}{2}} \mathrm{d}W_t^{(2)} \tag{40}$$

$$\begin{bmatrix} \mathrm{d}Q_{11t} \\ \mathrm{d}Q_{22t} \\ \mathrm{d}Q_{12t} \end{bmatrix} = \begin{bmatrix} \phi(\bar{Q}_{11} - Q_{11t}) \\ \phi(\bar{Q}_{22} - Q_{22t}) \\ \phi(\bar{Q}_{12} - Q_{12t}) \end{bmatrix} \mathrm{d}t + \sqrt{2}\vartheta \begin{bmatrix} Q_{11t} & 0 & 0 \\ 0 & Q_{22t} & 0 \\ 0 & 0 & Q_{12t}\sqrt{\frac{1+\rho_t^2}{2\rho_t^2}} \end{bmatrix} \Upsilon^{(2)}(\rho_t)^{\frac{1}{2}} \mathrm{d}W_t^{(2)}, \tag{41}$$

where

$$\Upsilon^{(1)}(\rho_t) = \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix}, \qquad \Upsilon^{(2)}(\rho_t) = \begin{bmatrix} 1 & \rho_t^2 & \sqrt{\frac{2\rho_t^2}{1+\rho_t^2}} \\ \rho_t^2 & 1 & \sqrt{\frac{2\rho_t^2}{1+\rho_t^2}} \\ \sqrt{\frac{2\rho_t^2}{1+\rho_t^2}} & \sqrt{\frac{2\rho_t^2}{1+\rho_t^2}} & 1 \end{bmatrix}$$

and

$$\rho_t = \frac{Q_{12t}}{\sqrt{Q_{11t}Q_{22t}}}.$$

Note that the drift in $[\mathrm{d}V_{1t}, \mathrm{d}V_{2t}]$ can be expressed in terms of a steady state drift parameter $\bar{c}$ with components $\bar{c}_i = c_i/\Lambda_{ii}$, i.e., $c_i - \Lambda_{ii} V_{it} = \Lambda_{ii}(\bar{c}_i - V_{it})$. Given the convergence rates of Proposition 1, the estimator of $\bar{c}$ should be less affected by the choice of the frequency, so that we will report estimation results also for this additional parameter.

The two partial systems $[\mathrm{d}V_{1t}, \mathrm{d}V_{2t}, \, \mathrm{d}Q_{12t}]$ and $[\mathrm{d}Q_{11t}, \mathrm{d}Q_{22t}, \mathrm{d}Q_{12t}]$ share the same correlation structure, $\Upsilon(\rho_t)$, and stochastic component, $\mathrm{d}W_t^{(2)}$.

We use an Euler discretization scheme of (39)-(41) and simulate 500 sample paths using a discretization interval $\Delta t = 1/640$ and length $k = 2000$ periods. The data is generated using the following parameterization: $c = [0.1, 0.15]'$, $A_{11} = 0.07$, $A_{22} = 0.10$, $\Lambda_{11} = 0.13$, $\Lambda_{22} = 0.10$, $\vartheta = 0.08$ and $\phi = 0.04$. This parameterization implies an unconditional variance of the marginal processes $\bar{c} = [1, 1.5]'$. As noted in Section 3.1, the drift of the correlation driving process has been parameterized as the product of a frequency invariant parameter $\bar{Q}$, representing the unconditional expectation of the process normalized by h, and a linear combination of frequency dependent parameters. Thus, $\bar{Q}_h = \bar{Q}\phi = \bar{Q}(1-\vartheta_h-\gamma_h)$. Furthermore, the diagonal elements of $\bar{Q}$ are fixed to one and not included in the estimation. The unconditional correlation, $\bar{Q}_{12}$, is set to 0.5. The square root of the correlation matrices of the diffusion, $\Upsilon^{(1)}(\rho_t)$ and $\Upsilon^{(2)}(\rho_t)$, are computed by spectral decomposition.

For each sample path we estimate the following model

$$\eta_{kh}^{(h)} = S_{kh}^{(h)-1}(Y_{kh}^{(h)} - Y_{(k-1)h}^{(h)}) \mid \mathcal{F}_{kh} \;\sim\; \mathrm{N}(0, hR_{kh}^{(h)}), \tag{42}$$

$$V_{(k+1)h}^{(h)} = c\,h + A\,h^{-1/2} S_{kh}^{(h)2}(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)}) + (I_N - A\sqrt{h} - \Lambda\,h)V_{kh}^{(h)}, \tag{43}$$

$$Q_{(k+1)h}^{(h)} = \bar{Q}\phi\,h + \vartheta\,h^{-1/2}(P_{kh}^{(h)}\,\eta_{kh}^{(h)}\,\eta_{kh}^{(h)\prime}\,P_{kh}^{(h)}) + (1 - \vartheta\sqrt{h} - \phi\,h)Q_{kh}^{(h)}, \tag{44}$$

that is (9)-(11) expressed as a function of frequency rescaled parameters. This is obtained by substituting (up to $o(1)$)

$$\begin{aligned}
c_h &= c\,h, && (45)\\
A_h &= A\sqrt{h}, && (46)\\
B_h &= (I_N - A\sqrt{h} - \Lambda\,h), && (47)\\
\vartheta_h &= \vartheta\sqrt{h}, && (48)\\
\gamma_h &= (1 - \vartheta\sqrt{h} - \phi\,h), && (49)
\end{aligned}$$

from Proposition 1, into (9)-(11). The model is estimated using data sampled at nine frequencies spanning from $h = 1$ to $h = 1/320$. The sequences of discrete time models are estimated by Gaussian QAML. The bias and root mean square error (RMSE) of parameter estimates are reported in Tables 1 and 2, respectively.

As the sampling frequency increases, the bias and RMSE tend to disappear at an appropriate rate for all parameters. For a given frequency, however, there are remarkable differences. For the parameters in the drift terms, bias and RMSE decrease as the sample size k increases, suggesting that the finite sample bias dominates the approximation bias, which confirms the results of Phillips and Yu (2009) for the Cox-Ingersoll-Ross diffusion process. However, this is not the case for the parameters A and ϑ linked to the innovation terms in V_t and Q_t, respectively. For these parameters, the approximation bias dominates the finite sample bias. Clearly, QAML is inconsistent when only the sample size is increased but not the frequency, which confirms the results of Wang (2002) for the univariate case. With sufficiently high frequency and sample size, however, the bias may be considered negligible in most practical situations.

Table 1: Bias of the diffusion parameters of the cDCC process

	$\bar{c}_1$=1.00		c_1=0.13		A_{11}=0.07		Λ_{11}=0.13	
$h \setminus k$	500	2000	500	2000	500	2000	500	2000
1	0.0013	-0.0012	0.1536	0.1156	-0.0425	-0.0488	0.1537	0.1159
1/4	0.0003	-0.0017	0.1065	0.0367	-0.0355	-0.0390	0.1071	0.0373
1/16	-0.0005	-0.0008	0.0298	0.0076	-0.0259	-0.0260	0.0300	0.0078
1/64	0.0000	-0.0005	0.0170	0.0067	-0.0148	-0.0146	0.0173	0.0068
1/320	-0.0001	-0.0005	0.0107	0.0026	-0.0072	-0.0071	0.0111	0.0027
	$\bar{c}_2$=1.50		c_2=0.15		A_{22}=0.10		Λ_{22}=0.10	
$h \setminus k$	500	2000	500	2000	500	2000	500	2000
1	-0.0048	0.0007	0.1489	0.0634	-0.0551	-0.0596	0.1002	0.0424
1/4	-0.0052	-0.0039	0.0662	0.0135	-0.0399	-0.0410	0.0451	0.0095
1/16	-0.0051	-0.0027	0.0245	0.0074	-0.0246	-0.0240	0.0173	0.0053
1/64	-0.0044	-0.0017	0.0198	0.0066	-0.0122	-0.0125	0.0142	0.0047
1/320	-0.0032	0.0001	0.0146	0.0033	-0.0054	-0.0058	0.0106	0.0024
	$\bar{Q}_{12}$=0.50				ϑ=0.08		ψ=0.04	
$h \setminus k$	500	2000			500	2000	500	2000
1	-0.0063	-0.0013			-0.0300	-0.0356	0.0622	0.0087
1/4	-0.0010	0.0012			-0.0211	-0.0214	0.0185	0.0043
1/16	0.0001	0.0017			-0.0118	-0.0114	0.0104	0.0033
1/64	0.0036	0.0020			-0.0054	-0.0055	0.0075	0.0031
1/320	0.0046	0.0033			-0.0022	-0.0022	0.0050	0.0023

Table 2: RMSE of the diffusion parameters of the cDCC process

	$\bar{c}_1$=1.00		c_1=0.13		A_{11}=0.07		Λ_{11}=0.13	
$h \setminus k$	500	2000	500	2000	500	2000	500	2000
1	0.0941	0.0385	0.3256	0.2765	0.0535	0.0517	0.3248	0.2768
1/4	0.0445	0.0226	0.2555	0.1246	0.0416	0.0403	0.2579	0.1264
1/16	0.0363	0.0186	0.0922	0.0309	0.0284	0.0265	0.0923	0.0311
1/64	0.0338	0.0169	0.0456	0.0202	0.0162	0.0149	0.0464	0.0204
1/320	0.0342	0.0165	0.0326	0.0131	0.0084	0.0074	0.0336	0.0134
	$\bar{c}_2$=1.50		c_2=0.15		A_{22}=0.10		Λ_{22}=0.10	
$h \setminus k$	500	2000	500	2000	500	2000	500	2000
1	0.1818	0.0712	0.3150	0.1846	0.0660	0.0620	0.2131	0.1240
1/4	0.1136	0.0573	0.1736	0.0490	0.0456	0.0422	0.1158	0.0335
1/16	0.1034	0.0515	0.0639	0.0271	0.0275	0.0248	0.0440	0.0188
1/64	0.1008	0.0498	0.0496	0.0204	0.0147	0.0131	0.0346	0.0146
1/320	0.1000	0.0496	0.0408	0.0158	0.0074	0.0063	0.0291	0.0118
	$\bar{Q}_{12}$=0.50				ϑ=0.08		ψ=0.04	
$h \setminus k$	500	2000			500	2000	500	2000
1	0.0785	0.0350			0.0420	0.0376	0.1619	0.0280
1/4	0.0578	0.0294			0.0261	0.0225	0.0419	0.0132
1/16	0.0548	0.0276			0.0148	0.0123	0.0247	0.0092
1/64	0.0537	0.0271			0.0083	0.0063	0.0204	0.0084
1/320	0.0510	0.0251			0.0045	0.0029	0.0180	0.0074

5 Conclusions

This paper considered weak diffusion limits of two conditional correlation GARCH specifications, namely the cDCC model of Aielli (2006) and the CCC model of Bollerslev (1990). For the cDCC-GARCH model, the diffusion limit is degenerate in the sense that the diffusion of the variances and that of the diagonal elements of the correlation driving process are pairwise governed by the same Brownian motion. We show that this result is due to the particular structure of the noise propagation system of the variances and of the correlation driving process. The CCC model, which can be obtained from cDCC under suitable parameter restrictions, admits a non-degenerate diffusion. Under an alternative set of conditions regarding the convergence rates of the parameters, we obtain diffusion limits characterized by a stochastic price process where either the variances, the correlations, or both, are time-varying but deterministic. Our Monte Carlo study confirms that estimation of the diffusion parameters by QAML is inconsistent for any fixed frequency, but may provide good approximations if the frequency and sample size are sufficiently large.

There are several possible extensions of this work. First of all, the assumption of Gaussian innovations may be relaxed. One may also extend the results to allow for volatility spillover. Furthermore, similar to Nelson (1990) it may be possible to derive the stationary distribution of the continuous time limit of returns, variances and correlations. Finally, it would be useful to extend the results of this paper to jump-diffusion processes, based on the results of Ethier and Kurtz (1986).

Appendix: Proofs

Proof of Proposition 1. The first step is to compute the increments of the process (9)-(11), that is

$$\begin{aligned}
Y_{kh}^{(h)} - Y_{(k-1)h}^{(h)} &= S_{kh}^{(h)}\eta_{kh}^{(h)} \\
V_{(k+1)h}^{(h)} - V_{kh}^{(h)} &= c_h + A_h S_{kh}^{(h)2} h^{-1}(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)}) + (B_h - I_N)V_{kh}^{(h)} \\
q_{(k+1)h}^{(h)} - q_{kh}^{(h)} &= (1-\vartheta_h-\gamma_h)\bar{q} + \vartheta_h h^{-1}\text{vech}(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)}) + (\gamma_h - 1)q_{kh}^{(h)},
\end{aligned}$$

where we have used that $\bar{q}_h = \bar{q}(1-\vartheta_h-\gamma_h)$.

Second, we compute the moments (conditioned on $\mathcal{F}_{kh} = \{kh,\ Y_0^{(h)}, ..., Y_{(k-1)h}^{(h)}, V_0^{(h)}, ..., V_{kh}^{(h)}, q_0^{(h)}, ..., q_{kh}^{(h)}\}$) to define suitable convergence conditions as required by Assumption 1. To simplify the notation, let us define the difference operator over an interval of size h as $\Delta : \Delta X_{kh}^{(h)} = X_{kh}^{(h)} - X_{(k-1)h}^{(h)}$. The first conditional moment per unit of time of the increments of (9)-(11) is given by

$$\begin{aligned}
h^{-1}\text{E}[\Delta\, Y_{kh}^{(h)}|\mathcal{F}_{kh}] &= S_{kh}^{(h)}\ \text{E}[\,\eta_{kh}^{(h)}] = 0 &(50)\\
h^{-1}\text{E}[\Delta\, V_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] &= h^{-1}c_h + A_h h^{-2} S_{kh}^{(h)2}\text{E}[\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)}|\mathcal{F}_{kh}] + h^{-1}(B_h - I_N)V_{kh}^{(h)} \\
&= h^{-1}c_h + h^{-1}(A_h + B_h - I_N)V_{kh}^{(h)} &(51)\\
h^{-1}\text{E}[\Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] &= h^{-1}(1-\vartheta_h-\gamma_h)\bar{q} + h^{-2}\vartheta_h\text{vech}(\text{E}[P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)}|\mathcal{F}_{kh}]) \\
&\quad + h^{-1}(\gamma_h - 1)q_{kh}^{(h)} \\
&= h^{-1}(1-\vartheta_h-\gamma_h)\bar{q} + h^{-1}(\vartheta_h + \gamma_h - 1)q_{kh}^{(h)}, &(52)
\end{aligned}$$

where $\text{E}[\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)}|\mathcal{F}_{kh}] = h\,1_N$ and $\text{vech}(P_{kh}^{(h)}\text{E}[\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}|\mathcal{F}_{kh}]P_{kh}^{(h)}) = h\text{vech}(P_{kh}^{(h)}R_{kh}^{(h)}P_{kh}^{(h)}) = hq_{kh}^{(h)}$.

To compute the second moments per unit of time, consider the following partition

$$\text{vech}\left(\text{Var}([\Delta Y_{kh}^{(h)\prime}, \Delta V_{(k+1)h}^{(h)\prime}, \Delta q_{(k+1)h}^{(h)\prime}]'|\mathcal{F}_{kh})\right) = \begin{bmatrix} \text{Var}(\Delta Y_{kh}^{(h)}|\mathcal{F}_{kh}) \\ \text{Cov}(\Delta Y_{kh}^{(h)}, \Delta V_{(k+1)h}^{(h)}|\mathcal{F}_{kh})' \\ \text{Cov}(\Delta Y_{kh}^{(h)}, \Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh})' \\ \text{Var}(\Delta V_{(k+1)h}^{(h)}|\mathcal{F}_{kh}) \\ \text{Cov}(\Delta V_{(k+1)h}^{(h)}, \Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh})' \\ \text{Var}(\Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh}) \end{bmatrix}.$$

The conditional variance of $\Delta Y_{kh}^{(h)}$ standardized by h is given by

$$h^{-1}\text{Var}[\Delta Y_{kh}^{(h)}|\mathcal{F}_{kh}] = h^{-1}S_{kh}^{(h)}\text{E}(\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}|\mathcal{F}_{kh})S_{kh}^{(h)} = S_{kh}^{(h)}R_{kh}^{(h)}S_{kh}^{(h)}. \qquad (53)$$

Similarly the conditional variance of $\Delta V_{(k+1)h}^{(h)}$ is given by

$$h^{-1}\mathrm{Var}[\Delta V_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] = A_h S_{kh}^{(h)2} h^{-3}\Big[\mathrm{E}[(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)})(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)})'|\mathcal{F}_{kh}]$$
$$-\mathrm{E}[(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)})|\mathcal{F}_{kh}]\mathrm{E}[(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)})|\mathcal{F}_{kh}]'\Big] S_{kh}^{(h)2} A_h'. \quad (54)$$
$$= 2h^{-1} A_h S_{kh}^{(h)2}(R_{kh}^{(h)} \odot R_{kh}^{(h)}) S_{kh}^{(h)2} A_h' \quad (55)$$

where the second equality uses that, under the conditional normality assumption, $\mathrm{E}[\eta_{kh,i}^{(h)}\eta_{kh,j}^{(h)}|\mathcal{F}_{kh}] = h^2(1 + 2R_{kh,ij}^{(h)2})$, $i, j \in \{1, ..., N\}$. Moreover, $\mathrm{E}[(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)})|\mathcal{F}_{kh}] = h1_N$.

The variance of $\Delta q_{(k+1)h}^{(h)}$ is given by

$$h^{-1}\mathrm{Var}[\Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] = \vartheta_h^2 h^{-3}\mathrm{E}[\mathrm{vech}(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)})\mathrm{vech}(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)})'|\mathcal{F}_{kh}]$$
$$- \vartheta_h^2 h^{-3}\mathrm{E}[\mathrm{vech}(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)})|\mathcal{F}_{kh}]\mathrm{E}[\mathrm{vech}(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)})|\mathcal{F}_{kh}]' \quad (56)$$
$$= h^{-1}\vartheta_h^2[D_N^+(P_{kh}^{(h)} \otimes P_{kh}^{(h)})D_N^{+\prime}K_{kh}^{(h)}D_N^+(P_{kh}^{(h)} \otimes P_{kh}^{(h)})D_N^{+\prime} - q_{kh}^{(h)}q_{kh}^{(h)\prime}]. \quad (57)$$

where the second equality uses that $\mathrm{E}[\mathrm{vech}(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)})|\mathcal{F}_{kh}] = hq_{kh}^{(h)}$, and where $K_{kh}^{(h)} = h^{-2}\mathrm{E}[\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime} \otimes \eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}|\mathcal{F}_{kh}]$ is the $(N(N+1)/2 \times N(N+1)/2)$ matrix of conditional fourth moments of $\eta_{kh}^{(h)}$ which, given the normality assumption of the innovations, is given by

$$K_{kh}^{(h)} = 2D_N^+(R_{kh}^{(h)} \otimes R_{kh}^{(h)})D_N^{+\prime} + \mathrm{vech}(R_{kh}^{(h)})\mathrm{vech}(R_{kh}^{(h)})',$$

which is a consequence of Theorem 10.2 of Magnus (1988), see the proof of Theorem 1 of Hafner (2003).

Finally, the conditional covariances are

$$h^{-1}\mathrm{Cov}[\Delta Y_{kh}^{(h)}, \Delta V_{(k+1)h}^{(h)})|\mathcal{F}_{kh}] = h^{-1}\mathrm{E}[(S_{kh}^{(h)}\eta_{kh}^{(h)})(A_h S_{kh}^{(h)2}h^{-1}(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)}))'|\mathcal{F}_{kh}]$$
$$= h^{-2}S_{kh}^{(h)}\mathrm{E}[\eta_{kh}^{(h)}(\eta_{kh}^{(h)} \odot \eta_{kh}^{(h)})'|\mathcal{F}_{kh}]S_{kh}^{(h)2}A_h = 0, \quad (58)$$

because all conditional third moments of $\eta_{kh}^{(h)}$ are equal to zero given the normality assumption. Furthermore, we have

$$h^{-1}\mathrm{Cov}[\Delta Y_{kh}^{(h)}, \Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] = h^{-2}\mathrm{E}[(S_{kh}^{(h)}\eta_{kh}^{(h)})(\vartheta_h \mathrm{vech}(P_{kh}^{(h)}\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime}P_{kh}^{(h)}))'|\mathcal{F}_{kh}]$$
$$= h^{-2}\vartheta_h S_{kh}^{(h)}\mathrm{E}[\eta_{kh}^{(h)}\,\mathrm{vech}(\eta_{kh}^{(h)}\eta_{kh}^{(h)\prime})'|\mathcal{F}_{kh}](D_N^+(P_{kh}^{(h)} \otimes P_{kh}^{(h)})D_N^{+\prime}) = 0 \quad (59)$$

and

$$
\begin{aligned}
&h^{-1}\mathrm{Cov}[\Delta V^{(h)}_{(k+1)h}, \Delta q^{(h)}_{(k+1)h}|\mathcal{F}_{kh}] = \\
&\quad = h^{-3}\mathrm{E}\left[\left(A_h S^{(h)2}_{kh}(\eta^{(h)}_{kh}\odot\eta^{(h)}_{kh})\right)\left(\vartheta_h \mathrm{vech}(P^{(h)}_{kh}\eta^{(h)}_{kh}\eta^{(h)\prime}_{kh}P^{(h)}_{kh})\right)'|\mathcal{F}_{kh}\right] \\
&\quad - h^{-3}\mathrm{E}\left[A_h S^{(h)2}_{kh}(\eta^{(h)}_{kh}\odot\eta^{(h)}_{kh})|\mathcal{F}_{kh}\right]\mathrm{E}\left[\vartheta_h \mathrm{vech}(P^{(h)}_{kh}\eta^{(h)}_{kh}\eta^{(h)\prime}_{kh}P^{(h)}_{kh})|\mathcal{F}_{kh}\right]' \\
&\quad = h^{-1}\vartheta_h A_h S^{(h)2}_{kh}\left[I^* K^{(h)}_{kh} D^{+}_{N}(P^{(h)}_{kh}\otimes P^{(h)}_{kh})D^{+\prime}_{N} - 1_N q^{(h)\prime}_{kh}\right]. \qquad (60)
\end{aligned}
$$

where the second equality uses $\eta^{(h)}_{kh}\odot\eta^{(h)}_{kh} = \mathrm{diag}(\eta^{(h)}_{kh}\eta^{(h)\prime}_{kh}) = I^*\mathrm{vech}(\eta^{(h)}_{kh}\eta^{(h)\prime}_{kh})$.

For the conditional moments (50)-(52) (drift) and (53), (55), and (57)-(60) (second moments) to converge to well behaved functions as $h\to 0$, as required by Assumption 1 *a*) and *b*), the following limits must exist and be finite

$$
\begin{aligned}
&\lim_{h\to 0} h^{-1}c_h = c && (61)\\
&\lim_{h\to 0} h^{-1}(A_h + B_h - I_N) = -\Lambda && (62)\\
&\lim_{h\to 0} h^{-1/2}A_h = A && (63)\\
&\lim_{h\to 0} h^{-1}(1-\vartheta_h-\gamma_h) = \phi && (64)\\
&\lim_{h\to 0} h^{-1/2}\vartheta_h = \vartheta, && (65)
\end{aligned}
$$

where c is a $(N\times 1)$ vector, A and Λ are $(N\times N)$ diagonal matrices and ϕ and ϑ are scalars with all elements positive and finite, such that $c>0$ (elementwise) ensures positivity of the variance process, $A>0$ and $\vartheta>0$ ensure that the rescaled second conditional moment of $V^{(h)}_{kh}$ and $q^{(h)}_{kh}$ does not vanish as $h\to 0$, while $\Lambda>0$ and $\phi>0$ ensure covariance stationarity of the return process. Finally, by straightforward computation as in Nelson (1990), under (61)-(65), Assumption 1 *c*) holds for $\delta=2$, i.e.,

$$
\begin{aligned}
h^{-1}\lim_{h\to 0} E\left[\left|(\Delta Y^{(h)}_{kh})_i\right|^4|\mathcal{F}_{kh}\right] &= 0, \forall i, i=1,...,N\\
h^{-1}\lim_{h\to 0} E\left[\left|(\Delta V^{(h)}_{(k+1)h})_i\right|^4|\mathcal{F}_{kh}\right] &= 0, \forall i, i=1,...,N\\
h^{-1}\lim_{h\to 0} E\left[\left|(\Delta q^{(h)}_{(k+1)h})_i\right|^4|\mathcal{F}_{kh}\right] &= 0, \forall i, i=1,...,N(N+1)/2
\end{aligned}
$$

which completes the proof. ■

Proof of Theorem 1.

The process (9)-(14) is Markovian with drift and second moment per unit of time given by (50)-(52) (drift) and (53), (55), and (57)-(60) (second moments), respectively. The theorem of weak convergence applies if Assumptions 1 to 4 hold. Proposition 1 provides suitable convergence conditions for Assumption 1 to hold, and the drift and diffusion matrix for the system of stochastic differential equations are defined.

Substituting (61)-(65) into (50)-(52) (first moments) and (53), (55), and (57)-(60) (second moments), we obtain

$$
\begin{aligned}
h^{-1}\mathrm{E}[\Delta Y_{kh}^{(h)}|\mathcal{F}_{kh}] &= 0 \\
h^{-1}\mathrm{E}[\Delta V_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] &= c - \Lambda V_{kh}^{(h)} + o(1) \\
h^{-1}\mathrm{E}[\Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] &= \phi(\bar{q} + q_{kh}^{(h)}) + o(1)
\end{aligned}
$$

for the drift, while for the second moment

$$
\begin{aligned}
& h^{-1}\mathrm{Var}[\Delta Y_{kh}^{(h)}|\mathcal{F}_{kh}] = S_{kh}^{(h)} R_{kh}^{(h)} S_{kh}^{(h)} \\
& h^{-1}\mathrm{Var}[\Delta V_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] = 2A_h S_{kh}^{(h)2}(R_{kh}^{(h)} \odot R_{kh}^{(h)}) S_{kh}^{(h)2} A_h + o(1) \\
& h^{-1}\mathrm{Var}[\Delta q_{(k+1)h}^{(h)}|\mathcal{F}_{kh}] = \vartheta^2[(D_N^+(P_{kh}^{(h)} \otimes P_{kh}^{(h)})D_N^{+\prime})K_{kh}^{(h)} \\
& \qquad (D_N^+(P_{kh}^{(h)} \otimes P_{kh}^{(h)})D_N^{+\prime}) - q_{kh}^{(h)} q_{kh}^{(h)\prime}] + o(1) \\
& h^{-1}\mathrm{Cov}[\Delta Y_{kh}^{(h)}, \Delta V_{(k+1)h}^{(h)}]|\mathcal{F}_{kh}) = 0 \\
& h^{-1}\mathrm{Cov}[\Delta Y_{kh}^{(h)}, \Delta q_{(k+1)h}^{(h)}]|\mathcal{F}_{kh}) = 0 \\
& h^{-1}\mathrm{Cov}[\Delta V_{(k+1)h}^{(h)}, \Delta q_{(k+1)h}^{(h)}]|\mathcal{F}_{kh}) = \vartheta A S_{kh}^{(h)2}[I^* K_{kh}^{(h)} \times \\
& \qquad D_N^+(P_{kh}^{(h)} \otimes P_{kh}^{(h)})D_N^{+\prime} - 1_N q_{kh}^{(h)\prime}] + o(1).
\end{aligned}
$$

Hence, as $h \to 0$, the functions (21) and (22) are solutions of (1) and (2) and represent the drift and the diffusion matrix of the diffusion process $X_t = [Y_t', V_t', q_t']'$. From the representation (24), we have that the columns $N+1, \ldots, 2N$ and $2N+1, \ldots, 3N$ are collinear. Thus, the diffusion matrix is singular with $\mathrm{rank}(a([Y_t', V_t', q_t']')) = N(N+3)/2 < \dim([a(Y_t', V_t', q_t']')) = N(N+5)/2$.

The scale matrix $\sigma(X_t, t)$ can be obtained by Cholesky or spectral decomposition of (22) so that Assumption 2 holds. We assume that the probability law of initial values in (14) satisfies Assumption 3 and that for each $h \geq 0$, $\nu_h\left([Y_0', V_0', q_0']' : V_0 > 0\right.$ (elementwise) and $\left.\epsilon' Q_0 \epsilon > 0, \forall \epsilon \in \mathbb{R}^N \backslash \{0\}\right) = 1$. Condition 1 is satisfied given Assumption 1 c) which ensures continuity of the sample paths of the limit process

X_t with probability one. Condition 2 holds since the diffusion matrix and the inner product of the drift and the state variable X are at most of order two in X. Thus Assumption 4 holds, which completes the proof. ■

Proof of Proposition 2. The proof follows directly from the proof of Proposition 1 under the parameter restriction $\vartheta_h = \gamma_h = 0 \ \forall h$, i.e., $R_{kh}^{(h)} = (\bar{Q} \odot I_N)^{-1/2} \bar{Q} (\bar{Q} \odot I_N)^{-1/2} = R \ \forall kh, \ k \in \mathbb{N}, \ h > 0$. ■

Proof of Theorem 2. The theorem of weak convergence applies by symmetry with the unrestricted model (see Theorem 1). In particular Assumption 1 holds analogously under the given parameter constraint. Hence substituting (61)-(63) into the two sets of equations (50)-(51) and (53), (55), and (58), as $h \to 0$, we obtain the following mappings

$$b(X_t, t) = \begin{bmatrix} 0 \\ c - \Lambda V_t \end{bmatrix} \tag{66}$$

and

$$a(X_t, t) = \begin{bmatrix} S_t R S_t & 0 \\ 0 & 2 A S_t^2 (R \odot R) S_t^2 A \end{bmatrix} \tag{67}$$

which are solution of (1) and (2) and represent the drift and the diffusion matrix of the diffusion process $X_t = [Y_t', V_t']'$. Under the assumption that the initial values satisfy, for each $h \geq 0$, $\nu_h([Y_0', V_0']' : V_0 > 0 \text{ (elementwise)}) = 1$ and $\epsilon' R \epsilon > 0, \forall \epsilon \in \mathbb{R}^N \backslash \{0\}$, the diffusion matrix in (67) is non-degenerate with $\text{rank}(a(X_t, t)) = N^2$. ■

Proof of Proposition 3. Assumption 1a) holds trivially. To show Assumption 1b), consider the limit, as $h \to 0$, of the moments of interests (55), (57) and (60). The case $\delta_1 = \delta_2 = 0$ is covered by Proposition 1. If $\delta_1 > 0$, then $\lim_{h \to 0} h^{-1} \text{Var}[\Delta V_{(k+1)h}^{(h)} | \mathcal{F}_{kh}] = 0$ provided that

$$\lim_{h \to 0} h^{-1/2} A_h = 0 \tag{68}$$

that is A_h is of order $h^{1/2+\delta_1}$, $\delta_1 > 0$.

Similarly, if $\delta_2 > 0$, then $\lim_{h \to 0} h^{-1} \text{Var}[\Delta q_{(k+1)h}^{(h)} | \mathcal{F}_{kh}] = 0$ provided that

$$\lim_{h \to 0} h^{-1/2} \vartheta_h = 0 \tag{69}$$

that is, ϑ_h is of order $h^{1/2+\delta_2}$, $\delta_2 > 0$.

Either $\delta_1 > 0$ or $\delta_2 > 0$, or both, also ensure that

$$\lim_{h \to 0} h^{-1} \mathrm{Cov}[\Delta V^{(h)}_{(k+1)h}, \Delta q^{(h)}_{(k+1)h} | \mathcal{F}_{kh}] = 0.$$

Hence, under (68) and (69) Assumption 1b) holds. Finally, Assumption 1c) can be shown similar to the case $\delta_1 = \delta_2 = 0$ (Proposition 1). ■

Proof of Theorem 3. Under Proposition 3, the theorem of weak convergence applies by analogy to Theorem 1. Furthermore, depending on the combination of convergence conditions for A_h and ϑ_h, we either obtain a diffusion with deterministic variances and stochastic correlations (i.e. $\delta_1 > 0$ and $\delta_2 = 0$), or stochastic variances and deterministic correlations (i.e. $\delta_1 = 0$ and $\delta_2 > 0$), or deterministic variances and correlations (i.e. $\delta_1 > 0$ and $\delta_2 > 0$). The drift and diffusion matrices can be derived from those of Proposition 1. ■

References

AIELLI, G. P. (2006): "Consistent Estimation of Large Scale Dynamic Conditional Correlations," Department of Statistics, University of Florence.

——— (2013): "Dynamic Conditional Correlations: On Properties and Estimation," *Journal of Business and Economic Statistics*, 31, 282–299.

ALEXANDER, C., AND E. LAZAR (2005): "On the Continuous Limit of GARCH," Discussion Papers in Finance 2005-13.

BARONE-ADESI, G., H. RASMUSSEN, AND C. RAVANELLI (2005): "An Option Pricing Formula for the GARCH Diffusion Model," *Computational Statistics and Data Analysis*, 49(2), 287–310.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

——— (1990): "Modeling the Coherence on Short-Run Nominal Exchange Rates: a Multivariate Generalized ARCH Model," *Review of Economics and Statistics*, 72, 498–505.

BOLLERSLEV, T., R. ENGLE, AND D. NELSON (1994): *Handbook of Econometrics*vol. IV, chap. Arch Models. North-Holland.

BROZE, L., O. SCAILLET, AND J. ZAKOIAN (1998): "Quasi Indirect Inference for Diffusion Processes," *Econometric Theory*, 14, 161–186.

CORRADI, V. (2000): "Reconsidering the Continuous Time Limit of the GARCH(1,1) Process," *Journal of Econometrics*, 96, 145–153.

DING, Z., C. GRANGER, AND R. ENGLE (1993): "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1, 83–106.

DROST, C., AND T. NIJMAN (1993): "Temporal Aggregation of GARCH Processes," *Econometrica*, 61, 909–927.

DROST, F., AND J. WERKER (1996): "Closing the Gap: Continuous Time GARCH Modeling," *Journal of Econometrics*, 74, 31–57.

DUAN, J. (1997): "Augmented GARCH(p,q) Process and its Diffusion Limit," *Journal of Econometrics*, 79, 97–127.

DUFFIE, D., AND K. SINGLETON (1993): "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61, 929–952.

ENGLE, R. (2002): "Dynamic Conditional Correlation - A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business and Economic Statistics*, 20, 339–350.

ENGLE, R., AND G. LEE (1996): "Estimating Diffusion Models of Stochastic Volatility," in *Modeling Stock Market Volatility: Bridging the Gap to Continuous Time*, ed. by P. Rossi, pp. 333–384. Academic Press.

ETHIER, S., AND T. KURTZ (1986): *Markov Processes: Characterization and Convergence*. John Wiley and Sons.

FORNARI, F., AND A. MELE (1997): "Weak Convergence and Distributional Assumptions for a General Class of Nonlinear ARCH Models," *Econometric Reviews*, 16, 205–229.

——— (2006): "Approximating Volatility Diffusions with CEV-ARCH Models," *Journal of Economic Dynamics & Control*, 30(6), 931–966.

GHYSELS, E., A. HARVEY, AND E. RENAULT (1996): *Maddala, G. (ed.): Handbook of Statistics* vol. 14, chap. Stochastic volatility, pp. 119–191. North-Holland.

HAFNER, C. (2003): "Fourth moment structure of multivariate GARCH models," *Journal of Financial Econometrics*, 1, 26–54.

——— (2008): "Temporal Aggregation of Multivariate GARCH Processes," *Journal of Econometrics*, 142, 467–483.

HAFNER, C., AND J. ROMBOUTS (2007): "Estimation of Temporally Aggregated Multivariate GARCH Models," *Journal of Statistical Computation and Simulation*, 77, 629–650.

HOBSON, D., AND L. ROGERS (1998): "Complete Models with Stochastic Volatility," *Mathematical Finance*, 8, 27–48.

Jeantheau, T. (2004): "A Link Between Complete Models with Stochastic Volatility and ARCH Models," *Finance and Stochastics*, 8, 111–131.

Jones, C. (2003): "The dynamics of stochastic volatility: evidence from underlying and options markets," *Journal of Econometrics*, 116, 181–224.

Kleppe, T., J. Yu, and H. Skaug (2010): "Estimating the GARCH diffusion: Simulated maximum likelihood in continuous time," SMU Research Collection, School of Economics, DP 1232.

Kushner, H. (1984): *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory.* The MIT Press.

Lewis, A. (2000): *Option Valuation under Stochastic Volatility.* Finance Press.

Lütkepohl, H. (1996): *Handbook of Matrices.* John Wiley and Sons.

Magnus, J. (1988): *Linear Structures.* London: Griffin.

Nelson, D. (1990): "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7–38.

——— (1994): "Asymptotic Filtering Theory for Multivariate ARCH Models," NBER Technical Working Papers.

Nelson, D., and D. Foster (1994): "Asymptotic Filtering Theory for Univariate ARCH Models," *Econometrica*, 62, 1–41.

Phillips, P., and J. Yu (2009): *T.G. Anderson et al. (ed.): Handbook of Financial Time Series* chap. Maximum likelihood and Gaussian estimation of continuous time models in finance, pp. 497–530. Springer.

Steele, J. (2001): *Stochastic Calculus and Financial Applications.* Springer.

Stentoft, L. (2011): "American option pricing with discrete and continuous time models: An empirical comparison," *Journal of Empirical Finance*, 18, 880–902.

Stroock, D., and S. Varadhan (1979): *Multidimensional Diffusion Processes.* Springer-Verlag.

Trifi, A. (2006): "Issues of Aggregation Over Time of Conditional Heteroscedastic Volatility Models: What Kind of Diffusion Do We Recover?," *Studies in Nonlinear Dynamics & Econometrics*, 10(4).

Wang, Y. (2002): "Asymptotic Nonequivalence of GARCH Models and Diffusions," *Annals of Statistics*, 30, 754–783.

Research Papers
2013

CREATES
Center for Research in Econometric Analysis of Time Series