

2000s-22

# Temporal Aggregation of Volatility Models

*Nour Meddahi, Éric Renault*

---

**Série Scientifique**  
*Scientific Series*

---



**CIRANO**  
Centre interuniversitaire de recherche  
en analyse des organisations

Montréal  
Juillet 2000

## CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

*CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.*

### Les organisations-partenaires / The Partner Organizations

- École des Hautes Études Commerciales
- École Polytechnique
- Université Concordia
- Université de Montréal
- Université du Québec à Montréal
- Université Laval
- Université McGill
- MEQ
- MRST
- Alcan Aluminium Ltée
- AXA Canada
- Banque Nationale du Canada
- Banque Royale du Canada
- Bell Québec
- Bombardier
- Bourse de Montréal
- Développement des ressources humaines Canada (DRHC)
- Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec
- Hydro-Québec
- Imasco
- Industrie Canada
- Pratt & Whitney Canada Inc.
- Raymond Chabot Grant Thornton
- Ville de Montréal

© 2000 Nour Meddahi et Éric Renault. Tous droits réservés. All rights reserved.

Reproduction partielle permise avec citation du document source, incluant la notice ©.

Short sections may be quoted without explicit permission, provided that full credit, including © notice, is given to the source.

Ce document est publié dans l'intention de rendre accessibles les résultats préliminaires de la recherche effectuée au CIRANO, afin de susciter des échanges et des suggestions. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs, et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.

*This paper presents preliminary research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.*

# Temporal Aggregation of Volatility Models\*

Nour Meddahi<sup>†</sup>, Éric Renault<sup>‡</sup>

## Résumé / Abstract

Dans cet article, nous considérons l'agrégation temporelle des modèles de volatilité. Nous introduisons une classe de modèles de volatilité semi-paramétrique dénommée SR-SARV et caractérisée par une variance stochastique ayant une dynamique autorégressive. Notre classe contient les modèles GARCH usuels ainsi que plusieurs variantes asymétriques. De plus, nos modèles à volatilité stochastique sont caractérisés par des moments conditionnels observables et à plusieurs horizons. La classe des modèles SR-SARV est une généralisation naturelle des modèles GARCH faibles. Notre extension présente quatre avantages : i) nous ne supposons pas que le moment d'ordre quatre est fini; ii) nous permettons des asymétries (de type skewness et effet de levier) qui sont exclues par les modèles GARCH faibles; iii) nous dérivons des restrictions sur des moments conditionnels utiles pour l'inférence non-linéaire; iv) notre cadre de travail nous permet d'étudier l'agrégation temporelle des modèles IGARCH ainsi que des modèles non linéaires comme le modèle EGARCH et les modèles exponentiels à volatilité stochastique en temps discret et continu.

*In this paper, we consider temporal aggregation of volatility models. We introduce a semiparametric class of volatility models termed square-root stochastic autoregressive volatility (SR-SARV) and characterized by an autoregressive dynamic of the stochastic variance. Our class encompasses the usual GARCH models and various asymmetric GARCH models. Moreover, our stochastic volatility models are characterized by observable multiperiod conditional moment restrictions. The SR-SARV class is a natural extension of the weak GARCH models. Our extension has four advantages: i) we do not assume*

---

\* Corresponding Author: Nour Meddahi, CIRANO, 2020 University Street, 25<sup>th</sup> floor, Montréal, Qc, Canada H3A 2A5 Tel.: (514) 985-4026 Fax: (514) 985-4039 email: meddahin@cirano.umontreal.ca

This is a revision of a part of Meddahi and Renault (1996), "Aggregations and Marginalization of GARCH and Stochastic Volatility Models". Some other results of this manuscript are now included in two companion papers, Meddahi and Renault (2000a) and Meddahi and Renault (2000b) entitled "Temporal and Cross-Sectional Aggregations of Volatility in Mean Models" and "Conditioning Information in Volatility Models" respectively. The authors thank Torben Andersen, Bryan Campbell, Marine Carrasco, Ramdam Dridi, Feike Drost, Jean-Marie Dufour, Ola Elarian, Rob Engle, Jean-Pierre Florens, René Garcia, Ramazan Gençay, Christian Gouriéroux, Stéphane Gregoir, Joanna Jasiak, Tom McCurdy, Theo Nijman, Enrique Sentana, Neil Shepard, Bas Werker, Jean-Michel Zakoian, two referees and a co-editor, and the participants of the Econometric Society meetings at Istanbul (1996) and Pasadena (1997), the Fourth Workshop of Financial Modeling and Econometric Analysis, Tilburg, December 1996, for their helpful comments. They also acknowledge fruitful discussions during seminars at CEMFI, CORE, CREST, North Carolina (Triangle seminar), Montréal, Oxford. The authors are solely responsible for any remaining errors. The first author acknowledges FCAR and MITACS for financial support.

<sup>†</sup> Université de Montréal and CIRANO

<sup>‡</sup> Université de Montréal and CIRANO

*that the fourth moment is finite; ii) we allow for asymmetries (skewness, leverage effect) that are excluded by the weak GARCH models; iii) we derive conditional moment restrictions which are useful for non-linear inference; iv) our framework allows us to study temporal aggregation of IGARCH models and non-linear models such as EGARCH and Exponential SV in discrete and continuous time.*

**Mots Clés :** GARCH, volatilité stochastique, espace-état, SR-SARV, agrégation, rendements d'actifs, processus de diffusion, générateur infinitésimal, fonctions propres

**Keywords:** GARCH, stochastic volatility, state-space, SR-SARV, aggregation, asset returns, diffusion processes, infinitesimal generator, Eigenfunctions

# 1 Introduction

Prices of financial assets, such as stocks, bonds or currencies, are available at many frequencies from intradaily to annual. When modeling volatility of the returns of such series, issues related to the effect of temporal aggregation and the choice of the observation frequency arise naturally. Basically, two modeling strategies can be considered: the model can be specified for the observable frequency by implicitly assuming that it is the correct model for this frequency (an assumption which is testable), or the model can be set for a high frequency, say continuous time, and observable restrictions can be derived for a lower frequency. Typically, models from the ARCH<sup>1</sup> family belong to the first class, while models in Drost and Nijman (1993) and Hansen and Scheinkman (1995) stem from the second strategy.<sup>2</sup> In the latter case, we say that a class of models is closed under temporal aggregation if it keeps the same structure, with possibly different parameter values, for any frequency.

Drost and Nijman (1993) consider the temporal aggregation of volatility models. They show that the usual GARCH models of Bollerslev (1986) are not closed under temporal aggregation. The main reason is that such models imply that the squared residual process is a semi-strong ARMA (where the innovation process is a martingale difference sequence, m.d.s.) which is not closed under temporal aggregation. The ARMA literature teaches us that weak ARMA models, where the innovation process is serially uncorrelated (weak white noise), are closed under temporal aggregation. Therefore, Drost and Nijman (1993) introduce the class of weak GARCH models characterized by the weak ARMA structure of the squared innovation process and show that it is closed under temporal aggregation.

However, weak GARCH models have several limitations. First, since weak GARCH models are characterized by the weak ARMA structure of the squared innovation process, Drost and Nijman (1993) assume that the fourth moment of the innovation process is finite. This seems to be empirically violated by several financial time series, especially by high frequency data.<sup>3</sup> Second, in the weak GARCH setting, only linear projections and not conditional expectations are considered. It is an important drawback if one considers that the conditional variance provides the relevant measure of risk. It is also a limitation for statistical purposes since the QMLE setting is violated. Indeed, in a Monte Carlo study we show clearly that QMLE is not consistent for temporally aggregated GARCH models.<sup>4</sup> Finally, for temporal aggregation of flow variables (e.g., returns), Drost and Nijman (1993)

---

<sup>1</sup>ARCH models were introduced by Engle (1982) and extended by Bollerslev (1986) to GARCH. For a review of the ARCH literature, see, e.g., Bollerslev, Engle and Nelson (1994), and Diebold and Lopez (1995).

<sup>2</sup>Hansen and Scheinkman (1995) consider continuous time stochastic differential equations and derive moment restrictions for a given frequency of data. Duffie and Glynn (1997) extend this to randomly sampled observations. Nelson bridges the gap between discrete time ARCH models and continuous time models by taking an approximating, filtering or smoothing approach: Nelson (1990, 1992, 1996), Nelson and Foster (1994).

<sup>3</sup>Recently, Davis and Mikosch (1998) show that for an ARCH(1) of Engle (1982) with infinite fourth moment, the standard estimator of the correlation between  $\varepsilon_t^2$  and its lags converges to a random variable.

<sup>4</sup>This is an important difference with Drost and Nijman (1992) who report simulation results which suggest that the QMLE of temporally aggregated GARCH is consistent or has a very small bias. Our results are different from theirs because we aggregate over a much longer period and we take empirically more relevant low frequency parameters.

exclude asymmetries such as skewed innovations and leverage effects (Black, 1976, Nelson, 1991).

The aim of this paper is to propose a new class of volatility models which closed under temporal aggregation and which avoids the limitations of the weak GARCH class. We follow the main idea of Drost and Nijman (1993) by considering an ARMA structure of the squared residuals. However, our approach is based on linear state-space modeling, that is, according to financial terminology, stochastic volatility (SV) modeling.<sup>5</sup> We consider the Square-Root Stochastic Autoregressive Volatility (SR-SARV) models characterized by AR dynamics of the conditional variance process. Special ARCH-type examples of SR-SARV include ARCH of Engle (1982) and GARCH of Bollerslev (1986), the asymmetric GARCH models of Glosten, Jagannathan and Runkle (1989), Engle and Ng (1993). The SR-SARV class is a natural generalization of the weak GARCH class that avoids their limitations. In particular, even if the variance is stochastic, we derive observable conditional moment restrictions for non-linear inference. When the fourth moment of the residual process is finite, these restrictions imply that the squared residual process is an ARMA process. Besides, we prove that any symmetric SR-SARV model with finite fourth moment is a weak GARCH. Hence, weak GARCH are SV processes rather than standard GARCH and our results generalize those of Drost and Nijman (1993) and of Drost and Werker (1996). Finally, our framework allows us to study temporal aggregation of IGARCH and non-linear models such as EGARCH and exponential SV.

Several models in the literature share the property of autoregression of the variance: GARCH models, structural GARCH models of Harvey, Ruiz and Sentana (1992), SV models of Barndorff-Nielsen and Shephard (1999) and the SR-SARV models of Andersen (1994).<sup>6</sup> Our class of models is closely related to the Andersen (1994) SR-SARV, and we adopt his terminology. However, while Andersen (1994) specifies a parametric setting, we take a semiparametric approach without any assumption on the probability distributions, because distributional assumptions are not closed under temporal aggregation. In the SV setting, it is usual and indeed necessary to specify the complete probability distribution which is required, e.g. for inference or forecasting, in the presence of non-linear transformations of latent variables (see, e.g., Gouriéroux and Jasiak, 1999). However, we consider here **linear models** and so we do not require any distributional assumptions. In particular, we derive **observable** multiperiod conditional moment restrictions (Hansen, 1985) for inference purposes.

Since Akaike (1974), it is well-known that there is an equivalence between **weak** ARMA and **weak** state-space models. In particular, given an ARMA process with finite variance, we can find a state-space model, generally not unique, such that the observable restrictions are the same for both models. In a companion paper, Meddahi and Renault (2000a), we extend this result to

---

<sup>5</sup>See Ghysels, Harvey and Renault (1996) and Shephard (1996) for a review.

<sup>6</sup>Several multivariate models in factor GARCH literature also share this property: Diebold and Nerlove (1989), Engle, Ng and Rothschild (1990), King, Sentana and Wadhvani (1994).

semi-strong models. However, there is not an equivalence between semi-strong ARMA models and semi-strong state-space models. More precisely, we show that while the semi-strong ARMA model admits a particular semi-strong state-space representation, the latter amounts to some multiperiod conditional moment restrictions which are less restrictive than those implied by a semi-strong ARMA. For instance, consider the ARMA(1,1) case. We show that  $z_t$  admits a semi-strong state representation if and only if there exist  $\omega$  and  $\gamma$  such that  $E[z_t - \omega - \gamma z_{t-1} \mid z_\tau, \tau \leq t-2] = 0$ .<sup>7</sup> It turns out that these weakened multiperiod conditional moment restrictions are closed under temporal aggregation. In other words, the previous **particular state-space representation** of the semi-strong ARMA(1,1) is robust to temporal aggregation while semi-strong ARMA models are not. Multiperiod conditional moment restrictions are very useful for inference (see Hansen and Singleton, 1996). When the variance of  $z_t$  is finite, these restrictions imply that  $z_t$  is a weak and not a semi-strong ARMA: it is in between.

Starting from the SR-SARV(1) class characterized by the AR(1) dynamics of the conditional variance process, we propose several extensions. In the spirit of GARCH (p,p) modeling, we introduce the SR-SARV(p) class: the variance process is the sum of the components (marginalization) of a positive multivariate VAR(1) of size  $p$ . This class contains the usual GARCH(p,p) model. Besides, the multiperiod restrictions fulfilled by the squared process are of  $p$  lags. When the fourth moment is finite, this implies that the squared process is an ARMA(p,p). In continuous time, this leads up to consider a SV model where the variance is a marginalization of a vector of size  $p$ , that is a multi-factor model for the variance (e.g., Heston, 1993; Duffie and Kan, 1996).<sup>8</sup> Exact discretization of such models is a SR-SARV(p), hence the process of squared residuals fulfills multiperiod restrictions.

We also consider the SR-SARV( $\infty$ ) class to study temporal aggregation of SV non-linear models. We do this by plugging the non-linear models in a linear SR-SARV model by considering a natural expansion of the volatility process. For instance, in the Gaussian exponential SV model of Taylor (1986), we expand the conditional variance process on the Hermite polynomials which are AR(1) and uncorrelated processes. In other words, the conditional variance process is a linear combination of an infinite number of AR(1) processes and hence a Gaussian exponential SV is SR-SARV( $\infty$ ). In continuous time, we do this expansion on the eigenfunctions of the infinitesimal generator of the volatility diffusion process (see, Hansen, Scheinkman and Touzi, 1998). We show that the SR-SARV( $\infty$ ) class is closed under temporal aggregation.

Finally, we consider temporal aggregation of IGARCH models. In this case, we consider the ISR-SARV class where we relax the assumption of integrability of the variance process while maintaining the stationarity assumption. We show that this class is closed under temporal aggregation.

---

<sup>7</sup>This restriction is less restrictive than saying that the innovation process of  $z_t$  is a m.d.s.

<sup>8</sup>Heston (1993) considers a SV model where the volatility is a Constant Elasticity of Variance (CEV) process introduced by Cox (1975). They are characterized by a linear drift and popular in finance for their nonnegativity.

The rest of the paper is organized as follows. In section 2, we summarize some classical results of ARMA theory and present the main results of Meddahi and Renault (2000a). In particular, we stress the relationships between ARMA and state-space representation. Furthermore, the latter is used to derive multiperiod conditional moment restrictions, exact discretization of continuous time models and temporal aggregation properties. Then we introduce in section 3 the SR-SARV(p) model in discrete and continuous time. Temporal aggregation, exact discretization and multiperiod conditional moment restrictions for volatility models are then deduced from the state-space representation. In particular, we characterize the relations between SR-SARV, semi-strong GARCH, weak GARCH and ARMA representations for squared innovations. Section 4 focuses more specifically on SR-SARV(1) processes to make more specific the characterization of the subclass of GARCH(1,1) and to discuss asymmetry issues (leverage effect and skewness). Section 5 considers temporal aggregation of IGARCH models, while section 6 studies temporal aggregation of non-linear models as exponential SV in discrete and continuous time. Section 7 provides a synthesis of the various models that we consider. A Monte Carlo study is presented in section 8 and we conclude in section 9. The proofs of the results are provided in the Appendix.

## 2 State-space and multiperiod ARMA models

In this section, we revisit standard ARMA theory to enhance its main lessons for volatility modeling. We particularly focus on the state-space representation of an ARMA model, the implied conditional moments restrictions (for inference purposes) and the discretization of autoregressive continuous time models. Finally, we give several examples of positive autoregressive processes.

### 2.1 State-space representation

In the time series literature (see, e.g., Brockwell and Davis, 1990), two types of ARMA processes are generally studied. The first one considers the case of strong white noise, that is independent and identically distributed (i.i.d.) innovations with a finite variance,  $D(0, \sigma^2)$ . Note that in some cases, the existence of the variance is not necessarily assumed,  $D(0)$ . In the second type, the innovations are only assumed to be second-order stationary and serially uncorrelated (weak white noise). On the other hand, several economic models imply conditional moment restrictions (first order conditions, rational expectations...). Thus the econometric literature often focuses on an intermediate type of ARMA models based on conditional expectations: the innovation process is a m.d.s. (semi-strong white noise).

**Definition 2.1. Strong, semi-strong and weak ARMA:** *Let  $\{z_t, t \in \mathbf{Z}\}$  be a stationary integrable process such that  $P(L)z_t = \omega + Q(L)\eta_t$ , with  $P(L) = 1 - \sum_{i=1}^p a_i L^i$ ,  $Q(L) = 1 - \sum_{j=1}^q b_j L^j$ , where  $L$  is the lag operator. We assume that  $a_p \neq 0$ ,  $b_q \neq 0$  and the polynomials  $P(L)$  and  $Q(L)$  have different roots which are outside the unit circle. We say that:*

- i)  $z_t$  is a strong ARMA( $p, q$ ) if the process  $\eta_t$  is i.i.d.  $D(0)$ ;
- ii)  $z_t$  is a semi-strong ARMA( $p, q$ ) if  $\eta_t$  is a m.d.s. ( $E[\eta_t | \eta_\tau, \tau \leq t-1] = 0$ );
- iii)  $z_t$  is a weak ARMA( $p, q$ ) if  $E[\eta_t] = 0$  and  $Cov[\eta_t, \eta_{t-h}] = 0$  for  $h \geq 1$ .

Note that in both strong and semi-strong cases, we only assume integrability of the process  $z_t$ . Moreover, the strong case implies the semi-strong one which implies, when the second moment of  $z_t$  is finite, the weak one. Under normality and homoskedasticity, the three notions are equivalent. Another approach to describe time series is based on state-space modeling:

**Definition 2.2. Strong, semi-strong and weak state-space representation:** Consider an integrable process  $\{z_t, G_t, t \in \mathbf{Z}\}$  such that

$$z_t = g_{t-1} + \eta_t, \text{ with} \tag{2.1}$$

$$g_t = e' G_t, \tag{2.2}$$

$$G_t = \Omega + \Gamma G_{t-1} + V_t, \tag{2.3}$$

where  $e \in \mathbb{R}^p$  and the eigenvalues of  $\Gamma$  are assumed to be smaller than one in modulus. Define the increasing filtration  $J_t = \sigma(z_t, G_t, m_t)$  where  $m_t$  is a given process.

i) When  $\{z_t, G_t\}$  is a strictly stationary process, and the process  $(\eta_t, V_t)$  is i.i.d. with zero mean and independent of  $J_{t-1}$ , we say that  $\{z_t\}$  admits the strong state-space representation of order  $p$   $\{G_t, \eta_t\}$  w.r.t.  $J_{t-1}$ ;

ii) when the process  $\{z_t, G_t\}$  is strictly stationary and the process  $(\eta_t, V_t)$  is a m.d.s. w.r.t  $J_t$ , that is

$$E[\eta_t | J_{t-1}] = 0, \tag{2.4}$$

$$E[V_t | J_{t-1}] = 0, \tag{2.5}$$

we say that  $\{z_t\}$  admits the semi-strong state-space representation of order  $p$   $\{G_t, \eta_t\}$  w.r.t.  $J_t$ ;

iii) when  $\{z_t, G_t\}$  is a stationary second-order process and the process  $(\eta_t, V_t)$  is weak white noise, and not correlated with  $(z_\tau, G_\tau, m_\tau)$ ,  $\tau < t-1$ , we say that  $\{z_t\}$  admits the weak state-space representation of order  $p$   $\{G_t, \eta_t\}$  w.r.t.  $J_t$ .

Again, in both strong and semi-strong state-space representations, we only assume that the processes are integrable, possibly with infinite variance. In such models, the dynamics of the process  $z_t$  are defined through the process  $g_t$  which is a marginalization of the VAR(1) process  $G_t$  of size  $p$ . Therefore  $g_t$  is a weak ARMA( $p, p-1$ ) (see, e.g., Lutkepohl, 1991).<sup>9</sup> The process  $G_t$  is possibly unobservable by the economic agent or by the econometrician. For instance,  $g_t$  can represent the (rational) expectation of an economic agent of a variable  $z_{t+1}$ . In this case,  $g$  is observable by the economic agent and not by the econometrician. As we already mentioned, Akaike (1974) shows that weak ARMA models are tantamount to weak state-space models. We extend this result in Meddahi and Renault (2000a) to semi-strong models by proving Propositions 2.1, 2.2 and 2.3 below:

---

<sup>9</sup>Note however that the above definition in terms of a VAR(1) process  $G_t$  of state variables is not tantamount to a definition directly in terms of a state process  $g_t$  ARMA( $p, p-1$ ); the important difference relies on the conditioning information set.

**Proposition 2.1 State-space representation of a semi-strong ARMA**

Let  $\{z_t, t \in \mathbf{Z}\}$  be a semi-strong ARMA( $p, p$ ) with a corresponding representation  $P(L)z_t = \omega + Q(L)\eta_t$  with  $P(L) = 1 - \sum_{i=1}^p a_i L^i$  and  $Q(L) = 1 - \sum_{i=1}^p b_i L^i$ . Define the processes  $\{G_t, g_t, v_t, t \in \mathbf{Z}\}$  by

$$G_{t-1} \equiv (E[z_{t+p-1} | I_{t-1}], E[z_{t+p-2} | I_{t-1}], \dots, E[z_t | I_{t-1}])', \quad (2.6)$$

$$g_{t-1} \equiv (0, 0, \dots, 0, 1)G_{t-1}, \quad (2.7)$$

and  $v_t \equiv z_t - g_{t-1}$  where  $I_t = \sigma(z_\tau, \tau \leq t)$ . Then  $z_t$  admits the semi-strong state-space representation  $\{G_t, v_t, t \in \mathbf{Z}\}$ . Moreover,  $G_t = \Omega + \Gamma G_{t-1} + V_t$  where  $\Omega = (\omega, 0, 0, \dots, 0)'$  and

$$\Gamma = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdot & \cdots & \cdots & \cdot \\ \cdot & \cdots & \cdots & \cdot \\ 0 & \cdots & 1 & 0 \end{bmatrix}. \quad (2.8)$$

Note that this result concerns semi-strong ARMA and therefore strong ARMA. However, it is not true for weak ARMA because the weak noise property is too poor to provide conditional moment restrictions like (2.4) and (2.5). These conditional moment restrictions are the only binding restrictions w.r.t. the Wold representation setting. Of course, any ARMA( $p, q$ ) can be written as an ARMA( $r, r$ ) with  $r = \max(p, q)$ . But the above property shows that the state-space representation requires something intermediate between weak and semi-strong ARMA( $p, p$ ) properties, characterized by the following multiperiod conditional moment restrictions of order  $p$  on observable variables:

**Proposition 2.2 State-space representation and multiperiod conditional moments restrictions** A stationary process  $\{z_t, t \in \mathbf{Z}\}$  admits a semi-strong state-space representation of order  $p$  iff there exist  $(p+1)$  real numbers  $\omega, a_1, \dots, a_p$ , such that the roots of  $1 - \sum_{i=1}^p a_i L^i$  are outside the unit circle and

$$E[z_t - \omega - \sum_{i=1}^p a_i z_{t-i} | z_\tau, \tau \leq t - p - 1] = 0. \quad (2.9)$$

To summarize, the state-space representation of order  $p$  characterizes a class of processes which contains strictly the class of semi-strong and strong ARMA( $p, p$ ) and is, when the second moment is finite, strictly included in the weak ARMA( $p, p$ ) one. For a review of examples where the multiperiod restrictions (2.9) occur or of the optimal instruments issue for (2.9), see Hansen and Singleton (1996).

On the other hand, temporal aggregation properties of ARMA models are proven **only** for weak ARMA,<sup>10</sup> see e.g., Palm and Nijman (1984) and Granger (1990) for a survey. However, the class of semi-strong VAR(1) is closed under temporal aggregation for stock variables. This allows us to prove that the state-space representation of order  $p$  is closed under temporal aggregation:

---

<sup>10</sup>Under normality and homoskedasticity, temporal aggregation holds for semi-strong and strong ARMA since they are equivalent to weak ARMA.

### Proposition 2.3 Temporal aggregation of the state-space representation

Let  $\{z_t, t \in \mathbf{Z}\}$  be a stationary process which admits a semi-strong state-space representation of order  $p$   $\{G_t, \eta_t, t \in \mathbf{Z}\}$  w.r.t.  $J_t = \sigma(z_\tau, G_\tau, m_\tau, \tau \leq t)$ . Define for a given integer  $m$  and real numbers  $(a_0, a_1, \dots, a_{m-1})$  the process  $\{z_{tm}^{(m)}, t \in \mathbf{Z}\}$  by  $z_{tm}^{(m)} \equiv \sum_{i=0}^{m-1} a_i z_{tm-i}$ . Then  $\{z_{tm}^{(m)}, t \in \mathbf{Z}\}$  admits a semi-strong state-space representation of order  $p$  w.r.t.  $J_{tm}^{(m)} = \sigma(z_{\tau m}^{(m)}, G_{\tau m}, m_{\tau m}, \tau \leq t)$ . More precisely, we have  $z_{tm}^{(m)} = g_{tm-m}^{(m)} + \eta_{tm}^{(m)}$  where

$$g_{tm-m}^{(m)} \equiv E[z_{tm}^{(m)} | J_{tm-m}] = e'(A^{(m)}G_{tm-m} + B^{(m)}), \quad (2.10)$$

$$\text{with } A^{(m)} = \sum_{i=0}^{m-1} a_i \Gamma^{m-i-1} \text{ and } B^{(m)} = \left( \sum_{i=0}^{m-1} a_i \left( \sum_{k=0}^{m-i-2} \Gamma^k \right) \right) \Omega. \quad (2.11)$$

Assume that  $e'A^{(m)} \neq 0$ , i.e.  $z_{tm}^{(m)}$  is not a m.d.s., then  $g_{tm-m}^{(m)} = e^{(m)'} G_{tm-m}^{(m)}$  with

$$e^{(m)} = A^{(m)'} e, \quad G_{tm-m}^{(m)} = G_{tm-m} + e^{(m)} (e^{(m)'} e^{(m)})^{-1} e' B^{(m)}. \quad (2.12)$$

Besides,  $G_{tm-m}^{(m)}$  is a VAR(1) process with an autoregressive matrix  $\Gamma^{(m)}$  given by

$$\Gamma^{(m)} = \Gamma^m. \quad (2.13)$$

This proposition means that while the semi-strong ARMA class is not closed under temporal aggregation, the particular state-space representation that we consider is.<sup>11</sup> Thus, the class of ARMA processes defined by the multiperiod conditional moment restrictions (2.9) is closed under temporal aggregation. Note that this class is endowed with richer properties than the weak ARMA which makes it more interesting for both financial and statistical purposes (see below).

Let us focus at this stage on the interpretations of the above results which will be particularly pertinent in the context of the temporal aggregation of volatility models. The resulting variable  $g_{tm}^{(m)}$  at the low frequency is the conditional mean of the aggregated process  $z_{(t+1)m}^{(m)}$  given the information at the **high frequency**  $\sigma(z_{tm-\tau}, G_{tm-\tau}, \tau \geq 0)$  (see the first part of (2.10)). Thus, it is an affine function of the initial state variable  $G_{tm}$  (second part of (2.10)).<sup>12</sup> Then, assuming that  $e'A^{(m)} \neq 0$ ,<sup>13</sup> we can rewrite  $g_{tm}^{(m)}$  as a marginalization of a new state variable  $G_{tm}^{(m)}$  which is indeed the original one plus a constant (see (2.12)).<sup>14</sup> Therefore, it is also a VAR(1) and the low frequency autoregressive coefficient is equal to the high frequency coefficient to the power  $m$  (see (2.13)). Thus persistence increases exponentially with the frequency. Finally, note that even when we consider temporal aggregation of flow variables, the corresponding state variable is considered as a stock variable. This is why the temporal aggregation result holds since semi-strong AR(1) stock variables are closed under temporal aggregation.

<sup>11</sup>Temporal aggregation of stock variables observed at the dates  $m, 2m, 3m, \dots, Tm$ , corresponds to  $a = (1, 0, 0, \dots, 0)'$  while for flow variables  $a = (1, \dots, 1)'$ .

<sup>12</sup>This is due to the Markovianity in mean and to the autoregressive form of  $g_{tm}$ .

<sup>13</sup>This assumption is not restrictive. Indeed,  $e'A^m = 0$  means that the process  $z_{tm}^{(m)}$  is a m.d.s.; Therefore, it is a degenerate state-space model.

<sup>14</sup>As usual, the state-space representation is not unique. For instance, we can consider  $\tilde{G}_{tm}^{(m)} = A^{(m)}G_{tm} + B^{(m)}$  as a state variable. In this case, if one assumes that the matrix  $A^{(m)}$  is non singular, then  $\tilde{G}_{tm}^{(m)}$  is a VAR(1) with an autoregressive matrix equal to  $\Gamma^{(m)}$ . In other words, we refer to the state  $G_{tm}^{(m)}$  rather than  $\tilde{G}_{tm}^{(m)}$  because we have to assume only that  $e'A^{(m)} \neq 0$  rather than  $A^{(m)}$  being non singular.

## 2.2 From continuous time to discrete time

Several models in financial economics are defined in continuous time. However, the data are typically available in discrete time. Therefore, for inference purposes, it is necessary to derive the implied restrictions fulfilled by the data. A natural approach is to derive the exact likelihood of the data from a fully parametric continuous time model. However this likelihood does not admit in general a closed form expression (see however Ait-Sahalia, 1998). Therefore alternative approaches are developed like non parametric methods (Ait-Sahalia, 1996), simulated methods (see Gouriéroux and Monfort, 1996), Bayesian methods (Elarian, Chib and Shephard, 1998) or GMM method (Hansen and Sheinkman, 1995). However, the template of continuous time process allowing one to derive the likelihood for discrete time data is the Ornstein-Uhlenbeck (OU) process:

$$dY_t = K(\Theta - Y_t)dt + \Sigma dW_t, \quad (2.14)$$

where  $Y_t \in \mathbb{R}^p$ ,  $\Theta \in \mathbb{R}^p$ ,  $K$  is a matrix of size  $(p \times p)$  and  $dW_t$  is a p-variate standard Wiener process. In this case, for any  $h > 0$ , the process  $\{Y_{\tau h}, \tau \in \mathbf{Z}\}$  is a conditionally Gaussian VAR(1) process with a conditional mean given by  $(Id - e^{-Kh})\Theta + e^{-Kh}Y_{(\tau-1)h}$ . Note however that the VAR structure of the conditional mean is indeed only due to the linear structure of the drift. Therefore, given a process  $\{Y_t, t \in \mathbb{R}^p\}$  defined by

$$dY_t = K(\Theta - Y_t)dt + \Sigma_t dW_t, \quad (2.15)$$

where the matrix  $\Sigma_t$  can depend on  $Y_t$  or on additional variables  $F_t$ ,<sup>15</sup> the process  $\{Y_{\tau h}, \tau \in \mathbf{Z}\}$  is a semi-strong VAR(1) process, that is:

$$E[Y_{th} | Y_{\tau h}, \tau \leq t-1] = (Id - e^{-Kh})\Theta + e^{-Kh}Y_{(t-1)h}. \quad (2.16)$$

Such processes will be of interest in our paper. Furthermore, we also require their positivity for volatility modeling purposes.

## 2.3 Autoregression and positivity

We consider three examples of autoregressive processes which are at the same time nonnegative.

- **Example 1:** Let us consider the process  $\{z_t, t \in \mathbb{R}\}$  which is the stationary solution of:

$$dz_t = k(\theta - z_t)dt + \delta(z_t)^\lambda dW_t, \quad (2.17)$$

where  $1/2 \leq \lambda \leq 1$  ensures that there exists a nonnegative stationary process solution of (2.17).<sup>16</sup> This is the class of CEV processes. When  $\lambda = 1/2$ , we say that it is the square-root process. Note that (2.17) is a univariate version of (2.15). Thus, from (2.16) we deduce that the **nonnegative** process  $\{z_{\tau h}, \tau \in \mathbf{Z}\}$  is a semi-strong **AR**(1) process.

- **Example 2:** Let  $\{x_t, t \in \mathbf{Z}\}$  the process defined by  $x_t = ax_{t-1} + u_t$  where  $|a| < 1$  and the  $u_t$  are i.i.d.  $\mathcal{N}(0, \sigma^2)$ . Define  $\{z_t, t \in \mathbf{Z}\}$  by  $z_t \equiv x_t^2$ . Then it is straightforward to show that

$$E[z_t | z_\tau, \tau < t] = a^2 z_{t-1} + \sigma^2 \quad \text{and} \quad \text{Var}[z_t | z_\tau, \tau < t] = 2\sigma^2(\sigma^2 + 2a^2 z_{t-1}). \quad (2.18)$$

<sup>15</sup>Of course, the choice of  $\Sigma_t$  is such that there exists a unique stationary solution of (2.15).

<sup>16</sup>Note that the existence of a stationary solution can be guaranteed without the restriction  $\lambda \leq 1$ ; see Conley et al. (1995).

The first result of (2.18) says that the nonnegative process  $z_t$  is an AR(1) while the second one implies that it is conditionally heteroskedastic. Note, however, that conditional heteroskedasticity is not necessary to ensure the nonnegativity of an AR process (see the following example). We can also adapt this example in continuous time. More precisely, consider  $\{x_t, t \in \mathbb{R}\}$  the stationary solution of the stochastic differential equation (SDE)  $dx_t = -kx_t dt + \sigma dW_t$ , with  $k > 0$ , and define the process  $\{z_t, t \in \mathbb{R}\}$  by  $z_t \equiv x_t^2$ . Then by the Ito Lemma we have  $dz_t = (\sigma^2 - 2kz_t)dt + 2x_t \sigma dW_t$ , which can be rewritten as  $dz_t = (\sigma^2 - 2kz_t)dt + 2\sqrt{z_t} \sigma d\tilde{W}_t$ . In other words,  $z_t$  is a constrained square-root process.

• **Example 3:** Let  $\{z_t, t \in \mathbf{Z}\}$  the process defined by  $z_t = \omega + \rho z_{t-1} + v_t$  where,  $0 < \omega$ ,  $0 \leq \rho < 1$  and  $v_t$  i.i.d.  $\mathcal{D}(0, \sigma^2)$ . The process  $z_t$  has the following MA( $\infty$ ) representation:  $z_t = \sum_{i=0}^{+\infty} \rho^i (v_t + \omega)$ . Thus, nonnegativity of  $z_t$  is ensured when  $v_t \geq -\omega$  almost surely. A particular example is an AR(1) plus a positive noise (e.g., Barndorff-Nielsen, Jensen and Sorensen, 1998; and Barndorff-Nielsen and Shephard, 1999). More precisely, let  $\tilde{v}_t$  be an integrable positive i.i.d. process. Define  $z_t$  as the stationary solution of  $z_t = \tilde{\omega} + \rho z_{t-1} + \tilde{v}_t$ , with  $\tilde{\omega} > 0$  and  $0 < \rho < 1$ . Then  $z_t = \omega + \rho z_{t-1} + v_t$  with  $\omega = \tilde{\omega} + E[\tilde{v}_t]$ ,  $v_t = \tilde{v}_t - E[\tilde{v}_t]$  and, hence,  $v_t \geq -E[\tilde{v}_t] \geq -\omega$ .

### 3 SR-SARV(p) model

In this section we introduce the Square Root Stochastic Autoregressive Volatility model of order p (SR-SARV(p)) in discrete and continuous times. The main feature of these models involves a state-space representation of order p for the squared (innovation) process. We prove the consistency between these two models by showing that exact discretization of continuous time SR-SARV(p) model is a discrete time SR-SARV(p) model. This result suggests that this class of models is closed under temporal aggregation and we therefore prove the aggregation result. Then we derive observable restrictions of our model. It provides multiperiod conditional moment restrictions of p lags which hold for the squared process. When the fourth moment is finite, it ensures an ARMA structure for the squared innovation process which is intermediate between weak and semi-strong. Finally we recall the definitions of semi-strong GARCH and weak GARCH and their links with the ARMA structure of the squared innovations.

#### 3.1 The model

##### 3.1.1 Discrete time SR-SARV(p) model

**Definition 3.1. Discrete time SR-SARV(p) model:** A stationary squared integrable process  $\{\varepsilon_t, t \in \mathbf{Z}\}$  is called a **SR-SARV(p) process** with respect to an increasing filtration  $J_t, t \in \mathbf{Z}$ , if:

- i) the process  $\varepsilon_t$  is adapted w.r.t.  $J_t$ , that is  $I_t \subset J_t$  where  $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$ ;
- ii)  $\varepsilon_t$  is a martingale difference sequence w.r.t.  $J_{t-1}$ , that is  $E[\varepsilon_t | J_{t-1}] = 0$ ;
- iii) the conditional variance process  $f_{t-1}$  of  $\varepsilon_t$  given  $J_{t-1}$  is a marginalization of a stationary  $J_{t-1}$ -adapted VAR(1) of size p:

$$f_{t-1} \equiv \text{Var}[\varepsilon_t | J_{t-1}] = e' F_{t-1}, \quad (3.1)$$

$$F_t = \Omega + \Gamma F_{t-1} + V_t, \text{ with } E[V_t | J_{t-1}] = 0, \quad (3.2)$$

where  $e \in \mathbb{R}^p$ ,  $\Omega \in \mathbb{R}^p$  and the eigenvalues of  $\Gamma$  have modulus smaller than one.

Observe that our definition is defined for a given information set  $J_t$ .  $J_t$  can strictly contain the information set  $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$ , which is the minimal information set. In particular,  $J_t$  may contain macroeconomic variables, information about other assets and markets, the volume of transactions, the spread, the order book and so on.<sup>17</sup> Indeed, we never assume that the econometrician observes the full information set  $J_t$  even if the economic agent does. Thus, the model is a Stochastic Volatility (SV) model since the conditional variance process is a function of possibly latent variables.

The process of interest  $\varepsilon_t$  is assumed to be a martingale difference sequence w.r.t the large information  $J_t$  and therefore w.r.t.  $I_t$ . Typically,  $\varepsilon_t$  could be the log-return of a given asset with a price at time  $t$  denoted by  $S_t$ :  $\varepsilon_t = \text{Log}(S_t/S_{t-1})$ . This assumption of m.d.s. is widespread in financial economics and related to the notion of informational efficiency of asset markets. However, we do not preclude predictable log-returns; in this case, our  $\varepsilon_t$  should be interpreted as the innovation process (see Meddahi and Renault, 1996). In addition,  $\varepsilon_t^2$  admits the state-space representation  $\varepsilon_t^2 = e'F_{t-1} + (\varepsilon_t^2 - E[\varepsilon_t^2 | J_{t-1}])$ . Therefore,  $\varepsilon_t^2$  is endowed with the state-space model properties like, e.g., multiperiod conditional moment restrictions. Note however that we do not assume that the fourth moment is finite. We only assume the integrability of the conditional variance process and, hence, the finiteness of the second moment. Observe also that we do not assume that the components of the vectors  $e$  and  $F_t$  are nonnegative. We need only that the process  $e'F_t$  is non negative. A sufficient but not necessary condition for this is that all the components of  $e$  and  $F_t$  are nonnegative. For instance consider the case of  $p = 2$ . Assume that  $e = (1, 1)'$  and  $F_t = (f_{1,t}, f_{2,t})'$  where  $f_{1,t}$  and  $f_{2,t}$  are positive and independent AR(1) processes (for instance following the example 3 in the previous section). Define  $\tilde{f}_{1,t} = f_{1,t} + E[f_{2,t}]$  and  $\tilde{f}_{2,t} = f_{2,t} - E[f_{2,t}]$  and  $\tilde{F}_t = (\tilde{f}_{1,t}, \tilde{f}_{2,t})'$ . Obviously  $\tilde{f}_{2,t}$  is not a nonnegative process while  $e'\tilde{F}_t$  is (since  $e'\tilde{F}_t = e'F_t$ ).

This model is related to Andersen (1994) SR-SARV and indeed we adopt his terminology. However, there are several differences between the models.<sup>18</sup> More precisely, Andersen (1994) considers a fully parametric model by specifying the complete distribution of the process  $(\varepsilon_t, F_t)'$ . The temporal aggregation requirement prevents us from completely specifying the probability distributions. Indeed, neither distributional assumptions nor homo-conditional moments restrictions (homoskewness, homokurtosis) are closed under temporal aggregation (see below). Furthermore, Andersen (1994) excludes leverage effects while we do not. Note that we do specify neither this

---

<sup>17</sup>Note also that  $\sigma(\varepsilon_\tau, f_\tau, \tau \leq t) \subset J_t$  since the process  $f_t$  is adapted w.r.t.  $J_t$ .

<sup>18</sup>Andersen (1994) considers the general class of SARV models where a function of the conditional variance process is a polynomial of an AR(1) Markov process. When this function is the square-root, Andersen (1994) calls it Square-Root (SR) SARV while he terms Exponential SARV when this function is the exponential one, corresponding to the Taylor (1986) and Harvey, Ruiz and Shephard (1994) lognormal SV model.

leverage effect  $Cov[\varepsilon_t, F_t]$  (which is a multivariate leverage effect) nor the high order moments of  $\varepsilon_t$  (third, fourth...) and  $V_t$ .<sup>19</sup> To summarize, we consider a semiparametric SV model.

Now we consider continuous time stochastic volatility models which are popular in finance due to their positivity; then, we prove that exact discretization of these processes are discrete time SR-SARV(p) ones.

### 3.1.2 Continuous time SR-SARV(p) model

**Definition 3.2. Continuous time SR-SARV(p) model:** *A continuous time stationary process  $\{y_t, t \in \mathbb{R}\}$  is called a SR-SARV(p) process with respect to an increasing filtration  $J_t, t \in \mathbb{R}$ , if and only if there exists a  $p$ -variate process  $F_t^c$  such that  $y_t$  is the stationary solution of*

$$d \begin{pmatrix} y_t \\ F_t^c \end{pmatrix} = \begin{pmatrix} 0 \\ K(\Theta - F_t^c) \end{pmatrix} dt + R_t dW_t, \quad (3.3)$$

where  $W_t$  is a  $(p+1)$ -variate standard Wiener process adapted w.r.t  $J_t$ ,  $K$  is a  $p \times p$  matrix with eigenvalues that have positive real parts<sup>20</sup> and  $R_t$  is a  $(p+1) \times (p+1)$  lower triangular matrix, such that the coefficient  $r_{11,t}$  is the square-root of  $r_{11,t}^2 \equiv \sigma_t^2 = e' F_t^c$ , with  $e \in \mathbb{R}_+^p$ .

The instantaneous conditional variance of  $(y_t, F_t^c)$  given  $J_t$  is  $R_t R_t'$ . The matrix  $R_t$  is lower triangular,<sup>21</sup> therefore the conditional variance of  $y_t$  given  $J_t$  is  $r_{11,t}^2$ . In other words, we follow the main idea of the discrete time SR-SARV(p) model, that is the conditional variance is a marginalization of a  $p$  dimensional VAR(1) positive process  $F_t^c$ . Note that as for the discrete time model, we have a semiparametric SV model since we do not define completely the matrix  $R_t$ . In particular, we allow for a leverage effect. Of course, the matrix  $R_t$  has to fulfill conditions ensuring existence and uniqueness of a stationary solution of the SDE (3.3). For instance, this is consistent with the Duffie and Kan (1996) setting of a multivariate square-root process such that each coefficient of  $R_t R_t'$  is of the form  $(1, F_t^c)' \tilde{e}$  with  $\tilde{e} \in \mathbb{R}^{p+1}$ .<sup>22</sup> For  $p = 1$ , we can consider a CEV process<sup>23</sup>

$$d\sigma_t^2 = k(\theta - \sigma_t^2)dt + \delta(\sigma_t^2)^\lambda dW_{2,t}, \quad \text{with } 1/2 \leq \lambda \leq 1. \quad (3.4)$$

Finally, note that the framework allows for models where there are additional factors in  $R_t$ .

We will now prove that the two previous definitions are indeed consistent since exact discretization of continuous time SR-SARV(p) model is a discrete time SR-SARV(p) one:

### Proposition 3.1 Exact discretization of continuous time SR-SARV(p)

<sup>19</sup> Andersen (1994) considers only one factor, so his model is related to a SR-SARV(1). However, he defined the volatility process as a function of a polynomial, say of degree  $p$ , of an AR(1) state-variable  $K_t$ . Thus, it is a marginalization of the vector  $(K_t, K_t^2, \dots, K_t^p)'$  which is indeed a VAR(1) of size  $p$ . In other words, Andersen (1994) considers implicitly a particular SARV(p) model.

<sup>20</sup> Indeed, a usual assumption, see e.g. Bergstrom (1990), page 53, is that the eigenvalues of  $K$  are distinct. Therefore  $K$  is diagonalisable, i.e. there exists a matrix  $H$  such that  $HKH^{-1} = \text{Diag}(\lambda_1, \dots, \lambda_p) \equiv \Lambda$ . As a consequence, for  $u > 0$ ,  $He^{-uK}H^{-1} = e^{-u\Lambda} = \text{Diag}(e^{-u\lambda_1}, \dots, e^{-u\lambda_p})$  with  $e^Z = \sum_{i=0}^{\infty} \frac{Z^i}{i!}$ . The positivity of the real parts of the eigenvalues  $K$  ensures the existence of  $e^{-uK} \forall u > 0$ .

<sup>21</sup> This Gram-Schmidt normalization rule is standard and can be maintained without loss of generality.

<sup>22</sup> See Duffie and Kan (1996) for sufficient conditions of existence of a stationary solution of (3.3) in this case.

<sup>23</sup> Since there is only one factor, we change the notations by taking  $F_t^c \equiv \sigma_t^2$ ,  $W_t = (W_{1,t}, W_{2,t})'$ .

Let  $\{y_t, t \in \mathbb{R}\}$  be a continuous time SR-SARV( $p$ ) process with a corresponding factor process  $\{F_t^c, t \in \mathbb{R}\}$ . Then, for any sampling interval  $h$ , the associated discrete time process  $\varepsilon_{th}^{(h)} = y_{th} - y_{(t-1)h}$ ,  $t \in \mathbf{Z}$ , is a SR-SARV( $p$ ) process w.r.t.  $J_{th}^{(h)}$ ,  $J_{th}^{(h)} = \sigma(\varepsilon_{\tau h}^{(h)}, F_{\tau h}^c, \tau \leq t, \tau \in \mathbf{Z})$ . The corresponding conditional variance process  $f_{(t-1)h}^{(h)} \equiv \text{Var}[\varepsilon_{th}^{(h)} | J_{(t-1)h}^{(h)}]$  is given by  $f_{th}^{(h)} = e'F_{th}^{(h)}$  with  $F_{th}^{(h)} = A^{(h)}F_{th}^c + B^{(h)}$ , where  $A^{(h)} = K^{-1}(Id - e^{-Kh})$  and  $B^{(h)} = (hId - A^{(h)})\Theta$ .

This proposition<sup>24</sup> means that exact discretization of the factor or stochastic volatility models of Duffie and Kan (1996), Constantinides (1992), Heston (1993), are also factor or SV models. Moreover, such models imply conditional moment restrictions based only on the observable variables (see later).

More recently, Barndorff-Nielsen and Shephard (1999) have considered a new class of continuous time stochastic volatility models, termed positive OU processes, that can be exactly discretized. Indeed we can prove that exact discretization of positive OU processes are SR-SARV (see Lemma A.3 in the Appendix). This positive OU processes are very useful in finance, since they allow us to simulate the integrated volatility which is important in option pricing of SV models (Hull and White, 1987).

The previous result suggests that the SR-SARV( $p$ ) class is closed under temporal aggregation. This is the main focus of interest of the paper and the purpose of the next subsection.

### 3.2 Temporal aggregation of SR-SARV( $p$ ) models

We consider here the general temporal aggregation of a given process. More precisely, let consider a process  $\{\varepsilon_t, t \in \mathbf{Z}\}$ , where it is assumed that we observe the process  $\{\varepsilon_{tm}^{(m)}, t \in \mathbf{Z}\}$  defined by

$$\varepsilon_{tm}^{(m)} = \sum_{k=0}^{m-1} a_k \varepsilon_{tm-k}, \quad (3.5)$$

with  $m \in \mathbf{N}^*$ ,  $a = (a_0, a_1, \dots, a_{m-1})' \in \mathbb{R}^m$ . Temporal aggregation of stock variables observed at the dates  $m, 2m, 3m, \dots, Tm$ , corresponds to  $a = (1, 0, 0, \dots, 0)'$ , while for flow variables  $a = (1, \dots, 1)'$ . This latter case is particularly suitable for log-returns and continuously-compounded interest rates.

#### Proposition 3.2 Temporal aggregation of SR-SARV( $p$ ) models

Let  $\varepsilon_t$  be a SR-SARV( $p$ ) process w.r.t. an increasing filtration  $J_t$  and a conditional variance process  $f_t = e'F_t$ . For a given integer  $m$ , the process  $\varepsilon_{tm}^{(m)}$  defined by (3.5) is a SR-SARV( $p$ ) w.r.t.  $J_{tm}^{(m)} = \sigma(\varepsilon_{\tau m}^{(m)}, F_{\tau m}, \tau \leq t)$ . More precisely, we have:

$$f_{tm-m}^{(m)} \equiv \text{Var}[\varepsilon_{tm}^{(m)} | J_{tm-m}] = e'(A^{(m)}F_{tm-m} + B^{(m)}), \quad (3.6)$$

$$\text{where } A^{(m)} = \sum_{k=0}^{m-1} a_k^2 \Gamma^{m-k-1} \text{ and } B^{(m)} = \left( \sum_{k=0}^{m-1} a_k^2 \left( \sum_{i=0}^{m-k-2} \Gamma^i \right) \right) \Omega. \quad (3.7)$$

Assume that  $e'A^{(m)} \neq 0$ , then  $f_{tm}^{(m)} = e^{(m)'}F_{tm}^{(m)}$  with

$$e^{(m)} = A^{(m)'}e, \quad F_{tm}^{(m)} = F_{tm} + e^{(m)}(e^{(m)'}e^{(m)})^{-1}e'B^{(m)}. \quad (3.8)$$

As well,  $F_{tm}^{(m)}$  is a VAR(1) process with an autoregressive matrix  $\Gamma^{(m)}$  given by

$$\Gamma^{(m)} = \Gamma^m. \quad (3.9)$$

---

<sup>24</sup>Note that in this proposition, the discrete time state variable is  $A^{(h)}F_{th}^c + B^{(h)}$  and not, as in Proposition 2.3,  $F_{th}^c$  plus a constant. The reason is that we are sure that the matrix  $A^{(h)}$  is not singular (see footnote 14).

Therefore, models where the conditional variance is a marginalization of a VAR(1) process are robust to temporal aggregation. Note that this result is not a direct application of the temporal aggregation property of the state-space representation (Proposition 2.3): here we consider the dynamics of  $(\sum_{i=0}^m a_i \varepsilon_{tm-i})^2$ , while in the previous case we had characterized the dynamics of  $\sum_{i=0}^m a_i z_{tm-i}$ . Actually, in section 2, the process  $z_t$  has a state-space representation of order  $p$  while here it is  $\varepsilon_t^2$  and not  $\varepsilon_t$ . However, the two results stem from the same intuition. Consider the initial process  $\varepsilon_t$  with the information  $J_t$  at high frequency and define the process at low frequency  $\varepsilon_{tm}^{(m)}$  by (3.5). Define  $f_{tm}^{(m)}$  as the conditional variance of  $\varepsilon_{(t+1)m}^{(m)}$  given the information at **high frequency**  $J_{tm}$  (first part of 3.6). This information is generally not observable either by the agent or by the econometrician and thus the variance is stochastic. But by something like a Markovian property,<sup>25</sup> the conditional variance  $f_{tm}^{(m)}$  is a function of  $F_{tm}$ . By the linearity of the model, this function is indeed affine (second part of (3.6)). Define the information at low frequency by  $J_{tm}^{(m)} \equiv \sigma(\varepsilon_{\tau m}^{(m)}, F_{\tau m}^{(m)}, \tau \leq t)$ . Then  $\varepsilon_{tm}^{(m)}$  is still a m.d.s. with respect to  $J_{tm}^{(m)}$  since  $E[\varepsilon_{tm+m}^{(m)} | J_{tm}^{(m)}] = 0$  and  $J_{tm}^{(m)} \subset J_{tm}$ . Of course, by definition, the conditional variance  $f_{tm}^{(m)}$  of  $\varepsilon_{tm+m}^{(m)}$  given  $J_{tm}^{(m)}$  is positive. Then assuming that  $e' A^{(m)} \neq 0$ ,<sup>26</sup> we can rewrite this conditional variance as a marginalization of a new state variable  $F_{tm}^{(m)}$ . The latter is a VAR(1) since it is the sum of a VAR(1) and a constant. Thus,  $\varepsilon_{tm}^{(m)}$  is a SR-SARV(1) w.r.t.  $J_{tm}^{(m)}$ . Finally, the autoregressive parameter of the VAR(1)  $F_{tm}^{(m)}$  is equal to the autoregressive parameter of the high frequency vector  $F_t$  to the power  $m$  (3.9). It means that the persistence increases exponentially with the frequency. Conversely, conditional heteroskedasticity vanishes when the frequency is low. This corresponds to a well-documented empirical evidence and was pointed out by Diebold (1988), Drost and Nijman (1993), and by Christoffersen and Diebold (2000) in a free model framework.

Temporal aggregation of conditionally heteroskedastic models was already considered by Drost and Nijman (1993) and lead to the weak GARCH paradigm while the links between continuous time SV models and weak GARCH were put forward by Drost and Werker (1996). In the next subsection, we will recap these results and characterize the links between weak GARCH and SR-SARV models.

### 3.3 Observable restrictions

#### 3.3.1 Multiperiod conditional moment restrictions

The SR-SARV is defined w.r.t. an increasing filtration  $J_t$ , which may not be observable by the economic agent or the econometrician. However, as in the previous section, since a SR-SARV( $p$ ) implies that  $\varepsilon_t^2$  has a state-space representation of order  $p$ , we can derive conditional moments fulfilled

---

<sup>25</sup>If one has in mind an underlying continuous time representation like (3.3), the low frequency process  $(y_{tm}^{(m)}, F_{tm}^{(m)})$  is Markovian. More generally, our setting ensures that the conditional variance  $f_{tm}^{(m)}$  depends on past information only through  $F_{tm}$ .

<sup>26</sup>As in the previous section, this assumption is not restrictive. The equality  $e' A^{(m)} = 0$  would mean that the process  $\varepsilon_{tm}^{(m)}$  is homoskedastic which is a degenerate SR-SARV model. In other words, temporal aggregation would cancel the volatility effect.

by the observable process  $\varepsilon_t$  given the minimal information  $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$ . This means that the information generated by the returns should belong, in any case, to the econometrician's information set. These restrictions are multiperiod ones of order  $p$ .

**Proposition 3.3 SR-SARV and multiperiod restrictions**

Let  $\{\varepsilon_t, t \in \mathbf{Z}\}$  be a stationary process. It admits a SR-SARV( $p$ ) representation w.r.t. an increasing filtration  $J_t$  if and only if there exist  $p+1$  reals  $\omega, \gamma_1, \dots, \gamma_p$ , such that the roots of  $1 - \sum_{i=1}^p \gamma_i L^i$  are outside the unit circle and

$$E[\varepsilon_t^2 - \omega - \sum_{i=1}^p \gamma_i \varepsilon_{t-i}^2 \mid \varepsilon_\tau, \tau \leq t - p - 1] = 0. \tag{3.10}$$

Note that this result is not a direct application of Proposition 2.2 since the information structures are different. However, the proof is similar. Therefore, when the fourth moment of  $\varepsilon_t$  is finite,  $\varepsilon_t^2$  is an ARMA( $p, p$ ) defined by (3.10), that is an ARMA property which is intermediate between weak and semi-strong. The (semi-strong) ARMA structure was the main idea of the ARCH models introduced by Engle (1982) and generalized by Bollerslev (1986). Indeed, the clustering effect in financial data that these models account for is directly related to the ARMA structure of the squared residuals.

For temporal aggregation purposes, Drost and Nijman (1993) introduce the weak GARCH models where the squared residuals process is a weak ARMA. Following the Drost and Nijman (1993) terminology, we precisely define below the various concepts and show how they are nested.

**3.3.2 GARCH( $p, q$ )**

**Definition 3.3. GARCH( $p, q$ ):** Let a stationary process  $\{\varepsilon_t, t \in \mathbf{Z}\}$  and define the processes  $\{h_t, u_t, t \in \mathbf{Z}\}$  by the stationary solution of

$$B(L)h_t = \omega + A(L)\varepsilon_t^2 \tag{3.11}$$

and  $u_t = \varepsilon_t / \sqrt{h_t}$ , with  $A(L) = \sum_{i=1}^q \alpha_i L^i$ ,  $B(L) = 1 - \sum_{i=1}^p \beta_i L^i$  where the roots of  $B(L) - A(L)$  and  $B(L)$  are assumed to be different and outside the unit circle. We say that:

i)  $\varepsilon_t$  is a strong GARCH( $p, q$ ) if the process  $u_t$  is i.i.d.  $D(0, 1)$ ;

ii)  $\varepsilon_t$  is a semi-strong GARCH( $p, q$ ) if the process  $u_t$  is such that

$$E[u_t \mid \varepsilon_\tau, \tau \leq t - 1] = 0 \quad \text{and} \quad \text{Var}[u_t \mid \varepsilon_\tau, \tau \leq t - 1] = 1; \tag{3.12}$$

iii)  $\varepsilon_t$  is a weak GARCH( $p, q$ ) if

$$EL[\varepsilon_t \mid H_{t-1}] = 0 \quad \text{and} \quad EL[\varepsilon_t^2 \mid H_{t-1}] = h_t, \tag{3.13}$$

where  $EL[x_t \mid H_{t-1}]$  denotes the best linear predictor of  $x_t$  on the Hilbert space,  $H_{t-1}$ , spanned by  $\{1, \varepsilon_\tau, \varepsilon_\tau^2, \tau \leq t - 1\}$ , that is

$$E[(x_t - EL[x_t \mid H_{t-1}])\varepsilon_{t-i}^r] = 0 \quad \text{for } i \geq 1 \quad \text{and } r = 0, 1, 2. \tag{3.14}$$

Note that in the strong and semi-strong cases, we do not assume that the fourth moment is finite while the weak setting requires this assumption.

**Proposition 3.4 Semi-strong GARCH and ARMA**

Let  $\{\varepsilon_t, t \in \mathbf{Z}\}$  be a m.d.s. ( $E[\varepsilon_t | \varepsilon_\tau, \tau \leq t - 1]$ ). It is a semi-strong GARCH( $p, q$ ) if and only if  $\varepsilon_t^2$  is a semi-strong ARMA( $\max\{p, q\}, p$ ) with an innovation process which is a m.d.s. w.r.t.  $I_t$ .

We specify in our Proposition that the innovation process of the squared process is a m.d.s. w.r.t.  $I_t$  and not only w.r.t.  $\tilde{I}_t = \sigma(\varepsilon_\tau^2, \tau \leq t)$  since the conditional variance process is defined given  $I_t$  (and not  $\tilde{I}_t$ ). Bollerslev (1988) already remarked that a strong GARCH( $p, q$ ) is a semi-strong ARMA( $\max\{p, q\}, p$ ). Note that strong GARCH implies only semi-strong ARMA: when  $\varepsilon_t^2/h_t$  is i.i.d., the ARMA process  $\varepsilon_t^2$  should in general be conditionally heteroskedastic.

Since  $\varepsilon_t^2$  is a semi-strong ARMA, it fulfills a multiperiod conditional moment restriction of order  $\max(p, q)$ .<sup>27</sup> Therefore, Proposition 3.3 implies that  $\varepsilon_t$  admits a SR-SARV( $\max\{p, q\}$ ) representation.

**Corollary 3.1 Semi-strong GARCH and SR-SARV**

Let  $\{\varepsilon_t\}$  be a semi-strong GARCH( $p, q$ ). Then  $\{\varepsilon_t\}$  is a SR-SARV( $\max\{p, q\}$ ) w.r.t.  $I_t$ .

Furthermore, the continuous time SR-SARV( $p$ ) is related to GARCH( $p, p$ ) model. To our knowledge, the relationship between GARCH( $p, p$ ) modeling of higher order ( $p > 1$ ) and continuous time stochastic volatility models was not clearly stated before in the literature, whatever the approach of diffusion approximating (Nelson, 1990), filtering (Nelson and Foster, 1994) or closing the GARCH Gap (Drost and Werker, 1996). Finally, the temporal aggregation of a GARCH model is a SR-SARV model. In other words, to close the class of GARCH processes, we have to plug it into the stochastic volatility class of models. This is not a surprising result since we know from that semi-strong ARMA are not closed under temporal aggregation.

In the next section, we give additional insights as to why GARCH models are not robust to temporal aggregation. Drost and Nijman (1993) already focused on this weakness of standard GARCH models. They give examples of strong and semi-strong GARCH which are not closed under temporal aggregation. Then, they introduce the weak GARCH model where the squared residuals are weak ARMA in order to benefit from the temporal aggregation of the weak ARMA structure.

**Proposition 3.5 Weak GARCH and ARMA**

Let  $H_{t-1}^s$  the Hilbert space spanned by  $\{1, \varepsilon_\tau^2, \tau \leq t - 1\}$  and  $\{h_t^s, \eta_t\}$  the processes defined by  $h_t^s = EL[\varepsilon_t^2 | H_{t-1}^s]$  and  $\eta_t = \varepsilon_t^2 - h_t^s$ . If  $\varepsilon_t$  is a weak GARCH( $p, q$ ) process, then  $h_t = h_t^s$  a.s. and, hence,  $\varepsilon_t^2$  is a weak stationary ARMA( $\max\{p, q\}, p$ ) process and

$$Cov(\eta_t, \varepsilon_\tau) = 0, \forall \tau < t. \tag{3.15}$$

Conversely, if  $\varepsilon_t^2$  is a weak stationary ARMA( $q, p$ ) process and (3.15) holds, then  $\varepsilon_t$  is a weak GARCH( $p, q$ ).

---

<sup>27</sup>More precisely, a semi-strong ARMA( $\tilde{q}, \tilde{p}$ ) implies a multiperiod conditional moment restrictions of order equal to  $\max\{\tilde{q}, \tilde{p}\}$ . Thus a semi-strong ARMA( $\max\{p, q\}, p$ ) implies a multiperiod restriction of order  $\max\{p, q\}$ .

Thus, the weak GARCH property is slightly more restrictive than the weak ARMA assumption for the squared residuals. In particular, (3.15) is like a symmetry assumption, which is implied by the maintained condition m.d.s. for  $\varepsilon_t$  when assuming semi-strong GARCH. In fact, Drost and Nijman (1993) take a “coherent” definition in the sense that they project both the residual and its square onto the same space  $H_{t-1}$ . However, the ARMA structure of the squared residuals was the main idea of the weak GARCH.<sup>28</sup> As we can already see, the class of weak ARMA strictly contains the class of ARMA models with a semi-strong state-space representation and finite variance. Therefore, it means that weak GARCH are in fact **Stochastic Volatility** models, i.e., Drost and Nijman (1993) plug also the class of GARCH models into the SV one.<sup>29</sup>

However, to show that weak GARCH class is closed under temporal aggregation for flow variables, Drost and Nijman (1993) maintain at least one of the following symmetry assumptions:

$$\forall h \in \mathbf{N}^*, \forall (a_k)_{1 \leq k \leq h} \in \{-1, 1\}^h, (\varepsilon_{t+k})_{1 \leq k \leq h} = (a_k \varepsilon_{t+k})_{1 \leq k \leq h} \text{ in distribution, or} \quad (3.16)$$

$$\forall 0 \leq i \leq j \ E[\varepsilon_t \varepsilon_{t-i} \varepsilon_{t-j}] = 0 \text{ and } \forall 0 \leq i \leq j \leq k, i \neq 0 \text{ or } j \neq k \ E[\varepsilon_t \varepsilon_{t-i} \varepsilon_{t-j} \varepsilon_{t-k}] = 0. \quad (3.17)$$

Such symmetry restrictions are indeed quite restrictive both from theoretical and empirical grounds. They preclude two types of asymmetry which appear relevant for financial data. First, even in the strong GARCH setting, the probability distribution of the standardized innovations  $\varepsilon_t/\sqrt{h_t}$  may be skewed. Second, since the weak GARCH models are SV ones (outside the standard GARCH class), another type of asymmetry (termed the leverage effect by Black, 1976, and popularized by Nelson, 1991) may matter. A clear distinction between these two types of asymmetric behavior of a general SR-SARV(1) process will be made in section 4 below. Equivalently, the leverage effect can be introduced in the continuous time setting by allowing the volatility matrix  $R_t$  to be non-diagonal, unlike the case considered by Drost and Werker (1996) and Andersen and Bollerslev (1998). Finally, note that our results concerning temporal aggregation and exact discretization are consistent with those of Drost and Nijman (1993) and Drost and Werker (1996).<sup>30</sup> In particular, the restrictions on the parameters are the same ( $\Gamma^{(m)} = \Gamma^m$ ).<sup>31</sup>

## 4 SR-SARV(1)

### 4.1 SR-SARV(1) and GARCH(1,1)

The GARCH(1,1) model is nowadays dominant w.r.t. any other ARCH or GARCH type model in the empirical finance literature. We now discuss in more detail its relationships with the general SR-SARV(1). In the previous section, we proved that a semi-strong GARCH(p,q) is also

---

<sup>28</sup>When Nijman and Sentana (1996) and Drost and Weker (1996) prove respectively that a marginalization of a multivariate GARCH and that the discretization of (3.3) for p=1 under (3.4) are weak GARCH, they only deal with the ARMA property of squared residuals.

<sup>29</sup>See the following section where we establish the exact links between SR-SARV and weak GARCH.

<sup>30</sup>Nevertheless, Drost and Werker (1993) consider only the one factor case.

<sup>31</sup>For more details, see Meddahi and Renault (1996).

a SR-SARV( $\max\{p, q\}$ ). We first characterize the SR-SARV(1) processes which are also semi-strong GARCH(1,1).

**Proposition 4.1 Semi-strong GARCH(1,1) and SR-SARV(1)**

Let  $\{\varepsilon_t, t \in \mathbf{Z}\}$  be a SR-SARV(1) process with a conditional variance process  $f_t$ .  $\varepsilon_t$  is a semi-strong GARCH(1,1) with  $\alpha > 0$  and  $\beta \geq 0$  if and only if: i)  $\varepsilon_t^2$  and  $f_t$  are conditionally perfectly positively correlated given  $J_{t-1}$ ; ii) the ratio  $Var[f_t | J_{t-1}]/Var[\varepsilon_t^2 | J_{t-1}]$  is constant and smaller or equal to  $\gamma^2$ . In this case:  $h_{t+1} = f_t$ ,  $J_t = I_t$  and  $\beta = \gamma - \alpha$  with  $\alpha = \sqrt{Var[f_t | J_{t-1}]/Var[\varepsilon_t^2 | J_{t-1}]}$ .

The first restriction means that GARCH models correspond to the degenerate case where there are no exogenous sources of randomness in the conditional variance. This degeneracy corresponds to GARCH only if it is a perfect linear conditional correlation. The second restriction is less known even though it was already coined by Nelson and Foster (1994). They observed that the most commonly used ARCH models assume that the variance of the variance rises linearly with the square of the variance, which is the main drawback of GARCH models in approximating SV models in continuous time. Thus, semi-strong GARCH setting implies nontrivial restrictions on conditional kurtosis dynamics.

On the other hand, Nelson (1991) stressed that one limitation of GARCH models is that only the magnitude and not the sign of unanticipated excess returns affects the conditional variance. Therefore, alternative asymmetric GARCH models have been introduced in the literature. For instance, Glosten, Jagannathan and Runkle (1989, GJR) introduce a model based on a GARCH formulation but accounting for the sign of the past residuals. More generally, asymmetric models have been studied and compared by Engle and Ng (1993) who consider the following models:

$$\text{GJR} : h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma S_{t-1}\varepsilon_{t-1}^2, \text{ where } S_t = 1 \text{ if } \varepsilon_t < 0, S_t = 0 \text{ otherwise} \quad (4.1)$$

$$\text{Asymmetric GARCH} : h_t = \omega + \alpha(\varepsilon_{t-1} + \gamma)^2 + \beta h_{t-1}; \quad (4.2)$$

$$\text{Nonlinear Asymmetric GARCH} : h_t = \omega + \alpha(\varepsilon_{t-1} + \gamma\sqrt{h_{t-1}})^2 + \beta h_{t-1}; \quad (4.3)$$

$$\text{VGARCH} : h_t = \omega + \alpha(\varepsilon_{t-1}/\sqrt{h_{t-1}} + \gamma)^2 + \beta h_{t-1}; \quad (4.4)$$

Let us also consider a related model considered by Heston and Nandi (1999):<sup>32</sup>

$$\text{Heston and Nandi} : h_t = \omega + \alpha(\varepsilon_{t-1}/\sqrt{h_{t-1}} - \gamma\sqrt{h_{t-1}})^2 + \beta h_{t-1}. \quad (4.5)$$

Actually, we show that all these models are in the SR-SARV(1) class.<sup>33</sup>

**Proposition 4.2 Asymmetric GARCH and SR-SARV(1)**

Let  $\{\varepsilon_t, t \in \mathbf{Z}\}$  be a m.d.s. and define  $h_t$  the conditional variance of  $\varepsilon_t$ , i.e.  $h_t \equiv Var[\varepsilon_t | \varepsilon_\tau, \tau \leq t-1]$ . Assume that  $h_t$  is given by (4.2), (4.3), (4.4), or by (4.5), then  $\varepsilon_t$  is a SR-SARV(1) model. If  $u_t = \varepsilon_t/\sqrt{h_t}$  is i.i.d., then the GJR model defined by (4.1) is also a SR-SARV(1).

<sup>32</sup>Heston and Nandi (1999) show that the diffusion limit of (4.5) is the stationary solution of (3.4) with  $\lambda = 1/2$ , i.e. the model considered by Heston (1993). Both the discrete time and the continuous time models provide closed-form option pricing formulas.

<sup>33</sup>Finally, Drost (1993) shows that symmetric QARCH of Sentana (1995) are weak GARCH. Indeed, it is easy to show that any QARCH is a SR-SARV model. This is also the case of the HARARCH model of Muller and al. (1997) since this model is a restricted QARCH.

## 4.2 SR-SARV(1) and weak GARCH(1,1)

We will now focus on the relationships between SR-SARV and weak GARCH. As already mentioned, Drost and Nijman (1993) prove the temporal aggregation property of symmetric weak GARCH (assuming (3.16) or (3.17)) which excludes the leverage effect and all the asymmetric models considered in Proposition 4.2. We specify two kinds of asymmetries for the SR-SARV model:

**Definition 4.1. Leverage effect and skewness:** Let  $\{\varepsilon_t, t \in \mathbf{Z}\}$  be a SR-SARV(1) process w.r.t. a filtration  $J_t$  with corresponding processes  $\{f_t, u_t, \nu_t\}$  where  $f_t = \omega + \gamma f_{t-1} + \nu_t$ . We say that:

i)  $\varepsilon_t$  does not present (conditional) leverage effect w.r.t.  $J_t$  if and only if

$$E[u_t \nu_t | J_{t-1}] = 0 \quad \text{or} \quad E[\varepsilon_t \varepsilon_{t+1}^2 | J_{t-1}] = 0; \quad (4.6)$$

ii)  $\varepsilon_t$  does not present (conditional) skewness w.r.t.  $J_t$  if and only if

$$E[u_t^3 | J_{t-1}] = 0 \quad \text{or} \quad E[\varepsilon_t^3 | J_{t-1}] = 0. \quad (4.7)$$

We show in the Appendix that the two conditions of (4.6) (and (4.7)) are equivalent. Now we can show that a SR-SARV model without leverage effect and skewness is a weak GARCH.

### Proposition 4.3 Weak GARCH(1,1) and SR-SARV(1)

If  $\varepsilon_t$  is a SR-SARV(1) process with finite fourth moment and without leverage effect and skewness, that is if (4.6) and (4.7) hold, then  $\varepsilon_t$  is a weak GARCH(1,1) process.

Therefore, there is no major difference between symmetric weak GARCH and symmetric SR-SARV. However, we do not prove an equivalence result, and it is clear that the class of symmetric weak GARCH is larger than one of symmetric SR-SARV. Indeed, one can interpret the weak GARCH model as a SV model, but not endowed with a sufficiently rich specification for statistical inference and economic interpretation. In addition, we have proved in section 3 that this weakness is not needed to close the GARCH gap with continuous time as in Drost and Werker (1996). In a sense, by introducing the SR-SARV, we have enriched the weak GARCH models by adding useful restrictions for financial and statistical interpretations. Furthermore, SR-SARV allows for asymmetries like the leverage effect and skewness. Indeed, the corresponding symmetry assumptions are closed under temporal aggregation.

### Proposition 4.4 Temporal aggregation, leverage effect and skewness

Let  $\{\varepsilon_t, t \in \mathbf{Z}\}$  be a SR-SARV process w.r.t. an increasing filtration  $J_t$  with corresponding processes  $\{f_t, u_t, \nu_t, t \in \mathbf{Z}\}$ . Define  $\varepsilon_{tm}^{(m)}$  by (3.5) and the corresponding SR-SARV(1) representation of Proposition 3.2,  $J_{tm}^{(m)}, \{f_{tm}^{(m)}, u_{tm}^{(m)}, \nu_{tm}^{(m)}\}$ . Then the symmetric SR-SARV class is closed under temporal aggregation. More precisely, we have:

$$E[u_t \nu_t | J_{t-1}] = 0 \implies E[u_{tm}^{(m)} \nu_{tm}^{(m)} | J_{tm-m}^{(m)}] = 0, \quad \text{and} \quad (4.8)$$

$$E[u_t \nu_t | J_{t-1}] = E[u_t^3 | J_{t-1}] = 0 \implies E[(u_{tm}^{(m)})^3 | J_{tm-m}^{(m)}] = 0. \quad (4.9)$$

This proposition means that our results generalize those of Drost and Nijman (1993) and Drost and Werker (1996), since symmetric SR-SARV are weak GARCH and are closed under temporal aggregation. Besides, the relationships between parameters at various frequencies, already stressed by these authors (particularly the persistence parameter) are maintained in our SR-SARV setting. Furthermore this proposition means that a symmetry assumption about the standardized innovation (a no-skewness effect) cannot be alleged without precluding leverage effect as well (see 4.9). Therefore, when one observes significant skewness at a low frequency, it may be due either to genuine skewness or to leverage effect at the high frequency, while the presence of the leverage effect at a low frequency implies the same feature at the high frequency.

**Proposition 4.5 Observable restrictions of leverage effect and skewness**

Let  $\varepsilon_t$  be a SR-SARV(1) w.r.t. an increasing filtration  $J_t$ .

i) If  $\varepsilon_t$  is without leverage effect ((4.6) holds), then

$$E[\varepsilon_t \varepsilon_{t+1}^2 \mid I_{t-1}] = 0. \tag{4.10}$$

ii) If  $\varepsilon_t$  is without skewness ((4.7) holds), then

$$E[\varepsilon_t^3 \mid I_{t-1}] = 0. \tag{4.11}$$

Therefore we can derive moments restrictions based on observable data which can be used to test the absence of leverage effect or skewness. Moreover, usual GARCH allows for leverage effect as soon as there is skewness since the conditions (4.6) and (4.7) are equivalent in this case. Indeed, in the introduction of his EGARCH paper, Nelson (1991) explicitly mentions that **symmetric** GARCH models do not take into account leverage effect.

## 5 Temporal aggregation of IGARCH models

We have considered in the previous sections temporal aggregation of volatility models with integrable volatility. However, some empirical evidence supports the Integrated GARCH model introduced by Engle and Bollerslev (1986). This process is not second-order stationary, since the second moment is infinite. This evidence is even more pronounced for high frequency data (5 and 10 minutes returns); see for instance Andersen and Bollerslev (1997a) and Gençay et al. (1998). While the second moment of the residuals is not finite, the notion of conditional variance is valid since the squared residual process is nonnegative and hence admits a conditional expectation. Moreover, we know that the GARCH(1,1) process is strictly stationary when  $E[\ln(\beta + \alpha u_t^2)] < 0$  (and  $\omega > 0$ ) with i.i.d. standardized residuals (see Nelson, 1990). This condition is ensured when  $\alpha + \beta = 1$ ,<sup>34</sup> that is for IGARCH(1,1). Therefore we can extend our notion of SR-SARV to nest the IGARCH class.

**Definition 5.1. Integrated SR-SARV(1) model:** A strictly stationary process  $\{\varepsilon_t, t \in \mathbf{Z}\}$  is called an **ISR-SARV(1) process** w.r.t.  $J_t$  if:

---

<sup>34</sup>By Jensen inequality, we have  $E[\ln(\beta + \alpha u_t^2)] < \ln E[\beta + \alpha u_t^2] = \ln(\alpha + \beta) = 0$ .

- i) the process  $\varepsilon_t$  is adapted w.r.t.  $J_t$ , that is  $I_t \subset J_t$  where  $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$ ;
- ii)  $\varepsilon_t$  is a martingale difference sequence w.r.t.  $J_{t-1}$ , that is  $E[\varepsilon_t | J_{t-1}] = 0$ ;
- iii) the conditional variance process  $f_{t-1}$  of  $\varepsilon_t$  given  $J_{t-1}$  is such that:

$$E[f_t | J_{t-1}] = \omega + f_{t-1}. \quad (5.1)$$

Obviously an IGARCH(1,1) is an ISR-SARV(1). Note that strict stationarity is not important for modeling purposes since we can remove it in the definition of an ISR-SARV. However, it is useful for inference. We now consider temporal aggregation of ISR-SARV:

**Proposition 5.1 Temporal aggregation of ISR-SARV(1)**

Let  $\varepsilon_t$  be an ISR-SARV(1) process w.r.t. an increasing filtration  $J_t$  and a conditional variance process  $f_{t-1}$ . The process  $\varepsilon_{tm}^{(m)}$  defined by  $\varepsilon_{tm}^{(m)} \equiv \sum_{k=0}^{m-1} a_k \varepsilon_{tm-k}$  is an ISR-SARV(1) w.r.t.  $J_{tm}^{(m)} = \sigma(\varepsilon_{\tau m}^{(m)}, f_{\tau m}, \tau \leq t)$ .

As a consequence, a temporally aggregated IGARCH process is also an integrated process but of SV type.<sup>35</sup> Empirically, the IGARCH model is rejected at low frequencies, e.g. monthly. Therefore by the aggregation result, one has to conclude that the model at high frequency is not an integrated one. A potential explanation of this is long memory in the volatility. For instance, Bollerslev and Mikkelsen (1998) (resp Comte and Renault, 1998) show via a Monte Carlo study that when the true model is FIGARCH (resp long memory continuous time SV), estimation of a GARCH model by QMLE suggests an IGARCH model. Temporal aggregation of long memory volatility models is beyond the scope of this paper; see Andersen and Bollerslev (1997b) and Bollerslev and Wright (1998).

## 6 Temporal aggregation of non-linear volatility models

In this section, we consider temporal aggregation of non-linear volatility models. While our approach is general, we present it in details in the second subsection for the exponential SV model in discrete time. Then we propose in the last subsection a generalization of our approach to any non-linear model in discrete or continuous time.

### 6.1 SR-SARV( $\infty$ )

As already mentioned, non-linear models are not robust to temporal aggregation while linear ones are. Therefore the natural way to close a non-linear model w.r.t temporal aggregation is to plug it into a class of linear models. We will do so by considering a natural expansion of the considered non-linear function of the underlying process, such as the Hermite expansion in the Gaussian case. So we still get a linear model but with an infinite number of components, which lead us to introduce the SR-SARV( $\infty$ ) class:

---

<sup>35</sup>Engle and Bollerslev (1986) consider temporal aggregation of IGARCH model with  $\omega = 0$  which is not, however, a strictly stationary process. Moreover, the variance process converges a.s. to a constant (Nelson, 1991).

**Definition 6.1. SR-SARV( $\infty$ ) model:** A stationary squared integrable process  $\{\varepsilon_t, t \in \mathbf{Z}\}$  is called a **SR-SARV( $\infty$ ) process** with respect to an increasing filtration  $J_t, t \in \mathbf{Z}$ , if:

- i) the process  $\varepsilon_t$  is adapted w.r.t.  $J_t$ , that is  $I_t \subset J_t$  where  $I_t = \sigma(\varepsilon_\tau, \tau \leq t)$ ;
- ii)  $\varepsilon_t$  is a martingale difference sequence w.r.t.  $J_{t-1}$ , that is  $E[\varepsilon_t | J_{t-1}] = 0$ ;
- iii) the conditional variance process  $f_{t-1}$  of  $\varepsilon_t$  given  $J_{t-1}$  is given by

$$f_{t-1} = \sum_{i=1}^{\infty} e_i f_{i,t-1}, \text{ where} \quad (6.1)$$

the sequence  $\{f_{i,t-1}, t \in \mathbf{Z}, i \in \mathbf{N}\}$  is such that  $\exists q$  with: a) the vector  $F_{q,t} = (f_{1,t}, \dots, f_{q,t})'$  is a stationary VAR(1) ( $E[F_{q,t} | J_{t-1}] = \Omega_q + \Gamma_q F_{q,t-1}$ ) and uncorrelated with any  $f_{n,t}$ ,  $n > q$ ; b)  $\forall n > q$ , the processes  $\{f_{n,t}\}$  are uncorrelated and univariate centered stationary AR(1) ( $E[f_{n,t} | J_{t-1}] = \gamma_i f_{n,t-1}$ ) where  $\sup_{i>q} |\gamma_i| < 1$  and the roots of  $\Gamma_{11}$  are outside the unit circle; c)

$$\sup_n E |f_{n,t}|^2 < +\infty \text{ and } \sum_{i=1}^{\infty} |e_i|^2 < +\infty. \quad (6.2)$$

In other words, with a misuse of terminology, we maintain the assumption that the conditional variance of  $\varepsilon_t$  given  $J_{t-1}$  is a marginalization of an infinite VAR(1),  $F_t \equiv (f_{1,t}, f_{2,t}, \dots)'$ . Note that the AR(1) assumption of  $f_{n,t}$  for  $n$  sufficiently large might be relaxed since it is not needed for the concept of SR-SARV( $\infty$ ). However, it is useful for temporal aggregation purposes. With a misuse of notation, we have  $E[F_t | J_{t-1}] = \Omega + \Gamma F_{t-1}$ . For temporal aggregation, we need to define power of infinite matrix, e.g.  $\Gamma^m$ . The sequence  $\{f_{i,t-1}, t \in \mathbf{Z}, i \in \mathbf{N}\}$  is well-suited for this. This is also the reason why we make the assumption on  $\sup_i |\gamma_i|$ . Under (6.2),  $\sum_{i=1}^p e_i f_{i,t}$  converges in mean square to  $\sum_{i=1}^{\infty} e_i f_{i,t}$ . We implicitly assume that the fourth moment of the residual is finite, because we consider an infinite dimensional Hilbert space. However, this assumption is not restrictive in most non-linear volatility models as Exponential SV and EGARCH.<sup>36</sup> Finally, note that recently Giraitis, Kokoszka and Leipus (2000) studied the probabilistic and statistical properties of a broad class of non-negative ARCH( $\infty$ ) models. We can now state the generalization of the Proposition 3.2 to the case of SR-SARV( $\infty$ ), i.e., SR-SARV( $\infty$ ) class is closed under temporal aggregation:

**Proposition 6.1 Temporal aggregation of SR-SARV( $\infty$ ) model**

Let  $\varepsilon_t$  be a SR-SARV( $\infty$ ) process w.r.t. an increasing filtration  $J_t$  and  $\{f_{1,t}, f_{2,t}, \dots\}$  the corresponding sequence defining the variance process. The process  $\varepsilon_{tm}^{(m)}$  defined by  $\varepsilon_{tm}^{(m)} \equiv \sum_{k=0}^{m-1} a_k \varepsilon_{t-m-k}$  is a SR-SARV( $\infty$ ) w.r.t.  $J_{tm}^{(m)} = \sigma(\varepsilon_{\tau m}^{(m)}, f_{1,\tau}, f_{2,\tau}, \dots, \tau \leq t)$ .

---

<sup>36</sup>This is an important difference between EGARCH and GARCH models. EGARCH share this property with SV models. The main reason is that in EGARCH and SV models, the (log of the) variance is the infinite sum of an i.i.d process. Therefore there are no restrictions on the variance of the (log) variance. In the GARCH case, the variance of the variance is a function of the variance. However in the VGARCH model defined by (4.4), it is easy to show that any moment is finite as soon as  $\beta < 1$ . For moment properties of various volatility models, see Carrasco and Chen (1999).

Let us apply this general result to the Exponential SV model by showing that these model is SR-SARV( $\infty$ ) and, hence, its SR-SARV( $\infty$ ) structure is closed under temporal aggregation.

## 6.2 Temporal aggregation of exponential SV models

In this section, we consider temporal aggregation of the exponential SV model considered by Taylor (1986) and popularized by Harvey, Ruiz and Shephard (1994). This model is defined by

$$\varepsilon_t = \exp(g_{t-1}/2) u_t, \text{ with} \quad (6.3)$$

$$g_t = \omega + \gamma g_{t-1} + v_t, \text{ where} \quad (6.4)$$

$(u_t, v_t)$  is i.i.d. and normal  $\mathcal{N}(0, \Sigma)$ , with  $\Sigma[1, 1] = 1$  and  $\Sigma[2, 2] = \sigma_v^2$ .<sup>37</sup>

Note that in order to accommodate leverage effect we do not assume that  $\Sigma[1, 2]$  is equal to zero. In this model, the log of the squared residual process has a linear strong state-space representation. Therefore we can not use Proposition 3.2 to address the temporal aggregation issue of such models since the conditional variance process of  $\varepsilon_t$  given  $J_{t-1} = \sigma(\varepsilon_\tau, g_\tau, \tau \leq t-1)$ , which is equal to  $\exp(g_{t-1})$ , is not linear in the state variable  $g_{t-1}$ . However we can adapt them to study the temporal aggregation of non-linear SV models by considering an expansion of the conditional variance  $\exp(g_{t-1})$  in terms of Hermite Polynomials. In other words, we will consider linear models with an infinite number of factors corresponding to a SR-SARV( $\infty$ ). More precisely, define  $\mu$  (resp  $\sigma^2$ ) as the mean (resp variance) of  $g_t$  ( $\mu = \omega/(1 - \gamma)$ ,  $\sigma^2 = \sigma_v^2/(1 - \gamma^2)$ ), and  $H_i(\cdot)$  the Hermite polynomials, characterized by

$$H_0(x) = 1, H_1(x) = x \text{ and } \forall i > 1, H_i(x) = xH_{i-1}(x) - (i-1)H_{i-2}(x). \quad (6.5)$$

It is well known that  $\{H_i(\cdot), i \in \mathbf{N}\}$  is an orthonormal basis of the squared integrable function of Gaussian processes (see e.g., Granger and Newbold, 1977, page 305), i.e.,

$$\forall f(\cdot) \text{ such that } E[f(g_t)^2] < \infty, \exists \psi_i, i = 0, 1, \dots, f(g_t) = \sum_{i=0}^{\infty} \psi_i H_i\left(\frac{g_t - \mu}{\sigma}\right), \text{ where} \quad (6.6)$$

$$\forall i \neq 0 E[H_i\left(\frac{g_t - \mu}{\sigma}\right)] = 0 \text{ and } E[H_i^2\left(\frac{g_t - \mu}{\sigma}\right)] = i!, \quad \forall i, j, i \neq j E[H_i\left(\frac{g_t - \mu}{\sigma}\right)H_j\left(\frac{g_t - \mu}{\sigma}\right)] = 0. \quad (6.7)$$

Moreover,  $\forall i \geq 1$ ,  $H_i\left(\frac{g_t - \mu}{\sigma}\right)$  is a semi-strong AR(1) with autoregressive coefficient equal to  $\gamma^i$ :

$$\forall i \geq 1, E[H_i\left(\frac{g_{t+1} - \mu}{\sigma}\right) | g_\tau, u_\tau, \tau \leq t] = \gamma^i H_i\left(\frac{g_t - \mu}{\sigma}\right). \quad (6.8)$$

Consider the normalized Hermite polynomials defined by  $H_i^*(\cdot) \equiv H_i(\cdot)/i!$ . The log-normal conditional variance  $\exp(g_{t-1})$  of  $\varepsilon_t$  is square-integrable and then

$$f_{t-1} = \text{Var}[\varepsilon_t | \varepsilon_\tau, g_\tau, \tau \leq t-1] = \exp(g_{t-1}) = \sum_{i=0}^{\infty} \psi_i^* H_i^*\left(\frac{g_{t-1} - \mu}{\sigma}\right) \quad (6.9)$$

with  $\psi^* = \psi\sqrt{i!}$ . Hence  $f_{t-1} = \sum_{i=1}^{\infty} e_i f_{i,t-1}$  where  $e_1 = 1$ ,  $f_{1,t-1} = \psi_0^* + \psi_1^* H_1^*\left(\frac{g_{t-1} - \mu}{\sigma}\right)$ ,  $e_i = \psi_i^*$  and  $f_{i,t-1} = H_i^*\left(\frac{g_{t-1} - \mu}{\sigma}\right)$  for  $i > 2$ . The elements of the sequence  $\{f_{i,t}, i\}$  are uncorrelated and are univariate semi-strong AR(1). This suggests that the exponential SV is a SR-SARV( $\infty$ ) w.r.t.

$J_t = \sigma(\varepsilon_\tau, g_\tau, \tau \leq t)$ :<sup>38</sup>

<sup>37</sup>Several methods for estimating this model are considered in the literature. For instance, Harvey, Ruiz and Shephard (1994) propose a QMLE method based on a Kalman-filtering approach while Jacquier, Polson and Rossi (1994) consider a Bayesian approach.

<sup>38</sup>Since  $H_1^*(x) = x$ , we have  $J_t = \sigma(\varepsilon_\tau, g_\tau, \tau \leq t) = \sigma(\varepsilon_\tau, H_i^*\left(\frac{g_\tau - \mu}{\sigma}\right), \tau \leq t, i = 1, 2, \dots) = \sigma(\varepsilon_\tau, H_1^*\left(\frac{g_\tau - \mu}{\sigma}\right), \tau \leq t)$ .

**Proposition 6.2 Exponential SV model is a SR-SARV( $\infty$ )**

The exponential SV process  $\{\varepsilon_t\}$  defined by (6.3-6.4) is a SR-SARV( $\infty$ ) w.r.t.  $J_t = \sigma(\varepsilon_\tau, g_\tau, \tau \leq t)$ .

Therefore, while the class of exponential SV models is not closed under temporal aggregation, it can be plugged into the SR-SARV( $\infty$ ) class which is. Of course the aggregated process is not an exponential SV model (in general).<sup>39</sup> Moreover, distributional assumptions like Gaussianity of the standardized residuals are not closed under temporal aggregation. Such assumptions are crucial in non-linear models, in particular for multi-step prediction. Note that, like exponential SV, Gaussian EGARCH are SR-SARV( $\infty$ ) by a similar argument.

**6.3 Temporal aggregation of non-linear models**

Consider now a more general non-linear model in discrete or continuous time. The main idea of the temporal aggregation of exponential volatility models is to plug this class into the SR-SARV( $\infty$ ) by taking an appropriate expansion of the conditional variance process. The approach used above in the Gaussian exponential SV case can be extended to non-Gaussian and continuous time settings. Note however that, while presented now in continuous time, the same arguments can be applied in discrete time as well. Consider any continuous time SV model defined by

$$dy_t = \sigma_t dw_t, \tag{6.10}$$

$$\sigma_t^2 = f(z_t), \tag{6.11}$$

$$dz_t = \mu(z_t)dt + \sigma(z_t)d\tilde{w}_t, \tag{6.12}$$

where  $(dw_t, d\tilde{w}_t)'$  is a bivariate Wiener process and assume that:

**Assumption A1** The process  $\{z_t\}$  is stationary and time reversible;

**Assumption A2** The infinitesimal generator of  $\{z_t\}$  has discrete spectrum, with eigenvalues  $\lambda_i$  and corresponding eigenfunctions  $f_i(t)$ ,  $i = 0, 1, \dots$ , such that  $\lambda_0 = 0$ ,  $\lambda_0 > \lambda_1 > \lambda_2 \dots$ ;

**Assumption A3** The function  $f(\cdot)$  is in the domain of the infinitesimal generator of  $\{z_t\}$ .

As shown by Hansen, Scheinkman and Touzi (1998), under appropriate boundary protocol, stationary **scalar** diffusions are time reversible.<sup>40</sup> Moreover, they give sufficient conditions related to the boundaries to ensure Assumption 2 and characterize the domain of the infinitesimal generator of  $\{z_t\}$ . Note that Assumption 3 implies that the process  $\sigma_t^2 = f(z_t)$  is square integrable.

Then, under A1-A3 (see, e.g., Darolles, Florens and Gouriéroux, 1998)

$$\sigma_t^2 = f(z_t) = \sum_{i=0}^{\infty} e_i f_i(z_t), \quad \text{with} \tag{6.13}$$

$$e_i = E[f(z_t)f_i(z_t)] \text{ and}$$

$$\forall i \geq 0, \forall h > 0, \quad E[f_i(z_{t-1+h}) | J_{t-1}] = \exp(-\lambda_i h) f_i(z_{t-1}) \tag{6.14}$$

---

<sup>39</sup>Note that one considers temporal aggregation of stock variables of Gaussian exponential SV, then the aggregated process is also a Gaussian exponential SV model.

<sup>40</sup>See Florens, Renault and Touzi (1998) for a discussion of the observable implications of reversibility.

where  $J_{t-1} = \sigma(\varepsilon_\tau, z_\tau, \tau \leq t-1)$ . This expansion was recently considered by Chen, Hansen and Scheinkman (1999) and Darolles, Florens and Renault (1998). In particular, they call the functions  $f_i(z_t)$  as the non-linear principal components.<sup>41</sup> The AR(1) dynamics (6.14) of these principal components will allow us to show that  $\varepsilon_t = \int_{t-1}^t \sigma_u dW_u$  is a SR-SARV( $\infty$ ) w.r.t.  $J_{t-1}$ . Define  $f_{t-1} \equiv \text{Var}[\varepsilon_t | \varepsilon_\tau, z_\tau, \tau \leq t-1]$ . Then

$$f_{t-1} = E\left[\int_{t-1}^t \sigma_u^2 | J_{t-1}\right] = \int_{t-1}^t \sum_{i=0}^{\infty} e_i E[f_i(z_u) | J_{t-1}] = \int_{t-1}^t \sum_{i=0}^{\infty} e_i \exp[-\lambda_i(u - (t-1))] f_i(z_{t-1}) du.$$

Hence,<sup>42</sup>  $f_{t-1} = \sum_{i=0}^{\infty} \tilde{e}_i f_i(z_{t-1})$ , where  $\tilde{e}_i = e_i \int_{t-1}^t \exp[-\lambda_i(u - (t-1))] du = e_i(1 - \exp(-\lambda_i))/\lambda_i$ . Note that  $|\tilde{e}_i| \leq |e_i|$ . By Parseval identity, we have  $\sum_{i=0}^{\infty} e_i^2 = E[f(z_t)^2]$  and hence  $\sum_{i=0}^{\infty} \tilde{e}_i^2 < \infty$ . Moreover  $E[f_i(t)^2] = 1 \forall i$ . Therefore,  $\{\varepsilon_t\}$  is a SR-SARV( $\infty$ ) w.r.t.  $J_{t-1}$ , i.e. any exact discretization of (6.10), (6.11) and (6.12) is, under A1-A3, a SR-SARV( $\infty$ ) which is closed under temporal aggregation.

## 7 A summary

In this section, we provide a synthesis of the various volatility models we have considered. Let  $\varepsilon_t$  be a stationary m.d.s. with respect to an increasing filtration  $J_t$ . Assume that the conditional variance of  $\varepsilon_t$  given  $J_{t-1}$ , is a linear combination of the components of a VAR(1) process of size  $p$  where  $p$  may be infinite. This is what we define as a SR-SARV( $p$ ) process w.r.t.  $J_t$ . This definition means that the squared residual is equal to this combination plus noise, i.e., we have a linear state-space representation for the squared residual. When the information set  $J_t$  is fully (resp partially) observed by the econometrician, we say that we have an ARCH-type (resp SV) model. The main result of this paper is that, for a given  $p$ , the SR-SARV( $p$ ) class is closed under temporal aggregation. Moreover, additional symmetry assumptions (e.g., no leverage effect) are closed under temporal aggregation.

Assume for the moment that  $p$  is finite and that the conditional variance is integrable. This class of models contains most of the linear volatility models considered in the literature, as GARCH (Engle, 1982, Bollerslev, 1986), asymmetric GARCH (Glosten, Jagannathan and Runkle, 1989, Engle and Ng, 1993), QARCH (Sentana, 1995), parametric SR-SARV (Andersen, 1994). The SR-SARV( $p$ ) is characterized by multiperiod conditional moment restrictions of type (3.10) fulfilled by  $\varepsilon_t^2$ . These restrictions are weaker than the standard restrictions stating that  $\varepsilon_t$  is a semi-strong GARCH( $p,p$ ), or equivalently that  $\varepsilon_t^2$  is a semi-strong ARMA( $p,p$ ). This is the reason why semi-strong GARCH models are not closed under temporal aggregation.

If we now assume in addition that the conditional variance is square-integrable, i.e. the fourth moment of  $\varepsilon_t$  is finite, then  $\varepsilon_t^2$  is a weak ARMA( $p,p$ ). This is less restrictive than the weak GARCH specification which maintains a symmetry assumption. If one makes the additional symmetry

---

<sup>41</sup>Their nonparametric extraction is considered in Chen, Hansen and Scheinkman (1998) and Darolles, Florens and Gouriéroux (1998).

<sup>42</sup>We adopt the convention that  $(1 - \exp(-\lambda_i))/\lambda_i = 1$  if  $\lambda_i = 0$ .

assumptions that the standardized residuals are not skewed and that there is no leverage effect, then the SR-SARV process is a weak GARCH. As a consequence, weak GARCH are SV models rather than ARCH-type ones. Moreover, the SR-SARV class is the natural generalization of the weak GARCH one which allows for asymmetries and to maintain conditional moment restrictions for inference purposes. Finally, our results generalize those of Drost and Nijman (1993) and Drost and Werker (1996), since these symmetry assumptions are closed under temporal aggregation; see Figure 1.

Assume now that  $p$  is infinite and that the conditional variance is square-integrable. This class of models contains most of the non-linear volatility models as EGARCH (Nelson, 1991), exponential SV in discrete time (Taylor, 1986) and continuous time (Hull and White, 1987). More generally, this class contains any volatility model such that the conditional variance process can be decomposed in a Hilbertian basis of square-integrable AR(1) processes. Special examples are any Gaussian volatility models in discrete time as well as continuous time diffusion models such that the volatility process is Markovian, time-reversible, and the infinitesimal generator of the volatility has a discrete spectrum.

Finally, the ISR-SARV( $p$ ),  $p$  finite, corresponds to the case where the conditional variance process is not integrable. This class contains the IGARCH models and is closed under temporal aggregation.

## 8 A Monte Carlo study

### 8.1 Three examples

In this section, we consider the estimation of three examples of SR-SARV(1) processes.

#### M1: Aggregated GARCH with finite fourth moment

We consider a Gaussian GARCH model at high frequency, that is

$$y_t = \mu^1 + \varepsilon_t = \mu^1 + \sqrt{h_t}u_t, \quad \text{with } h_t = \omega^1 + \alpha^1\varepsilon_{t-1}^2 + \beta^1h_{t-1} \quad (8.1)$$

where  $u_t$  is i.i.d.  $\mathcal{N}(0, 1)$  with  $(\mu^1, \omega^1, \alpha^1, \beta^1, \gamma^1) = (0, 2.8\text{E-}06, .0225, .9770, .9995)$  where  $\gamma^1 = \alpha^1 + \beta^1$ .

We choose these parameters such that after aggregation as flow over  $m$  periods with  $m=400$ , we obtain a weak GARCH model with the coefficients  $(\mu^0, \omega^0, \alpha^0, \beta^0) = (0, 0.4, 0.206, 0.594, 0.8)$ . The persistence parameter at the high frequency,  $\gamma^1$ , is conformable to the empirical study of Andersen and Bollerslev (1997a).  $\alpha^1$  and  $\beta^1$  are chosen such that after temporal aggregation,  $\alpha^0$  and  $\beta^0$  are close to those of a specification considered by Nijman and Sentana (1996).

#### M2: Aggregated GARCH with infinite fourth moment

We consider a Gaussian GARCH model for high frequency data with the coefficients  $(\mu^1, \omega^1, \alpha^1, \beta^1, \gamma^1) = (0, 2.03\text{e-}06, .02960, .9700, .9996)$  such that after temporal aggregation, with  $m = 400$ ,  $(\mu^0, \omega^0, \gamma^0) = (0, 0.3, 0.85)$ . The coefficients  $\alpha^0$  and  $\beta^0$  are not defined since the fourth moment is infinite<sup>43</sup> and the weak GARCH formulas do not apply. However, an automatic application

---

<sup>43</sup>The fourth moment is finite for a strong Gaussian GARCH(1,1) when  $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$  (Bollerslev, 1986).

of the weak GARCH formulas for  $\alpha^{(m)}$ ,  $\beta^{(m)}$ , with  $m=400$ , still provide coefficient  $\alpha^0$  and  $\beta^0$  which are real, positive and smaller than one:  $\alpha^0 = .3276$  and  $\beta^0 = .5224$ .

### M3: Cross-sectional aggregated GARCH

Following Nijman and Sentana (1996), we consider the cross-sectional aggregation of two independent GARCH(1,1) processes with the same persistence parameter. Nijman and Sentana (1996) show that the aggregated process is a weak GARCH(1,1).<sup>44</sup> In Lemma A.4 in the Appendix, we show that such process is a SR-SARV(1). Even if the scope of our paper is temporal aggregation and not cross-sectional aggregation, we consider this example for two reasons. First of all, Nijman and Sentana (1996) show via a Monte Carlo study that the QMLE is not consistent in that case. We show here that our estimation method is consistent. Second, the literature uses high frequency data to study temporal aggregation (see e.g., Andersen and Bollerslev, 1997, Andersen and al., 1999). In this context, the price is often defined as the mean of the Bid and the Ask quotes. Therefore there is cross-sectional aggregation. Since we consider the empirical study of Andersen and Bollerslev (1997a) as a benchmark, we also consider cross-sectional aggregation. See the companion paper Meddahi and Renault (2000b) for a general study of cross-sectional aggregation of volatility models. The individual GARCH processes that we consider are Gaussian with the following parameters:  $(\mu^1, \omega^1, \alpha^1, \beta^1, \gamma^1) = (0, .1, .35, .5, .85)$  and  $(\mu^2, \omega^2, \alpha^2, \beta^2, \gamma^2) = (0, .1, .05, .8, .85)$ . This corresponds to the sixth example of Nijman and Sentana (1996) in Table 4.<sup>45</sup> The aggregated process is a weak GARCH with the parameters  $(\mu^0, \omega^0, \alpha^0, \beta^0, \gamma^0) = (0, .2, .281, .569, .85)$ .

Finally, when the aggregated model is a weak GARCH, that is in the first and the third examples, we also consider a Gaussian GARCH(1,1) with the same parameter  $(\mu^0, \omega^0, \alpha^0, \beta^0, \gamma^0)$  to compare our inference method with the Maximum Likelihood method. We call these models **M01** and **M03**.

## 8.2 Estimation method

Consider a process  $y_t$  defined as a constant plus a SR-SARV(1) process  $\varepsilon_t$  ( $y_t = \mu^0 + \varepsilon_t$ ). The estimation method should be based on the m.d.s. assumption for  $\varepsilon_t$

$$E[y_t - \mu^0 \mid I_{t-1}] = 0 \tag{8.2}$$

and the multiperiod conditional moment restriction of  $\varepsilon_t^2$ :

$$E[\varepsilon_t^2 - \omega^0 - \gamma^0 \varepsilon_{t-1}^2 \mid I_{t-2}] = 0. \tag{8.3}$$

We propose a two-step method. We estimate the constant  $\mu^0$  as the sample mean  $\hat{\mu}_T$  and then we estimate the parameters  $(\omega^0, \gamma^0)$  by using (8.3) after replacing  $\varepsilon_t$  by  $\hat{\varepsilon}_t = y_t - \hat{\mu}_T$ . This two-step method is consistent but not, in general, efficient. However, in our case  $\varepsilon_t$  is symmetric and hence there is no loss of efficiency. Moreover, we have in mind high frequency applications which lead us to

---

<sup>44</sup>In general, when the persistence parameters are not the same, the aggregated process is a weak GARCH(2,2) and a SR-SARV(2); see Nijman and Sentana (1996) and Meddahi and Renault (1996).

<sup>45</sup>The coefficients  $\omega_i$ ,  $i = 1, 2$ , are not provided in their study since they are scale parameters of the variance.

consider very large sample size. Following Drost and Nijman (1992) and Nijman and Sentana (1996), we choose  $T = 80,000$ .<sup>46</sup> The sample mean  $\hat{\mu}_T$  would be a consistent estimator for  $\mu^0$  even if we only assume that  $\varepsilon_t$  is a weak white noise as in weak GARCH. For the second step, we first consider the instruments  $(1, \hat{\varepsilon}_{t-2}^2, \dots, \hat{\varepsilon}_{t-r}^2)$ ,  $r = 9$ , and compute the corresponding optimal GMM estimator (Hansen, 1982).<sup>47</sup> This is basically the Yule-Walker estimation of the weak ARMA process  $\varepsilon_t^2$ .<sup>48</sup> Therefore our method is also consistent for weak GARCH. However an implicit assumption is that the eighth moment is finite (when we compute the GMM optimal weighting matrix). We call this estimator  $\hat{\theta}^{GMM1}$ .

To take advantage of the additional information (w.r.t. Yule-Walker equations) provided by the conditional moment restrictions (8.3), we consider another GMM estimator denoted by GMM2.<sup>49</sup> Define  $z_t$  and  $\hat{z}_t$  by  $z_t = \varepsilon_t^2 - \omega^0 - \gamma^0 \varepsilon_{t-1}^2$  and  $\hat{z}_t = \hat{\varepsilon}_t^2 - \hat{\omega}^{GMM1} - \hat{\gamma}^{GMM1} \hat{\varepsilon}_{t-1}^2$  respectively. When the fourth moment is finite,  $z_t$  is a weak MA(1)  $z_t = \eta_t - \beta^0 \eta_{t-1}$ . Therefore, we follow Hannan and Rissanen (1982) to estimate  $\beta$ . We regress  $\hat{z}_t$  on  $(1, \hat{z}_{t-1}, \dots, \hat{z}_{t-q})$  with  $q = 50$ . We get the residual  $\hat{\eta}_t$ . When  $q \rightarrow +\infty$ ,  $\hat{\eta}_t \rightarrow \eta_t$  since  $z_t$  is a weak MA(1) and hence a weak AR( $\infty$ ). Thus the regression of  $\hat{z}_t$  on  $-\hat{\eta}_{t-1}$  provides a consistent estimate for  $\beta^0$  called  $\hat{\beta}^{GMM1}$ . We define  $\alpha^0$  (resp  $\hat{\alpha}^{GMM1}$ ) as  $\gamma^0 - \beta^0$  (resp  $\hat{\gamma}^{GMM1} - \hat{\beta}^{GMM1}$ ) and  $\hat{h}_t^{GMM1} = \hat{\omega}^{GMM1} + \hat{\alpha}^{GMM1} \hat{\varepsilon}_{t-1}^2 + \hat{\beta}^{GMM1} \hat{h}_{t-1}^{GMM1}$  with  $\hat{\varepsilon}_0^2 = \hat{h}_0^{GMM1} = \hat{\omega}^{GMM1} / (1 - \hat{\gamma}^{GMM1})$ . Then we estimate  $(\omega^0, \gamma^0)$  by using (8.3) with the instruments  $(1, \hat{\varepsilon}_{t-2}^2, \hat{\varepsilon}_{t-3}^2, \dots, \hat{\varepsilon}_{t-r}^2) / (\hat{h}_{t-1}^{GMM1})^2$  and  $r = 9$ . This is a GLS-type correction which take into account the heteroskedasticity of  $z_t$ . Note however that  $h_{t-1}^2$  is not the conditional variance of  $z_t$  given  $I_{t-2}$ .<sup>50</sup>

We also use the Hannan-Rissanen method with this second estimator and obtain  $\hat{\alpha}^{GMM2}, \hat{\beta}^{GMM2}, \hat{h}_t^{GMM2}$ . When the fourth moment is not finite, the asymptotic properties of Hannan-Rissanen estimators are unknown. However the second GMM estimator is consistent for  $(\mu^0, \omega^0, \gamma^0)$  since  $(1, \hat{\varepsilon}_{t-2}^2, \hat{\varepsilon}_{t-3}^2, \dots, \hat{\varepsilon}_{t-r}^2) / (\hat{h}_{t-1}^{GMM1})^2$  is an admissible instrument.

We also estimate all the models by the Gaussian QML method which is the ML method for the models **M01** and **M03**.

### 8.3 Monte Carlo results

In Table 1, we report the results of the three estimation methods. Consider first the parameter of interest of the SR-SARV(1) model, namely  $\theta^0 = (\omega^0, \gamma^0)$ . In the five models, both  $\theta^{GMM1}$  and  $\theta^{GMM2}$  are such that the true values fall within the (Monte Carlo) confidence intervals. However for

<sup>46</sup>Such large sample size makes negligible the lack of efficiency of the two-step method.

<sup>47</sup>Note that in this case, if we estimate the parameter by the unconditional moment restriction  $E[m_t] = 0$ , the optimal weighting matrix is  $Var(m_t) + cov(m_t, m_{t-1}) + cov(m_t, m_{t-1})'$  since the multiperiod restrictions are of order 2. In the empirical study, we use the weight suggested by Newey and West (1987).

<sup>48</sup>Recently, Francq and Zakoian, 1998, proposes a two-step method to estimate weak GARCH which is related to the Yule-Walker method. A limitation of this method is that it does not take into account the heteroskedacticity in the mean and, more importantly, in the variance process.

<sup>49</sup>Another advantage of GMM2 w.r.t. GMM1 is to relax the restrictive assumption of finite eighth moment.

<sup>50</sup>Efficient IV estimation is very difficult in multiperiod restrictions: see Hansen, Heaton and Ogaki (1988).

models **M1** and **M2**, the first estimator presents a systematic bias ( $\omega^0$  is overestimated while  $\gamma^0$  is underestimated). The bias is even more important in model **M2** probably because in this model the fourth moment is not finite and hence the Yule-Walker method does not work. Both in terms of bias and variance, the second estimator  $\theta^{GMM2}$  strictly dominates  $\theta^{GMM1}$ . On the other hand, the lack of efficiency of the GMM estimators w.r.t. ML (in the models **M01** and **M03**) is not dramatic. Finally, the QMLE is not consistent in temporal aggregation and cross-sectional aggregation examples (models **M1**, **M2** and **M3**). To confirm this point, we simulate models **M1** and **M2** with a sample size equal to 150,000 and use the QML method. In Table 2, we report the results over 50 replications. The QMLE is clearly not consistent. This is an important difference with Drost and Nijman (1992) who report simulation results which suggest that the QMLE of temporally aggregated GARCH is consistent or has a very small bias. Our results are different from theirs because they consider aggregation only over short periods ( $m=2,4,8,16$ ) while we consider a much longer one ( $m=400$ ). Moreover, the low frequency parameters are small in their case. For instance, the volatility persistence parameter is .663 with  $m=8$  and .44 with  $m=16$ , while we consider persistence parameters equal to .8 and .85 which are empirically more relevant. See Meddahi (2000a) for more details.

Consider now the parameters  $\alpha^0$  and  $\beta^0$  which are well defined when the fourth moment is finite. In models **M1** and **M3**, both  $(\hat{\alpha}^{GMM1}, \hat{\beta}^{GMM1})$  and  $(\hat{\alpha}^{GMM2}, \hat{\beta}^{GMM2})$  are such that the true values fall within the (Monte Carlo) confidence intervals. However both present much more bias than their sum. Although, they work quite well for the strong GARCH models (**M01** and **M03**) in terms of bias, the lack of efficiency w.r.t. MLE is more significant than the one of their sum. In summary, the GMM method that take into account the heteroskedasticity, that is  $\theta^{GMM2}$ , works very well for the parameters of interest ( $\omega^0, \gamma^0$ ) in all the models while the estimation of  $(\alpha^0, \beta^0)$  (when they are well defined) needs more investigation.

An important outlet of the estimation of the parameters of the volatility process is to produce volatility filters and forecasts. Therefore we study the difference between the conditional variance of an SR-SARV(1) process, i.e.  $f_t = E[\varepsilon_{t+1}^2 | J_t]$ , with the linear volatility, i.e.  $h_{t+1} = EL[\varepsilon_{t+1}^2 | H_t^s]$ . Note that  $h_{t+1}$  is well-defined when the fourth moment is finite and that this is the volatility provided by the weak GARCH representation. Observe that  $h_{t+1} = EL[f_t | H_t]$ . Hence we consider the  $R^2$  of the regression of  $f_t$  onto the constant and  $h_t$  as a measure of the difference between the two processes. We provide in the Appendix, Lemma A.5, the theoretical formula of the  $R^2$  for any SR-SARV(1) with finite fourth moment. Table 3 gives the theoretical  $R^2$  as well as the Monte Carlo results about the  $R^2$  coefficients of the regressions of the true volatility process on the various filtered ones, that is  $h_t^{GMM1}, h_t^{GMM2}, h_t^{QMLE}$ , as well as  $h_t$  computed with the true parameter. Observe that the theoretical  $R^2$  in Model 1 is .59, that is the linear prediction of  $f_t$  (and  $\varepsilon_{t+1}^2$ ) leads in this case to

a severe underestimation of the dynamics of the conditional variance. Moreover the difference in the one-step ahead forecasting of the volatility is not so important between the four forecasts.<sup>51</sup> An important point to keep in mind in this respect is that a volatility forecast at horizon  $h$  will use the estimated value of  $\gamma^h$ . Therefore, the gains in estimation of  $\gamma$  we have obtained thanks to our conditional moment restrictions will be very sensible for multistep forecasting for two reasons: on the one hand, we are safeguarded against the aforementioned possible bias of QMLE; on the other hand, we take advantage of a significant reduction of variance w.r.t. the Yule-Walker estimation method of weak GARCH. Moreover, our comparison consider only linear filters of the volatility process.<sup>52</sup> The SR-SARV model could lead to better volatility filters by going further in non-linear filtering. More precisely, define  $z_t = \varepsilon_t^2 - \omega - \gamma\varepsilon_{t-1}^2$ . While we know that  $E[z_t | I_{t-2}] = 0$ , we would like to go further in estimating  $E[z_t | I_{t-1}]$ . A more accurate filter of the conditional variance could then be  $\hat{\omega} + \hat{\gamma}\hat{\varepsilon}_{t-1}^2 + \hat{E}[\hat{z}_t | I_{t-1}]$ . A well-suited nonparametric procedure is left for future research.

## 9 Conclusion

In this paper, we have considered the temporal aggregation of volatility models. We introduce a semiparametric class of volatility models termed square-root stochastic autoregressive volatility (SR-SARV) characterized by an autoregressive dynamic of the stochastic variance. Our class encompasses the usual GARCH models of Bollerslev (1986), the asymmetric GARCH models of Glosten, Jagannathan and Runkle (1989) and Engle and Ng (1993). Moreover, even if the volatility is stochastic, that is may involve a second source of randomness, the considered models are characterized by observable multiperiod conditional moment restrictions (Hansen, 1985). The SR-SARV class is a natural extension of the weak GARCH models of Drost and Nijman (1993). It extends the weak GARCH class since it does not assume that the fourth moment is finite and, moreover, allows for asymmetries (skewness, leverage effects). On the other hand, it provides a statistical structure which remains true the concept of conditional variance and maintains the validity of conditional moment restrictions, which are useful for inference. Finally we also consider temporal aggregation of IGARCH models and non-linear models as EGARCH and Exponential SV in discrete and continuous time.

In Meddahi and Renault (2000b), we show that our class of SR-SARV is robust to information reduction and, following Nijman and Sentana (1994), to marginalization and contemporaneous aggregation. Finally, Meddahi (2000b) considers SR-SARV-M models where conditional mean and variance may share some factors in common. He shows that this class is closed under temporal aggregation and that temporal aggregation of these models creates automatically a leverage effect.

---

<sup>51</sup>As previously noticed in another context by Nelson (1992), a wrong model can produce a forecast as good as one produced by the true model.

<sup>52</sup>Moreover, the shortcoming of a direct comparison of volatility forecasts is that parameter uncertainty is not taken into account; see e.g. Diebold and Mariano (1995) and West (1996).

## References

- Ait-Sahalia, Y. (1996), "Nonparametric Pricing of Interest Rate derivative Securities," *Econometrica*, 64, 527-560.
- Ait-Sahalia, Y. (1998), "Maximum-Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach," unpublished manuscript, Princeton University.
- Akaike, H. (1974), "Markovian Representation of Stochastic Processes and its Application to the Analysis of Autoregressive Moving Average Processes," *Ann. Inst. Stat. Math.*, 26, 363-387.
- Andersen, T.G. (1994), "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling," *Mathematical Finance*, 4, 75-102.
- Andersen, T.G. and T. Bollerslev (1997a), "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115-158.
- Andersen, T.G. and T. Bollerslev (1997b), "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of Finance*, 52, 975-1005.
- Andersen, T.G. and T. Bollerslev (1998), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885-905.
- Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (1999), "The Distribution of Exchange rate Volatility," *Journal of the American Statistical Association*, forthcoming.
- Barndorff-Nielsen, O.E., Jensen J.L. and M. Sorensen (1998), "Some Stationary Processes in Discrete and Continuous Time," *Advanced Applied Probability*, 30, 989-1007.
- Barndorff-Nielsen, O.E. and N. Shephard (1999), "Non-Gaussian OU based Models and some of their uses in Financial Economics," unpublished paper, Nuffield College, Oxford.
- Black, F. (1976), "Studies in Stock Price Volatility Changes," *Proceedings of the 1976 Business Meeting of the Business and Statistics Section, American Statistical Association*, 177-181.
- Bergstrom, A.R. (1990), *Continuous Time Econometric Modelling*, Oxford University Press.
- Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T. (1988), "On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Processes," *Journal of Times Series Analysis*, 9, 121-131.
- Bollerslev, T., R. Engle and D.B. Nelson (1994), "ARCH Models," *Handbook of Econometrics*, Vol IV.
- Bollerslev, T. and H.L. Mikkelsen (1996), "Modeling and Pricing Long Memory in Stock Market Volatility," *Journal of Econometrics*, 73, 151-184.
- Bollerslev, T. and J.H. Wright (1998), "Semiparametric Estimation of Long-Memory Volatility Dependencies: The Role of High Frequency," *Journal of Econometrics*, forthcoming.
- Brockwell, P.J. and R.A. Davis (1990), "Time Series: Theory and Methods," Second Edition, Springer Verlag.
- Carrasco, M. and X. Chen (1999), " $\beta$ -mixing and Moment Properties of Various GARCH, Stochastic Volatility and ACD Models," CREST DP 9941.

- Chen, X., L.P. Hansen and J. Scheinkman (1998), "Shape-Preserving Estimation of Diffusions," unpublished manuscript, University of Chicago.
- Chen, X., L.P. Hansen and J. Scheinkman (1999), "Principal Components and the Long Run," unpublished manuscript, London School of Economics.
- Christoffersen, P.F. and F.X. Diebold (2000), "How Relevant is Volatility Forecasting for Financial Risk Management," *Review of Economics and Statistics*, 81, 12-22.
- Comte, F. and E. Renault (1998), "Long Memory in Continuous Time Stochastic Volatility Models," *Mathematical Finance*, 8, 291-323.
- Conley, T., L.P. Hansen, E. Luttmer and J. Scheinkman (1997), "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, 10, 525-577.
- Constantinides, J. (1992), "A Theory of the Nominal Term Structure of Interest Rates," *Review of Financial Studies*, 5, 531-552.
- Cox, J.C. (1975), "Notes on Option Pricing I: Constant Elasticity of Variance Diffusions," Discussion Paper, Stanford University.
- Darolles, S., J.P. Florens and C. Gouriéroux (1998), "Kernel-Based Nonlinear Canonical Analysis," CREST DP 9855.
- Darolles, S., J.P. Florens and E. Renault (1998), "Nonlinear Principal Components and Inference on a Conditional Expectation Operator," unpublished manuscript, CREST.
- Davis, R.A. and T. Mikosch (1998), "The Sample Autocorrelations of Heavy-Tailed Processes with Application to ARCH," *Annals of Statistics*, 26, 2049-2080.
- Diebold, F.X. (1988), *Empirical Modeling of Exchange Rates*, Springer Verlag.
- Diebold, F.X. and J.A. Lopez (1995), "Modeling Volatility Dynamics," in *Macroeconometrics*, 427-466, edited by K.V. Hoover, Kluwer Academic Publishers.
- Diebold, F.X. and R.S. Mariano (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253-263.
- Diebold, F.X. and M. Nerlove (1989), "The Dynamics of Exchange Rate Volatility: A Multivariate Latent ARCH Model," *Journal of Applied Econometrics*, 4, 1-21.
- Drost, F.C. (1993), "Temporal Aggregation of Time-Series," in J. Kaehler and P. Kugler eds, *Econometric Analysis of Financial Markets*, 11-21.
- Drost, F.C. and TH.E. Nijman (1992), "Temporal Aggregation of GARCH processes," Tilburg University DP 9240.
- Drost, F.C. and TH.E. Nijman (1993), "Temporal Aggregation of GARCH processes," *Econometrica*, 61, 909-927.
- Drost, F.C. and B.J.M. Werker (1996), "Closing the GARCH Gap: Continuous Time GARCH Modeling," *Journal of Econometrics*, 74, 31-58.
- Duffie, D. and P. Glynn (1997), "Estimation of Continuous-Time Markov Processes Sampled at Random Time Intervals," unpublished paper, Stanford University.
- Duffie, D. and R. Kan (1996), "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 379-406.
- Elarian, O., S. Chib and N. Shephard (1998), "Likelihood Inference for Discretely Observed Non-Linear Diffusions," unpublished manuscript, Nuffield College, Oxford.
- Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.
- Engle, R.F. and T. Bollerslev (1986), "Modelling the Persistence of Conditional Variances," *Econometric Reviews*, 5, 1-50.
- Engle, R.F., and V.K. Ng (1993), "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749-1778.

- Engle, R.F., V.K. Ng and M. Rothschild (1990), "Asset Pricing with a Factor ARCH Structure: Empirical estimates for Treasury Bills," *Journal of Econometrics*, 45, 213-237.
- Florens, J.P., E. Renault and N. Touzi (1998), "Testing for Embeddability by Stationary Reversible Continuous-Time Markov Processes," *Econometric Theory*, 14, 744-769.
- Francq, C. and J.M. Zakoian (1998), "Estimating Weak GARCH Representations," *Econometric Theory*, forthcoming.
- Gençay, R., G. Ballocci, M. Dacorogna, R. Olsen and P. Pictet (1998), "Real-Time Trading Models and Statistical Properties of Foreign Exchange Rates," unpublished paper, University of Windsor.
- Giraitis, L., P. Kokoszka and R. Leipus (2000), "Stationary ARCH Models: Dependence structure and Central Limit Theorem," *Econometric Theory*, 16, 3-22.
- Ghysels, E., A.C. Harvey and E. Renault (1996), "Stochastic Volatility," in: G.S. Maddala and C.R. Rao eds, *in Handbook of Statistics*, vol 14, 119-191.
- Glosten, L.R., R. Jagannathan and D. Runkle (1989), "On the Relation between the Expected Value of the Volatility of the Nominal Excess Return on Stocks," *The Journal of Finance*, Vol XLVIII, 1779-1801.
- Gouriéroux, C. and J. Jasiak (1999), "Dynamic Factor Models," CREST DP 9908.
- Gouriéroux, C. and A. Monfort (1996), *Simulation-Based Econometric Methods*, CORE Lectures, Oxford.
- Granger, C.W.J (1990), "Aggregation of Time Series Variables: A Survey," in T. Barker and H. Pesaran (Eds), *Disaggregation in Econometric Modelling*, pp. 17-34, Ratlege, New-York.
- Granger, C.W.J and P. Newbold (1977), *Forecasting Economic Time Series*, Academic Press.
- Hannan, E.J. and J. Rissanen (1982), "Recursive Estimation of Mixed Autoregressive Moving Average Order," *Biometrika*, 69, 81-94.
- Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.
- Hansen, L.P. (1985), "A Method for Calculating Bounds in the Asymptotic Covariance Matrix of Generalized Method of Moments Estimators," *Journal of Econometrics*, 30, 203-238.
- Hansen, L.P, J.C. Heaton and M. Ogaki (1988), "Efficiency Bounds Implied by Multiperiod Conditional Moment Restrictions," *Journal of the American Statistical Association*, 83, 863-871.
- Hansen, L.P. and J. Scheinkman (1995), "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes," *Econometrica*, 63, 767-804.
- Hansen, L.P., J. Scheinkman and N. Touzi (1998), "Spectral Methods for Identifying Scalar Diffusions," *Journal of Econometrics*, 86, 1-32.
- Hansen, L.P. and K.J. Singleton (1996), "Efficient Estimation of Linear Asset-Pricing Models With Moving Average Errors," *Journal of Business and Economic Statistics*, 14, 53-68.
- Harvey, A.C., E. Ruiz and E. Sentana (1992), "Unobserved Component time series models with ARCH disturbances," *Journal of Econometrics*, 52, 129-157.
- Harvey, A.C., E. Ruiz and N. Shephard (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies*, 61, 247-264.
- Heston, S.L. (1993), "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327-344.

- Heston, S.L. and S. Nandi (1999), "A Closed-Form GARCH Option Pricing Model," unpublished manuscript, Federal Reserve Bank of Atlanta.
- Hull, J. and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, Vol XLII, 281-300.
- Jacquier, E., N.G. Polson and P.E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, 12, 371-389.
- King, M., E. Sentana and S. Wadhvani (1994), "Volatility and Links between National Stock Markets," *Econometrica*, 62, 901-933.
- Lutkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Springer Verlag.
- Meddahi, N. (2000a), "QMLE for Aggregated GARCH: A Monte Carlo Study," in preparation.
- Meddahi, N. (2000b), "Temporal and Cross-sectional Aggregations of Volatility in Mean Models," in preparation.
- Meddahi, N. and E. Renault (1996), "Aggregation and Marginalization of GARCH and Stochastic Volatility Models," GREMAQ DP 96.30.433, Université de Toulouse and CRDE DP 3597, Université de Montréal.
- Meddahi, N. and E. Renault (2000a), "State Space and Multiperiod ARMA Models," in preparation.
- Meddahi, N. and E. Renault (2000b), "Conditionning Information in Volatility Models," in preparation.
- Muller, U.A., M. Dacorogna, R.D. Davé, R.B. Olsen, O.V. Pictet and J.E. von Weizsacker (1997), "Volatilities of Different Time Resolutions - Analyzing the Dynamics of Market Components," *Journal of Empirical Finance*, 4, 213-239.
- Nelson, D.B. (1990), "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7-39.
- Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347-370.
- Nelson, D.B. (1992), "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model," *Journal of Econometrics*, 52, 61-90.
- Nelson, D.B. (1996), "Asymptotically Optimal Smoothing with ARCH Models," *Econometrica*, 64, 561-573.
- Nelson, D.B. and D.P. Foster (1994), "Asymptotic Filtering Theory for Univariate ARCH models," *Econometrica*, 62, 1-41.
- Newey W.K. and K.D. West (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.
- Nijman, TH. and E. Sentana (1996), "Marginalization and Contemporaneous Aggregation of Multivariate GARCH Processes," *Journal of Econometrics*, 71, 71-87.
- Palm, F.C. and TH. Nijman (1984), "Missing Observations in the Dynamic Regression Model," *Econometrica*, 52, 1415-1435.
- Sentana, E. (1995), "Quadratic ARCH Models," *Review of Economic Studies*, 62, 639-661.
- Shephard, N. (1996), "Statistical Aspects of ARCH and Stochastic Volatility," in D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (Eds), *Time Series Models in Econometrics, Finance and Other Fields*, pp. 1-67, London, Chapman and Hall.
- Taylor, S.J. (1986), *Modeling Financial Time Series*, John Wiley.
- West, K. (1996), "Asymptotic Inference About Predictive Ability," *Econometrica*, 64, 1067-1084.

Table 1. GMM and QML estimation of various SR-SARV(1) models

Model	Method	$\mu^0$	$\omega^0$	$\alpha^0$	$\beta^0$	$\gamma^0$
<b>M1-M01</b>		0	0.4	.206	.594	.8
<b>M1</b>	<b>GMM1</b>	0.0001937 (0.005612)	0.5256 (0.1149)	0.1773 (0.04910)	0.5549 (0.07239)	0.7322 (0.05900)
	<b>GMM2</b>	0.0001937 (0.005612)	0.4064 (0.03587)	0.1773 (0.04910)	0.6170 (0.04667)	0.7942 (0.02015)
	<b>QMLE</b>	0.0004368 (0.005546)	0.4629 (0.02318)	0.1963 (0.006935)	0.5688 (0.01547)	0.7651 (0.01219)
<b>M01</b>	<b>GMM1</b>	0.0002450 (0.004660)	0.4150 (0.03230)	0.2080 (0.01960)	0.5850 (0.02870)	0.7920 (0.01620)
	<b>GMM2</b>	0.0002450 (0.004660)	0.4000 (0.01950)	0.2080 (0.01960)	0.5920 (0.02070)	0.8000 (0.01040)
	<b>MLE</b>	0.0001540 (0.004470)	0.4010 (0.01310)	0.2060 (0.004630)	0.5930 (0.008970)	0.8000 (0.006660)
<b>M2</b>		0	0.3	-	-	.85
<b>M2</b>	<b>GMM1</b>	-0.0003623 (0.005174)	0.5580 (0.1665)	0.2103 (0.08995)	0.4919 (0.1317)	0.7022 (0.08557)
	<b>GMM2</b>	-0.0003623 (0.005174)	0.3283 (0.05215)	0.2103 (0.08996)	0.6181 (0.08116)	0.8284 (0.02959)
	<b>QMLE</b>	0.0001183 (0.004887)	0.3906 (0.03870)	0.2899 (0.04298)	0.5059 (0.03156)	0.7958 (0.01959)
<b>M3-M03</b>		0	0.2	.281	.569	.85
<b>M3</b>	<b>GMM1</b>	-0.0007530 (0.004260)	0.2380 (0.06270)	0.2030 (0.06570)	0.6190 (0.09790)	0.8210 (0.04760)
	<b>GMM2</b>	-0.0007530 (0.004260)	0.2020 (0.02390)	0.2030 (0.06580)	0.6450 (0.06360)	0.8480 (0.01970)
	<b>QMLE</b>	-0.0006850 (0.003910)	0.2520 (0.01110)	0.1220 (0.004770)	0.6840 (0.01090)	0.8070 (0.008940)
<b>M03</b>	<b>GMM1</b>	-0.0005740 (0.003720)	0.2240 (0.03340)	0.2740 (0.03790)	0.5560 (0.05160)	0.8310 (0.02610)
	<b>GMM2</b>	-0.0005740 (0.003720)	0.2010 (0.008280)	0.2740 (0.03790)	0.5740 (0.03820)	0.8480 (0.007640)
	<b>MLE</b>	-0.0004090 (0.003130)	0.2010 (0.005300)	0.2800 (0.006140)	0.5690 (0.007470)	0.8490 (0.004790)

NOTE. The model **M1** (resp **M2**) corresponds to the temporally aggregated process  $\varepsilon_{tm}^{(m)} = \sum_{i=1}^m \varepsilon_{(t-1)m+i}$ , where  $\varepsilon_t$  is a Gaussian strong GARCH(1,1) with finite (resp infinite) fourth moment and  $m = 400$ ; the SR-SARV(1) coefficients at low frequency are  $(\mu^0, \omega^0, \gamma^0) = (0, .4, .8)$  (resp  $(0, .3, .85)$ ); the weak GARCH coefficients are  $(\mu^0, \omega^0, \alpha^0, \beta^0, \gamma^0) = (0, .4, .206, .594, .8)$  for **M1** and not defined for **M2** since the fourth moment is not finite. The model **M3** is defined as the cross-sectional aggregation of two independent strong Gaussian GARCH(1,1) processes with finite fourth moment and with the same persistence parameter ( $\gamma = \alpha + \beta$ ); the SR-SARV(1) coefficients are  $(0, .2, .85)$  and the weak GARCH ones are  $(0, .2, .281, .569, .85)$ . The Model **M01** (resp **M03**) is a Gaussian strong GARCH(1,1) which coefficients are the weak GARCH ones of **M1** (resp **M3**). GMM1 and GMM2 are the GMM estimators based on (8.3) described in the text. The sample size (after aggregation) is 80,000. The reported statistics are based on 100 replications. For each cell, the first number shows the mean and the second the standard deviation error (in parentheses).

**Table 2. QML estimation of temporally aggregated GARCH models**

Model	Method	$\mu^0$	$\omega^0$	$\alpha^0$	$\beta^0$	$\gamma^0$
M1		0	0.4	.206	.594	.8
M1	QMLE	0.000246 (0.003856)	0.4618 (0.01238)	0.1959 (0.004394)	0.5695 (0.008015)	0.7654 (0.006655)
M2		0	0.3	-	-	.85
M2	QMLE	.00004578 (0.003367)	0.3733 (0.01197)	0.3078 (0.009687)	0.4959 (0.01158)	0.8037 (0.007555)

NOTE. See Table 1 for the characteristics of the models. The sample size is 150,000. The statistics are based on 100 replications. Each cell contains the mean and the standard deviation error (in parentheses).

**Table 3. R2 of the regression of the variance onto the linear variance**

Model	Method	in sample	out of sample
<b>M1: <math>R^2=.580</math></b>	<b>GMM1</b>	0.3623 (0.06792)	0.3050 (0.05249)
	<b>GMM2</b>	0.3605 (0.06025)	0.3043 (0.05304)
	<b>QMLE</b>	0.3643 (0.06711)	0.3068 (0.05437)
	<b>True Parameter</b>	0.3646 (0.07027)	0.3175 (0.1044)
<b>M01: <math>R^2=1</math></b>	<b>GMM1</b>	0.9980 (0.004390)	0.6060 (0.05450)
	<b>GMM2</b>	0.9990 (0.002550)	0.6150 (0.05010)
	<b>MLE</b>	1.000 (0.0002360)	0.6160 (0.04900)
	<b>True Parameter</b>	1.000 (0.00002319)	0.99989 (.001702)
<b>M3: <math>R^2=.902</math></b>	<b>GMM1</b>	0.7560 (0.06720)	0.4830 (0.1470)
	<b>GMM2</b>	0.7630 (0.06910)	0.5040 (0.1430)
	<b>QMLE</b>	0.7590 (0.06670)	0.5310 (0.1340)
	<b>True Parameter</b>	0.7620 (0.0771)	0.6480 (0.1340)
<b>M03: <math>R^2=1</math></b>	<b>GMM1</b>	0.9960 (0.008150)	0.6490 (0.06990)
	<b>GMM2</b>	0.9980 (0.004610)	0.6670 (0.06410)
	<b>MLE</b>	1.000 (0.0001220)	0.6630 (0.05150)
	<b>True Parameter</b>	1.000 (0.0000153)	0.9992 (0.002558)

NOTE. See Table 1 for the characteristics of the models. For each model, we regress the true variance  $f_t$  onto the constant and the linear variance  $h_{t+1}$  which is the weak GARCH variance and compute the corresponding  $R^2$ . For each model we give the theoretical  $R^2$ . The linear variance is computed by using the three estimators  $\hat{\theta}^{GMM1}$ ,  $\hat{\theta}^{GMM2}$  and  $\hat{\theta}^{QMLE}$  and by the true parameter  $\theta^0$ . The in-sample size is 80,000 while the out-of-sample one is 2,000. The statistics are based on 100 replications. Each cell contains the mean and the standard deviation error (in parentheses).

## APPENDIX

**Lemma A.1.** Consider the processes  $z_t, G_t, V_t, \eta_t$  defined as in definition 2.2 (state-space representation) and define  $J_t$  as the sigma algebra  $J_t = \sigma(z_\tau, G_\tau, m_\tau, \tau \leq t)$ . Then

$$G_{t+i} = \left( \sum_{j=0}^{i-1} \Gamma^j \right) \Omega + \Gamma^i G_t + \sum_{j=0}^{i-1} \Gamma^j V_{t+i-j}, \quad \text{with} \quad (\text{A.1})$$

$$E[\eta_{t+i} | J_t] = 0, \quad E[V_{t+i} | J_t] = 0 \quad \text{and} \quad (\text{A.2})$$

$$E[G_{t+i} | J_t] = \left( \sum_{j=0}^{i-1} \Gamma^j \right) \Omega + \Gamma^i G_t. \quad (\text{A.3})$$

**Proof.** By recursive calculus, (2.3) implies (A.1). On the other hand, (2.4) (resp (2.5)) implies that  $\forall i > 0$ ,  $E[\eta_{t+i} | J_t] = 0$  (resp  $E[V_{t+i} | J_t] = 0$ ), that is (A.2) and, hence, (A.3).  $\square$

**Lemma A.2.** Consider  $\{z_t\}$  a univariate strong AR(1)  $z_t = \omega + \gamma z_{t-1} + \eta_t$  with  $\eta_t$  i.i.d. and assume that, for  $n \in \mathbf{N}$ ,  $E[|\eta_t|^n]$  is finite. Then  $\forall i \leq n$   $Z_{i,t} = (z_t, z_t^2, \dots, z_t^i)'$  is a semi-strong VAR(1).

**Proof.** Let  $i \leq n$ . Then  $z_t^i = \sum_{j=0}^i \binom{i}{j} \gamma^j z_{t-1}^j (\omega + \eta_t)^{i-j}$ . Define  $a_{i,j} = E[(\omega + \eta_t)^{i-j}]$ . Then  $E[z_t^i | z_\tau, \tau \leq t-1] = \sum_{j=0}^i a_{i,j} \binom{i}{j} \gamma^j z_{t-1}^j$ . Hence  $Z_{i,t}$  is a semi-strong VAR(1). Note that the autoregressive matrix is lower triangular and that the diagonal coefficient are  $(\gamma^1, \dots, \gamma^i)$ .  $\square$

**Proof of Proposition 3.1.** From (3.3), we have  $dy_t = \sqrt{e' F_t^c} dW_{1t}$  where  $W_{1t}$  is the first component of  $W_t$ . Therefore  $\varepsilon_{th}^{(h)} = \int_{th-h}^{th} \sqrt{e' F_u^c} dW_{1u}$  and  $f_{(t-1)h}^{(h)} \equiv \text{Var}[\varepsilon_{th}^{(h)} | J_{(t-1)h}^{(h)}] = E[\int_{th-h}^{th} e' F_u^c du | J_{(t-1)h}^{(h)}] = e' \int_{th-h}^{th} E[F_u^c | J_{(t-1)h}^{(h)}] du$ .

Consider the equation (3.3), then we have  $dF_t^c = K(\Theta - F_t^c)dt + M_{22} R_t dW_t$  where  $M_{22}$  is the  $p \times (p+1)$  matrix defined by  $M_{22} = (0, I_p)$ . Therefore, the continuous time version of (A.1) is

$$\forall h > 0, F_{t+h}^c = (Id - e^{-Kh})\Theta + e^{-Kh} F_t^c + e^{-Kh} \int_t^{t+h} e^{K(u-t)} M_{22} R_t dW_u. \quad (\text{A.4})$$

Hence, as in the proof of Proposition 2.3,  $f_{th-h}^{(h)} = e' F_{th-h}^{(h)}$  with  $F_{th-h}^{(h)} = \int_{th-h}^t \left\{ (Id - e^{-K[u-(th-h)]})\Theta + e^{-K[u-(th-h)]} F_{th-h}^c \right\} du = A^{(h)} F_{th-h}^c + B^{(h)}$  where  $A^{(h)} = K^{-1}(Id - e^{-Kh})$  and  $B^{(h)} = (hId - A^{(h)})\Theta$ . Since  $\{F_{th}^c, t \in \mathbf{Z}\}$  is a VAR(1) due to (A.4) and since  $A^{(h)}$  is non singular,  $\{F_{th}^{(h)}, t \in \mathbf{Z}\}$  is also a VAR(1) with the same autoregressive matrix than  $\{F_{th}^c\}$  that is  $e^{-Kh}$ .  $\square$

**Lemma A.3. Positive OU SV Model is a SR-SARV**  $\{x_t^*, \sigma_t\}$  the stationary solutions of  $dx_t = \sigma_t dw_t$  and  $\sigma_t^2 = e^{-\lambda t} \sigma_0^2 + \int_0^t e^{-\lambda(t-s)} dz(\lambda s)$  where  $\sigma_0^2 = \int_{-\infty}^0 e^s dz(s)$ ,  $\{z_t, t \in \mathbb{R}\}$  is an integrable homogenous Lévy process and  $\lambda$  a positive number. Then  $\{\varepsilon_t, t \in \mathbf{Z}\}$  defined by  $\varepsilon_t = x_t^* - x_{t-1}^*$  is a SR-SARV w.r.t.  $J_t = \sigma(\varepsilon_\tau, \sigma_\tau, \tau \leq t)$ .

**Proof.** We have  $\varepsilon_t = \int_{t-1}^t \sigma_u dW_u$ . Hence  $E[\varepsilon_t | J_{t-1}] = 0$ . Since  $\sigma_t^2 = e^{-\lambda t} \sigma_0^2 + \int_0^t e^{-\lambda(t-s)} dz(\lambda s)$ , we deduce that  $\sigma_{t+u}^2 = e^{-\lambda u} \sigma_t^2 + \int_t^{t+u} e^{-\lambda(t+u-s)} dz(\lambda s)$ , which is exactly the univariate version of

(A.4). Therefore, following the Proof of Proposition 3.1, we can show that  $\varepsilon_t$  is a SR-SARV(1) w.r.t.

$$J_t = \sigma(\varepsilon_\tau, \sigma_\tau, \tau \leq t). \square$$

**Proof of Proposition 3.2.** We show that the points of the Definition 3.1. are fulfilled: i) by definition of  $J_{tm}^{(m)}$ ; ii) we have  $J_{tm}^{(m)} \subset J_{tm}$ . Hence,  $E[\varepsilon_{tm}^{(m)} | J_{tm-m}^{(m)}] = \sum_{i=0}^{m-1} a_i E[E[\varepsilon_{tm-i} | J_{tm-i-1} | J_{tm-m}^{(m)}]] = 0$ , that is  $\varepsilon_{tm}^{(m)}$  is a m.d.s. w.r.t.  $J_{tm-m}^{(m)}$ ; iii) we have:  $Var[\varepsilon_{tm}^{(m)} | J_{tm-m}^{(m)}] = E[(\varepsilon_{tm}^{(m)})^2 | J_{tm-m}^{(m)}] = \sum_{i=0}^{m-1} a_i^2 E[\varepsilon_{tm-i}^2 | J_{tm-m}^{(m)}] + 2 \sum_{0 \leq i < j \leq m-1} a_i a_j E[\varepsilon_{tm-i} \varepsilon_{tm-j} | J_{tm-m}^{(m)}]$   
 $= \sum_{i=0}^{m-1} a_i^2 E[E[\varepsilon_{tm-i}^2 | J_{tm-i-1} | J_{tm-m}^{(m)}]] + 2 \sum_{0 \leq i < j \leq m-1} a_i a_j E[\varepsilon_{tm-j} E[\varepsilon_{tm-i} | J_{tm-i-1} | J_{tm-m}^{(m)}]]$   
 $= E[\sum_{i=0}^{m-1} a_i^2 f_{tm-i-1} | J_{tm-m}^{(m)}]$ . This is a conditional expectation of an aggregation of a state space model. Following the proof of Proposition 2.3, we have  $Var[\varepsilon_{tm}^{(m)} | J_{tm-m}^{(m)}] = E[e'(A^{(m)} F_{tm-m} + B^{(m)}) | J_{tm-m}^{(m)}]$  where  $A^{(m)}$  and  $B^{(m)}$  are defined by (3.7). By definition of  $J_{tm-m}^{(m)}$ ,  $F_{tm-m}^{(m)}$  is adapted w.r.t.  $J_{tm-m}^{(m)}$ . Hence,  $Var[\varepsilon_{tm}^{(m)} | J_{tm-m}^{(m)}] = e'(A^{(m)} F_{tm-m} + B^{(m)}) = e^{(m)'} F_{tm-m}^{(m)}$  where  $e^{(m)}$  and  $F_{tm-m}^{(m)}$  are defined by (3.8). As in Proposition 2.3,  $F_{tm-m}^{(m)}$  is a VAR(1) with autoregressive matrix  $\Gamma^m$ .  $\square$

**Proof of Proposition 3.3.** We follow the proof of Proposition 2.2. of Meddahi and Renault (2000a). Consider  $\{\varepsilon_t, t \in \mathbf{Z}\}$  a SR-SARV(p). Hence  $\varepsilon_t^2 = f_{t-1} + \eta_t$ . where  $\{f_t\}$  admits a state-space representation  $\{F_t, \eta_t\}$  w.r.t  $J_t$ . We have  $F_t = \Omega + \Gamma F_{t-1} + V_t \Rightarrow (Id - \Gamma L)F_t = \Omega + V_t \Rightarrow Det(Id - \Gamma L)F_t = (Id - \Gamma L)^*(\Omega + V_t)$  where  $L$  is the Lag Operator,  $Det(\cdot)$  is the determinant function and  $(Id - \Gamma L)^*$  is the adjoint matrix of  $(Id - \Gamma L)$ . Hence :  $Det(Id - \Gamma L)f_t = Det(Id - \Gamma L)e'F_t = e'(Id - \Gamma)^*\Omega + e'(Id - \Gamma L)^*V_t$  We have:  $Deg(e'(Id - \Gamma L)^*) \leq p - 1$  where  $Deg(\cdot)$  is the maximal degree of the lag polynomials, coefficients of the matrix. Hence  $E[Det(Id - \Gamma L)f_t - e'(Id - \Gamma)^*\Omega | J_{t-p}] = 0$ . Thus  $E[Det(Id - \Gamma L)\varepsilon_{t+1}^2 - e'(Id - \Gamma)^*\Omega | J_{t-p}] = 0$  since  $\varepsilon_{t+1}^2 = f_t + \eta_{t+1}$  and the (maximal) degree of  $Det(Id - \Gamma L)$  is p. Define  $a_1, \dots, a_p$  by  $1 - \sum_{i=1}^p a_i L^i = Det(Id - \Gamma L)$  and the real  $\omega$  by  $\omega = e'(Id - \Gamma)^*\Omega$ . By definition 3.1, the eigenvalues of  $\Gamma$  are smaller than one in modulus. Therefore the roots of  $1 - \sum_{i=1}^p a_i L^i$  are outside the unit circle. Finally,  $\sigma(\varepsilon_\tau, \tau \leq t - p) \subset J_{t-p}$ . Hence  $E[\varepsilon_{t+1}^2 - \omega - \sum_{i=1}^p a_i \varepsilon_{t+1-i}^2 | z_\tau, \tau \leq t - p] = 0$ , that is (3.10).

Conversely, consider a process  $\varepsilon_t$  such that (3.10). Define  $F_{t-1}$  by  $F_{t-1} = (E[\varepsilon_{t+p-1}^2 | I_{t-1}], E[\varepsilon_{t+p-2}^2 | I_{t-1}], \dots, E[\varepsilon_t^2 | I_{t-1}])'$ . Thus  $\varepsilon_t^2 = (0, 0, \dots, 0, 1)F_{t-1} + v_t$  with  $E[v_t | I_{t-1}] = 0$ .

For  $i = 2, \dots, p$ , we have again  $E[F_t(i) | I_{t-1}] = E[\varepsilon_{t+p+1-i}^2 | I_{t-1}] = F_{t-1}(i - 1)$ .

$E[F_t(1) | I_{t-1}] = E[\varepsilon_{t+p}^2 | I_{t-1}] = E[(\varepsilon_{t+p}^2 - \omega - \sum_{i=1}^p a_i \varepsilon_{t+p-i}^2) + \omega + \sum_{i=1}^p a_i \varepsilon_{t+p-i}^2 | I_{t-1}]$   
 $= \omega + \sum_{i=1}^p a_i E[\varepsilon_{t+p-i}^2 | I_{t-1}] = \omega + \sum_{i=1}^p a_i F_{t-1}(i)$ . Hence,  $E[F_t | I_{t-1}] = \Omega + \Gamma F_{t-1}$ . As a conclusion,  $\{\varepsilon_t^2\}$  has a state space representation  $\{F_t, v_t\}$  w.r.t.  $I_t$ . On the other hand,  $\varepsilon$  is a m.d.s. w.r.t.  $I_t$ .

Thus,  $\varepsilon_t$  is a SR-SARV(p) w.r.t.  $I_t$ .  $\square$

**Proof of Proposition 3.4.** Let  $\varepsilon_t$  be a semi-strong GARCH(p,q) defined by (3.11) and (3.12). Then  $(B(L) - A(L))\varepsilon_t^2 = \omega + B(L)\eta_t$  with  $\eta_t = \varepsilon_t^2 - h_t$ . By assumption, the roots  $B(L) - A(L)$  and  $B(L)$  are not common and are outside the unit circle. Finally,  $\eta_t$  is a m.d.s. w.r.t.  $I_t$  since

$E[\eta_t | I_{t-1}] = E[\varepsilon_t^2 | I_{t-1}] - h_t = 0$ . Conversely, consider a m.d.s.  $\{\varepsilon_t\}$  such that  $Q(L)\varepsilon_t^2 = \omega + P(L)\eta_t$  where  $Q(L) = 1 - \sum_{i=1}^q a_i L^i$ ,  $P(L) = 1 - \sum_{i=1}^p b_i L^i$ ,  $a_q \neq 0$ ,  $b_p \neq 0$  and  $p \leq q$ . Assume that  $E[\eta_t | I_{t-1}] = 0$ . Define  $h_t$  by  $h_t \equiv E[\varepsilon_t^2 | I_{t-1}]$ . Hence  $h_t = \omega + (1 - Q(L))\varepsilon_t^2 + (P(L) - 1)\eta_t$  and  $\eta_t = \varepsilon_t^2 - h_t$ . Thus  $h_t = \omega + (P(L) - Q(L))\varepsilon_t^2 + (1 - P(L))h_t$  and  $P(L)h_t = \omega + (P(L) - Q(L))\varepsilon_t^2$ . By assumption, the roots of  $P(L)$  and  $P(L) - (Q(L) - P(L))$ , i.e.  $Q(L)$ , are not common and outside the unit circle. Define  $u_t$  by  $u_t \equiv \varepsilon_t/\sqrt{h_t}$ . We have  $E[u_t | I_{t-1}] = 0$  since  $\varepsilon_t$  is a m.d.s.; moreover,  $Var[u_t | I_{t-1}] = E[\varepsilon_t^2 | I_{t-1}]/h_t = 1$ , i.e. (3.12).  $\square$

**Proof of Proposition 3.5.** Since  $H_t^s \subset H_t$ ,  $EL[h_t | H_{t-1}^s] = h_t^s$ . But  $h_t = \omega/B(1) + B(L)^{-1}A(L)\varepsilon_t^2$  and hence  $h_t \in H_{t-1}^s$ . Thus  $h_t = h_t^s$ . Therefore  $\varepsilon_t^2$  is a weak ARMA (since  $B(L)h_t^s = \omega + A(L)\varepsilon_t^2$ ) and  $cov(\eta_t, \varepsilon_\tau) = 0 \forall \tau < t$ .

Conversely, assume that  $\varepsilon_t^2$  is a weak ARMA and (3.15). We have:  $h_t = EL[\varepsilon_t^2 | H_{t-1}] = h_t^s + EL[\eta_t | H_{t-1}]$ . By definition of  $\eta_t$ ,  $\forall \tau < t$ ,  $cov(\varepsilon_\tau^2, \eta_t) = 0$ . Therefore, by combination with (3.15),  $\forall z \in H_{t-1}$ ,  $cov(z, \eta_t) = 0$ . Thus  $EL[\eta_t | H_{t-1}] = 0$  and  $h_t = h_t^s$  and  $\varepsilon_t$  is a weak GARCH.  $\square$

**Proof of Proposition 4.1.** Let us consider  $\varepsilon_t$  a GARCH(1,1). Let  $f_{t-1} = h_t = E[\varepsilon_t^2 | I_{t-1}]$  and  $u_t = \frac{\varepsilon_t}{\sqrt{h_t}}$ . By definition,  $E[u_t | I_{t-1}] = 0$  and  $E[u_t^2 | I_{t-1}] = 1$ . while  $f_t$  is an  $I_t$ -adapted AR(1) process. with an innovation process:  $\nu_t = \alpha f_{t-1}(u_t^2 - 1)$ . Then, given  $I_{t-1}$ ,  $\varepsilon_t^2$  and  $\nu_t = \alpha f_{t-1}(\frac{\varepsilon_t^2}{f_{t-1}} - 1)$  are conditionally perfectly positively correlated (since  $\alpha > 0$ ). Thus, this is also the case for  $\varepsilon_t^2$  and  $f_t = \omega + \gamma f_{t-1} + \nu_t$ . Moreover:  $Var[f_t | J_{t-1}] = Var[\nu_t | J_{t-1}] = \alpha^2 Var[\varepsilon_t^2 | J_{t-1}]$  with  $\alpha^2 \leq \gamma^2 = (\alpha + \beta)^2$  since  $\beta \geq 0$ .

Conversely, let us now consider a SR-SARV(1) process  $\varepsilon_t$  which fulfills the two restrictions of Proposition 4.1. By the first restriction, we know that:  $f_t = a_t \varepsilon_t^2 + b_t$ ,  $a_t, b_t \in J_{t-1}$ , with  $(Var[f_t | J_{t-1}])^{1/2} = a_t (Var[\varepsilon_t^2 | J_{t-1}])^{1/2}$ .

Thus, by the second restriction, we know that  $a_t$  is a positive constant  $\alpha$  smaller or equal to  $\gamma$ . Therefore:  $f_t = \alpha \varepsilon_t^2 + b_t$  and  $E[f_t | J_{t-1}] = \alpha f_{t-1} + b_t$ . By identification with the AR(1) representation of  $f_t$ , we conclude that:  $b_t = \omega + \beta f_{t-1}$  where  $\beta = \gamma - \alpha \geq 0$ . Thus:  $f_t = \omega + \alpha \varepsilon_t^2 + \beta f_{t-1}$ , which proves that  $f_t$  is also  $I_t$ -adapted (see  $0 < \alpha \leq \gamma < 1$ ). Then we know by Proposition 2.1 that  $\varepsilon_t$  is also a SR-SARV(1) process w.r.t.  $I_t$  and  $f_t = Var[\varepsilon_{t+1} | I_t]$ . Therefore, with:  $h_t = f_{t-1} = Var[\varepsilon_t | I_{t-1}]$  we do get the GARCH(1,1) representation:  $h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$ .  $\square$

**Proof of Proposition 4.2.** Define  $u_t$  as the standardized residuals ( $u_t = \varepsilon_t/\sqrt{h_t}$ ). Straightforward calculus show that all the models can be rewritten as  $h_t = \omega^* + \gamma^* h_{t-1} + \nu_{t-1}$  with:

GJR:  $\omega^* = \omega$ ,  $\gamma^* = \alpha + \beta + \gamma S$ ,  $\nu_{t-1} = \alpha(\varepsilon_{t-1}^2 - h_{t-1}) + \gamma(S_{t-1}\varepsilon_{t-1}^2 - S h_{t-1})$ , where  $S = E[S_t u_t^2 | I_{t-1}]$ .

Asymmetric GARCH:  $\omega^* = \omega + \alpha \gamma^2$ ,  $\gamma^* = \alpha + \beta$ ,  $\nu_{t-1} = \alpha(\varepsilon_{t-1}^2 - h_{t-1}) + 2\alpha \gamma \varepsilon_{t-1}$ .

Nonlinear Asymmetric GARCH:  $\omega^* = \omega$ ,  $\gamma^* = \alpha(1 + \gamma^2) + \beta$ ,  $\nu_{t-1} = \alpha h_{t-1}(u_{t-1}^2 - 1 + 2\gamma u_{t-1})$ .

VGARCH:  $\omega^* = \omega + \alpha(1 + \gamma^2)$ ,  $\gamma^* = \beta$ ,  $\nu_{t-1} = \alpha(u_{t-1}^2 - 1 + 2\gamma u_{t-1})$ .

Heston-Nandi:  $\omega^* = \omega + \alpha$ ,  $\gamma^* = \alpha\gamma^2 + \beta$ ,  $\nu_{t-1} = \alpha(u_{t-1}^2 - 1 - 2\gamma\varepsilon_{t-1})$ .

By the restrictions  $E[\varepsilon_{t-1} | I_{t-2}] = E[u_{t-1} | I_{t-2}] = 0$ ,  $E[\varepsilon_{t-1}^2 | I_{t-2}] = h_{t-1}$  and  $E[u_{t-1}^2 | I_{t-2}] = 1$ , we have  $E[\nu_{t-1} | I_{t-2}] = 0$ , that is  $\varepsilon_t$  is a SR-SARV(1).  $\square$

**Proof of the equivalence of the two conditions of (4.6) and (4.7).** i) We have:  $E[\varepsilon_t \varepsilon_{t+1}^2 | J_{t-1}] = \sqrt{f_{t-1}} E[u_t E[\varepsilon_{t+1}^2 | J_t] | J_{t-1}] = \sqrt{f_{t-1}} E[u_t f_t | J_{t-1}] = \sqrt{f_{t-1}} E[u_t(\omega + \gamma f_{t-1} + \nu_t) | J_{t-1}] = \sqrt{f_{t-1}} E[u_t \nu_t | J_{t-1}]$ . Hence  $E[u_t \nu_t | J_{t-1}] = 0 \iff E[\varepsilon_t \varepsilon_{t+1}^2 | J_{t-1}] = 0$  since  $f_t \neq 0$  almost surely.

ii) We have:  $E[\varepsilon_t^3 | J_{t-1}] = (f_{t-1})^{3/2} E[u_t^3 | J_{t-1}]$ . Hence  $E[u_t^3 | J_{t-1}] = 0 \iff E[\varepsilon_t^3 | J_{t-1}] = 0$ .  $\square$

**Proof of Proposition 4.3.** The SR-SARV(1) property implies, by Proposition 3.3, that  $\varepsilon_t^2$  fulfill the multiperiod restrictions (3.10) with  $p = 1$ . Define  $\omega_t$  by  $\omega_t = \varepsilon_t^2 - \omega - \gamma\varepsilon_t^2$ . We have  $E[\omega_t | I_{t-2}] = 0$  and  $\omega_t$  is a square integrable processe since  $\varepsilon_t$  has a finite fourth moment. Therefore  $\omega_t$  is a weak MA(1) and hence  $\varepsilon_t^2$  is a weak ARMA(1,1). Therefore, by Proposition 3.5,  $\varepsilon_t$  is a weak GARCH(1,1) if and only if (3.15) is fulfilled. But, since by the ARMA representation of  $\varepsilon_t^2$ , the Hilbert space  $H_t^s$  coincides with the Hilbert space spanned by 1,  $\eta_\tau$ ,  $\tau \leq t$ , the condition (3.15) is implied by the following symmetry property of the process  $\varepsilon$ :  $Cov(\varepsilon_{t'}, \varepsilon_t^2) = 0 \forall t, t'$  that is  $E(\varepsilon_{t'} \varepsilon_t^2) = 0 \forall t, t'$ . Thus, we are going to prove this symmetry property. Indeed, we will prove the stronger result (which will be useful in the following):

$$E[\varepsilon_{t'} \varepsilon_t^2 | J_\tau] = 0 \forall t, t' \text{ and } \tau = \text{Min}(t, t') - 1 \quad (\text{A.5})$$

If  $t' > t$ , then  $E[\varepsilon_{t'} \varepsilon_t^2 | J_{t-1}] = E[\varepsilon_t^2 E[\varepsilon_{t'} | J_{t'-1}] | J_{t-1}] = 0$  since  $\varepsilon_{t'}$  is an m.d.s. w.r.t.  $J_{t'-1}$ .

If  $t' = t$ , then  $E[\varepsilon_{t'} \varepsilon_t^2 | J_{t-1}] = E[\varepsilon_t^3 | J_{t-1}] = f_{t-1}^{3/2} E[u_t^3 | J_{t-1}] = 0$  by (4.7).

If  $t' < t$ , then  $E[\varepsilon_{t'} \varepsilon_t^2 | J_{t'-1}] = E[\varepsilon_{t'} f_{t-1} E[u_t^2 | J_{t-1}] | J_{t'-1}] = E[\varepsilon_{t'} f_{t-1} | J_{t'-1}]$ . Since  $f_t$  is an AR(1), we have  $f_{t-1} = \sum_{i=0}^{\infty} \gamma^i \nu_{t-1-i} + E[f_{t-1}]$ . Hence  $E[\varepsilon_{t'} f_{t-1} | J_{t'-1}] = \sum_{i=0}^{\infty} \gamma^i E[\nu_{t-1-i} \varepsilon_{t'} | J_{t'-1}]$ . But: if  $i \geq t - t'$ , then  $E[\nu_{t-1-i} \varepsilon_{t'} | J_{t'-1}] = \nu_{t-1-i} E[\varepsilon_{t'} | J_{t'-1}] = 0$ ; if  $i = t - t' - 1$ , then  $E[\nu_{t-1-i} \varepsilon_{t'} | J_{t'-1}] = E[\nu_{t'} \varepsilon_{t'} | J_{t'-1}] = \sqrt{f_{t'-1}} E[u_{t'} \nu_{t'} | J_{t'-1}] = 0$  by (4.6); finally, if  $i < t - t' - 1$ , then  $E[\nu_{t-1-i} \varepsilon_{t'} | J_{t'-1}] = E[\varepsilon_{t'} E[\nu_{t-1-i} | J_{t-i-2}] | J_{t'-1}] = 0$  since  $\nu_t$  is an m.d.s. w.r.t.  $J_t$ . Hence,  $E[\varepsilon_{t'} f_{t-1} | J_{t'-1}] = 0$ , which achieves the proof of Proposition 4.3.  $\square$

**Proof of Proposition 4.4.**  $E[u_{tm}^{(m)} \nu_{tm}^{(m)} | J_{tm-m}^{(m)}] = \frac{a^{(m)}}{\sqrt{f_{tm-m}^{(m)}}} E[\sum_{0 \leq i, j \leq m-1} a_{im} \gamma^j \varepsilon_{tm-i} \nu_{tm-j} | J_{tm-m}^{(m)}]$ . But (see third case of the proof of Proposition 4.3), (4.6) implies that  $E[\varepsilon_{tm-i} \nu_{tm-j} | J_{tm-m}^{(m)}] = 0$  for  $i, j = 0, 1, \dots, m-1$ . Thus,  $E[\varepsilon_{tm-i} \nu_{tm-j} | J_{tm-m}^{(m)}] = 0$  and hence  $E[u_{tm}^{(m)} \nu_{tm}^{(m)} | J_{tm-m}^{(m)}] = 0$ , i.e. (4.8).

$E[(u_{tm}^{(m)})^3 | J_{tm-m}^{(m)}] = \frac{1}{(f_{tm-m}^{(m)})^{3/2}} E[\sum_{0 \leq i, j, k \leq m-1} a_{im} a_{jm} a_{km} \varepsilon_{tm-i} \varepsilon_{tm-j} \varepsilon_{tm-k} | J_{tm-m}^{(m)}]$ . Let  $(i, j, k)$  as  $i \leq j \leq k \leq m-1$ . If  $i < j \leq k$ , then  $E[\varepsilon_{tm-i} \varepsilon_{tm-j} \varepsilon_{tm-k} | J_{tm-m}^{(m)}] = E[\varepsilon_{tm-j} \varepsilon_{tm-k} E[\varepsilon_{tm-i} | J_{tm-i-1} | J_{tm-m}^{(m)}]] = 0$ . If  $i = j = k$ , then  $E[\varepsilon_{tm-i} \varepsilon_{tm-j} \varepsilon_{tm-k} | J_{tm-m}^{(m)}] = E[(f_{tm-i-1})^{3/2} E[(u_{tm-i})^3 | J_{tm-i-1}]] = 0$ .

$J_{tm-i-1} \mid J_{tm-m}^{(m)} = 0$ . If  $i = j < k$ , then  $E[\varepsilon_{tm-i}\varepsilon_{tm-j}\varepsilon_{tm-k} \mid J_{tm-m}^{(m)}] = E[E[\varepsilon_{tm-k}(\varepsilon_{tm-i})^2 \mid J_{tm-m}^{(m)}] \mid J_{tm-m}^{(m)}] = 0$  by (A.5). So we have:  $E[(u_{tm}^{(m)})^3 \mid J_{tm-m}^{(m)}] = 0$ .  $\square$

**Proof of Proposition 4.5.** The second part of (4.6) implies (4.10). The second part of (4.7) implies (4.11).  $\square$

**Proof of Proposition 5.1.** This is exactly the same proof of Proposition 3.2 by taking  $\Gamma = 1$ .

**Proof of Proposition 6.1.** All the results of the proof of Proposition 3.2. still hold, in particular the restrictions i) and ii) of the definition of a SR-SARV( $\infty$ ). To achieve the proof, we have to show that  $f_{tm-m}^{(m)} = \text{Var}[\varepsilon_{tm}^{(m)} \mid J_{tm-m}^{(m)}]$  follows the restriction iii). Define  $\tilde{e} = (e_1, \dots, e_q)$ . Thus  $f_t = \tilde{e}'_q F_{q,t} + \sum_{j=q+1}^{\infty} e_j f_{j,t}$ . Then, following the Proof of Proposition 3.2, we have:

$$f_{tm-m}^{(m)} = E[\sum_{i=0}^{m-1} a_i^2 f_{tm-i-1} \mid J_{tm-m}^{(m)}] = \sum_{i=0}^{m-1} a_i^2 E[f_{tm-i-1} \mid J_{tm-m}^{(m)}].$$
 On the other hand,

$$E[f_{tm-i-1} \mid J_{tm-m}^{(m)}] = \tilde{e}'_q E[F_{q,tm-i-1} \mid J_{tm-m}^{(m)}] + \sum_{j=q+1}^{\infty} e_j E[f_{j,tm-i-1} \mid J_{tm-m}^{(m)}]$$

$$= \tilde{e}'_q (\sum_{k=0}^{m-i-2} \Omega + \Gamma_q^{m-i-1} F_{q,tm-m}) + \sum_{j=q+1}^{\infty} e_j \gamma_j^{m-i-1} f_{j,tm-m}.$$
 As a consequence,

$$f_{tm-m}^{(m)} = \sum_{i=0}^{m-1} a_i^2 \tilde{e}'_q (\sum_{k=0}^{m-i-2} \Omega + \Gamma_q^{m-i-1} F_{q,tm-m}) + \sum_{i=0}^{m-1} a_i^2 (\sum_{j=q+1}^{\infty} e_j \gamma_j^{m-i-1} f_{j,tm-m})$$

$$= \sum_{i=0}^{m-1} a_i^2 \tilde{e}'_q (\sum_{k=0}^{m-i-2} \Omega + \Gamma_q^{m-i-1} F_{q,tm-m}) + \sum_{j=q+1}^{\infty} \tilde{e}_j f_{j,tm-m} \text{ with } \tilde{e}_j = e_j \sum_{i=0}^{m-1} a_i^2 \gamma_j^{m-i-1}.$$

From the proof of Proposition 3.2, we know that  $\sum_{i=0}^{m-1} a_i^2 \tilde{e}'_q (\sum_{k=0}^{m-i-2} \Omega + \Gamma_q^{m-i-1} F_{q,tm-m})$  is a stationary VAR(1). It is obviously not correlated with any  $f_{j,t}$ ,  $j > q$ .

For  $j > q$ ,  $|\tilde{e}_j| \leq (\sum_{i=0}^{m-1} a_i^2) |e_j|$ . Hence, if  $\sum_{j=0}^{\infty} |e_j| < \infty$  (resp  $\sum_{j=0}^{\infty} |e_j|^2 < \infty$ ), then  $\sum_{j=0}^{\infty} |\tilde{e}_j| < \infty$  (resp  $\sum_{j=0}^{\infty} |\tilde{e}_j|^2 < \infty$ ). As a conclusion, the SR-SARV( $\infty$ ) class is closed under temporal aggregation.  $\square$

**Proof of Proposition 6.2.** We have  $f_{t-1} = \text{Var}[\varepsilon_t \mid J_{t-1}] = \sum_{i=1}^{\infty} f_{i,t-1}$ . By construction, for any  $i$ ,  $f_{i,t-1}$  is a semi-strong AR(1). Moreover, these processes are uncorrelated. By Parseval identity,  $\sum_{i=0}^{\infty} \psi_i^{*2} = E[\exp(2g_{t-1})]$ . Hence  $\sum_{i=1}^{\infty} e_i^2 < \infty$ . On the other hand,  $\forall i > 1$ ,  $E[H_i^{*2} (\frac{g_{t-1} - \mu}{\sigma})] = 1$ . Thus  $\sup_n E f_{n,t}^2 < +\infty$ . As a conclusion,  $\varepsilon_t$  is a SR-SARV( $\infty$ ) w.r.t.  $J_t$ .  $\square$

**Lemma A.4.** Consider two independent SR-SARV(1) processes,  $\{\varepsilon_{i,t}\}$ ,  $i = 1, 2$ , with the corresponding information sets  $J_{i,t}$  and conditional variance processes  $f_{i,t} = \omega_i + \gamma f_{i,t-1} + v_{i,t}$ . Then  $\varepsilon_t \equiv \varepsilon_{1,t} + \varepsilon_{2,t}$  is a SR-SARV(1) processe w.r.t.  $J_t = \sigma(\varepsilon_{\tau}, f_{1,\tau}, f_{2,\tau}, \tau \leq t)$ .

**Proof.** Obviously  $\varepsilon_t$  is  $J_t$ -adapted and a m.d.s. Let  $f_{t-1} \equiv \text{Var}[\varepsilon_t \mid J_{t-1}]$ . We have  $f_{t-1} = f_{1,t} + f_{2,t}$ . Thus,  $(1 - \gamma L)f_{t-1} = \omega_1 + \omega_2 + v_{1,t} + v_{2,t}$ , that is  $f_{t-1}$  is an AR(1) and, hence,  $\varepsilon_t$  is a SR-SARV(1) process w.r.t.  $J_t$ .  $\square$

**Lemma A.5.** Consider  $\varepsilon_t$  an SR-SARV(1) w.r.t.  $J_t$  defined by  $\varepsilon_t = \sqrt{f_{t-1}} u_t$ ,  $f_t = \omega + \gamma f_{t-1} + v_t$  where  $E[u_t \mid J_{t-1}] = E[v_t \mid J_{t-1}] = 0$  and  $E[u_t^2 \mid J_{t-1}] = 1$ . Assume that the fourth moment of  $\varepsilon_t$  is finite. Define  $h_{t+1}$  as the weak GARCH volatility, i.e.  $h_{t+1} \equiv EL[\varepsilon_{t+1}^2 \mid H_t^s] = \omega + \alpha \varepsilon_t^2 + \beta h_t$  where  $H_t$  is the Hilbert space spanned by  $\{1, \varepsilon_{\tau}^2, \tau \leq t\}$ . Then, the  $R^2$  of the regression of  $f_t$  onto the constant and  $h_{t+1}$  is given by  $R^2 = 1 - \frac{\text{Var}[\psi_t]}{(1 - \beta^2)\text{Var}[f_t]}$  with  $\psi_t = v_t - \alpha \eta_t$  where  $\eta_t$  is defined by

$$\varepsilon_t^2 = f_{t-1} + \eta_t.$$

**Proof.** We have  $\varepsilon_t^2 = f_{t-1}u_t^2 = f_{t-1} + f_{t-1}(u_t^2 - 1) = f_{t-1} + \eta_t$ . The dynamics of  $f_{t-1}$  imply that  $\varepsilon_t^2 = \omega + \gamma\varepsilon_{t-1}^2 + \eta_t - \gamma\eta_{t-1} + v_{t-1} = \omega + \gamma\varepsilon_{t-1}^2 - \beta r_{t-1} + r_t = h_t + r_t$ . Observe that  $(1 - \beta L)r_t = (1 - \gamma L)\eta_t + v_{t-1} = (1 - \beta L)\eta_t + (v_{t-1} - \alpha\eta_{t-1})$ . Hence  $r_t = \eta_t + (1 - \beta L)^{-1}\psi_{t-1}$ . In other words,  $f_{t-1} - h_t = r_{t-1} - \eta_{t-1} = (1 - \beta L)^{-1}\psi_{t-1}$ . Remark that in the semi-strong GARCH(1,1) case,  $v_t = \alpha\eta_t$  and, hence,  $\psi_t = 0$  a.s. and  $\eta_t = r_t$  a.s.

Note that since  $\eta_t$  and  $v_t$  are m.d.s., we have  $\forall i \neq 0, \text{cov}(\psi_t, \psi_{t-i}) = \text{cov}(v_t - \alpha\eta_t, v_{t-i} - \alpha\eta_{t-i}) = 0$ . Hence  $\text{Var}[(1 - \beta L)^{-1}\psi_t] = \frac{1}{1 - \beta^2}\text{Var}[\psi_t]$ . As a consequence, the  $R^2$  of the regression of  $f_t$  onto the constant and  $h_t$  is given by

$$R^2 = 1 - \frac{\text{Var}[f_t - h_{t+1}]}{\text{Var}[f_t]} = 1 - \frac{\text{Var}[(1 - \beta L)^{-1}\psi_t]}{\text{Var}[f_t]} = 1 - \frac{\text{Var}[\psi_t]}{(1 - \beta^2)\text{Var}[f_t]}.$$

In practice, one have to compute the variance of  $\psi_t$  to obtain the  $R^2$ . We do this for the examples  $M1$  and  $M3$  considered in the Monte Carlo study. Recall (from Bollerslev, 1986) that for a strong Gaussian GARCH(1,1) with finite fourth moment,  $\varepsilon_t = \sqrt{f_{t-1}}u_t$  where  $u_t$  is i.i.d.  $\mathcal{D}(0,1)$  and  $f_{t-1} = \omega + \alpha\varepsilon_{t-1}^2 + \beta f_{t-2}$ , we have:  $E[\varepsilon_t^2] = \frac{\omega}{1 - \gamma}$ ,  $\text{Var}[\eta_t] = E[f_{t-1}](E[u_t^4] - 1)$  and  $E[f_{t-1}^2] = \omega^2 \frac{(1 + \gamma)}{(1 - \gamma)(1 - \beta^2 - 2\alpha\beta - 3\alpha^2)}$ .

**Application to the temporal aggregation case.** Define  $\varepsilon_{tm}^{(m)} = \sum_{i=1}^m \varepsilon_{(t-1)m+i}$  where  $\varepsilon_t$  is a strong GARCH(1,1) with finite fourth moment. with  $\varepsilon_t = \sqrt{f_{t-1}}u_t$ ,  $f_{t-1} = \omega + \alpha\varepsilon_{t-1}^2 + \beta f_{t-2}$ . Straightforward calculus which are available upon request show that:  $\eta_{tm}^{(m)} = \sum_{i=1}^m (\frac{\alpha(1 - \gamma^{m-i})}{1 - \gamma} + 1)\eta_{(t-1)m+i} + \sum_{1 \leq i, j \leq m, i \neq j} \varepsilon_{(t-1)m+i} \varepsilon_{(t-1)m+j}$ ,  $v_{tm}^{(m)} = a^{(m)}(\sum_{i=0}^{m-1} \gamma^i v_{tm-i})$ , where  $a^{(m)} = \frac{1 - \gamma^m}{1 - \gamma}$ . Then, the variance of  $\psi_{tm}^{(m)} = v_{tm}^{(m)} - \alpha^{(m)}\eta_{tm}^{(m)}$  is given by  $\text{Var}[\psi_{tm}^{(m)}] = \text{Var}[v_{tm}^{(m)}] + (\alpha^{(m)})^2 \text{Var}[\eta_{tm}^{(m)}] - 2\alpha^{(m)} \text{Cov}(v_{tm}^{(m)}, \eta_{tm}^{(m)})$  where  $\text{Var}[v_{tm}^{(m)}] = (a^{(m)})^2 \frac{1 - \gamma^{2m}}{1 - \gamma^2} \alpha^2 \text{Var}[\eta_t]$ ,  $\text{Var}[\eta_{tm}^{(m)}] = \sum_{i=1}^m (\frac{\alpha(1 - \gamma^{m-i})}{1 - \gamma} + 1)^2 \text{Var}[\eta_t] + m(m-1)(\frac{\omega}{1 - \gamma})^2$  and  $\text{Cov}(v_{tm}^{(m)}, \eta_{tm}^{(m)}) = a^{(m)}\alpha(\sum_{i=1}^m \gamma^{m-i}(\frac{\alpha(1 - \gamma^{m-i})}{1 - \gamma} + 1))\text{Var}[\eta_t]$ .

**Application to the cross-sectional aggregation case.** Define  $\varepsilon_t$  by  $\varepsilon_t = \varepsilon_{1,t} + \varepsilon_{2,t}$  where  $\varepsilon_{1,t}$  and  $\varepsilon_{2,t}$  are independent and strong GARCH(1,1) with finite fourth moment, with  $\varepsilon_{i,t} = \sqrt{f_{i,t-1}}u_{i,t}$ ,  $f_{i,t-1} = \omega_i + \alpha_i\varepsilon_{i,t-1}^2 + \beta_i f_{i,t-2}$ ,  $i = 1, 2$ . We have  $\varepsilon_t = \sqrt{f_{t-1}}u_t$  where  $f_{t-1} = f_{1,t-1} + f_{2,t-1}$  and  $u_t = \frac{\varepsilon_t}{\sqrt{f_{t-1}}}$ . Hence  $v_{t-1} = v_{1,t-1} + v_{2,t-1} = \alpha_1 r_{1,t-1} + \alpha_2 r_{2,t-1}$  and  $\eta_t = \eta_{1,t} + \eta_{2,t} + 2\varepsilon_{1,t}\varepsilon_{2,t}$ . Therefore  $\psi_t = v_t - \alpha\eta_t = (\alpha_1 - \alpha)\eta_{1,t} + (\alpha_2 - \alpha)\eta_{2,t} - 2\alpha\varepsilon_{1,t}\varepsilon_{2,t}$ .

We have  $\text{Cov}(\eta_{1,t}, \eta_{2,t}) = \text{Cov}(\eta_{1,t}, \varepsilon_{1,t}\varepsilon_{2,t}) = \text{Cov}(\eta_{2,t}, \varepsilon_{1,t}\varepsilon_{2,t}) = 0$  and  $\varepsilon_{1,t}$  and  $\varepsilon_{2,t}$  independent. Thus,  $\text{Var}[\psi_t] = (\alpha_1 - \alpha)^2 \text{Var}[\eta_{1,t}] + (\alpha_2 - \alpha)^2 \text{Var}[\eta_{2,t}] + 4\alpha^2 E[\varepsilon_{1,t}^2]E[\varepsilon_{2,t}^2]$ . Finally, we have  $\text{Var}[f_t] = E[f_{1,t}^2] + E[f_{2,t}^2] + 2E[f_{1,t}]E[f_{2,t}] - (\frac{\omega}{1 - \gamma})^2$ .  $\square$

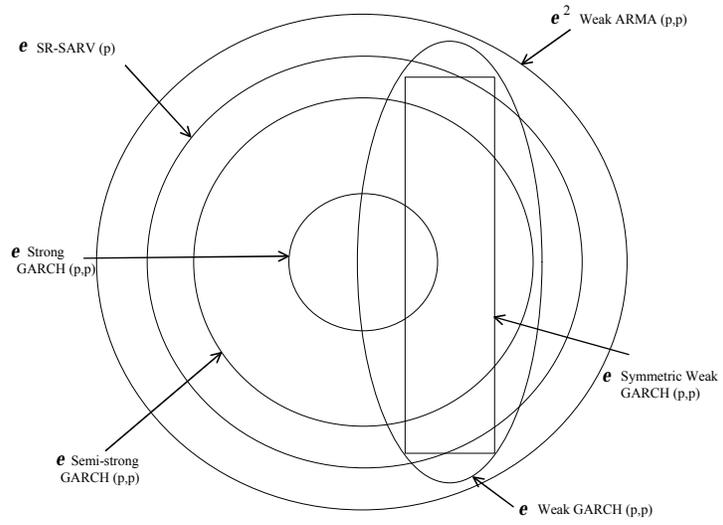


Figure 1:

Fig. 1. The relationships between the models when the fourth moment of  $\epsilon$  is finite