

# Memory and persistence in models of volatility in financial time series

Submitted by

**Xiaoyu Li**

to the **University of Exeter**

as a thesis for the degree of

Doctor of Philosophy in

Economics

June 2016

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: .....

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my first supervisor, Professor James Davidson, for his patient guidance, continuous encouragement and support of my PhD research. He has been the best advisor for my PhD studies in my mind. I would also like to thank my second supervisor, Dr Andreea Halunga, for her useful comments and suggestions.

I gratefully acknowledge the funding from the ESRC.

I am grateful to the staff of the Business School at the University of Exeter, especially to Kate Gannon and Helen Bell, for their support and help throughout my PhD studies. I also thank my colleagues and friends for all their encouragement and support.

I also thank the participants at the following conferences, who made comments on my research: University of Exeter Economics Seminar 2011-2016; 7th and 9th International Conference on Computational and Financial Econometrics, London, 2013 and 2015; 11th International Vilnius Conference on Probability Theory and Mathematical Statistics, 2014.

Finally, I would like to thank my whole family for their love and support, especially my parents and my husband, Yuchen Wei.

## ABSTRACT

This thesis first investigates the moment and memory properties of exponential-type conditional heteroscedasticity models. This primarily includes exponential generalised autoregressive conditional heteroscedastic (EGARCH) models, the fractionally integrated EGARCH model of Bollerslev and Mikkelsen (1996) (FIEGARCH(BM)), the hyperbolic EGARCH (HYEGARCH) model and the FIEGARCH(DL) model, as presented in Chapter 2. The moment conditions of these models are derived from previous literature, and the memory properties are measured by using the near-epoch dependence (NED) functions of an independent process approach. The existence of moments supports the limited memory properties of these models. This study shows that exponential autoregressive conditional heteroscedastic (EARCH)( $\infty$ ) processes may exhibit geometric memory, hyperbolic memory or long memory. The EGARCH is a case of a geometric memory process. The FIEGARCH(BM) and HY/FIEGARCH(DL) processes can exhibit hyperbolic memory or long memory, depending on the sign of the memory parameter. The study also derives the functional central limit theorem (FCLT) or fractional FCLT for the relevant processes in these exponential-type conditional heteroscedasticity models. Finally, the results of the simulation show that the HYEGARCH model has a hyperbolic memory and that the FIEGARCH(DL) model can capture long memory in absolute return series.

Next, the study investigates the asymptotic properties of the quasi-maximum likelihood estimator (QMLE) in autoregressive moving average (ARMA) models with EGARCH or HY/FIEGARCH(DL) errors in Chapter 3. This part of the study aims to investigate the asymptotic theory of the ARMA(1, 1)-EGARCH(1, 1) models and that of the pure HY/FIEGARCH(DL) models. First, the literature

on the asymptotic properties of the ARMA-GARCH and EGARCH processes is reviewed. The conditions for the consistency and asymptotic normality of the QMLE of the ARMA-EGARCH models are then demonstrated. This analysis also provides an investigation of that of the QMLE in the HY/FIEGARCH(DL) processes. A Monte Carlo simulation is used to study the properties of the QMLE in the pure HY/FIEGARCH(DL) processes.

Lastly, in a study co-authored with Professor James Davidson, we derive a simple sufficient condition for strict stationarity in the ARCH( $\infty$ ) class of processes with conditional heteroscedasticity. The concept of persistence in these processes is explored, and is the subject of a set of simulations showing how persistence depends on both the pattern of the lag coefficients of the ARCH model and the distribution of the driving shocks. The results are used to argue that an alternative to the usual method of ARCH/GARCH volatility forecasting should be considered.

## TABLE OF CONTENTS

Acknowledgements . . . . .	2
List of Tables . . . . .	11
List of Figures . . . . .	12
<b>1 Introduction</b>	<b>13</b>
1.1 Basic concepts . . . . .	20
<b>2 Moment and memory properties of the exponential-type conditional heteroscedasticity models</b>	<b>24</b>
2.1 Introduction . . . . .	24
2.2 Literature review . . . . .	28
2.2.1 Relevant volatility models . . . . .	28
2.2.2 Literature on the moment property of conditional heteroscedasticity models . . . . .	30
2.2.3 The memory property of conditional heteroscedasticity models	32
2.2.4 The FCLT and fractional FCLT of conditional heteroscedasticity models . . . . .	33
2.3 Exponential-type conditional heteroscedasticity models . . . . .	35

2.4	Moment properties of the exponential-type conditional heteroscedasticity models . . . . .	39
2.5	Memory properties of the exponential-type conditional heteroscedasticity models . . . . .	43
2.5.1	Memory properties of the EGARCH-type models . . . . .	43
2.5.2	The HYEGARCH and FIEGARCH(DL) models . . . . .	48
2.6	FCLT and fractional FCLT for the EGARCH-type models . . . . .	50
2.6.1	The FCLT for the EGARCH( $p, q$ ) process . . . . .	52
2.6.2	Investigation of the FCLT and fractional FCLT for the FIEGARCH(BM) process . . . . .	54
2.6.3	Investigation of the FCLT and fractional FCLT for the HYEGARCH and FIEGARCH(DL) processes . . . . .	55
2.7	Simulation for the memory properties of HYEGARCH and FIEGARCH(DL) processes . . . . .	57
2.7.1	Data generating process . . . . .	57
2.7.2	Estimation procedure and simulation results . . . . .	58
2.8	Conclusion . . . . .	59
2.9	Appendix A . . . . .	61
2.9.1	Proof of Theorem 2.5.2 . . . . .	61

2.9.2	Proof of Theorem 2.5.3 . . . . .	65
2.9.3	Proof of Theorem 2.6.4 . . . . .	66
2.9.4	Proof of Theorem 2.6.5 . . . . .	67
2.9.5	Proof of Theorem 2.6.6 . . . . .	68

**3 Asymptotic theory of the QMLE in ARMA models with EGARCH and HY/FIEGARCH errors 71**

3.1	Introduction . . . . .	71
3.2	Literature review . . . . .	73
3.2.1	Stationarity and ergodicity properties of the relevant volatility models . . . . .	74
3.2.2	Asymptotic theory of the QMLE in ARCH/GARCH models	75
3.2.3	Asymptotic theory of the QMLE in ARMA-type models with GARCH-type errors . . . . .	76
3.2.4	Asymptotic theory in the EGARCH-type models . . . . .	77
3.2.5	Finite sample properties of the EGARCH-type models . . . . .	80
3.3	The models and the QMLE . . . . .	81
3.3.1	The ARMA(1,1)-EGARCH(1,1) model . . . . .	81
3.3.2	The HY/FIEGARCH models . . . . .	83
3.3.3	Quasi-maximum-likelihood estimator . . . . .	85

3.4	The invertibility of the EGARCH-type processes . . . . .	87
3.4.1	Invertibility of the EGARCH( $p, q$ ) process . . . . .	87
3.4.2	Invertibility of the ARMA-EGARCH process . . . . .	95
3.4.3	Invertibility of the HYEGARCH process . . . . .	98
3.4.4	The invertibility of the derivatives of the ARMA-EGARCH process . . . . .	101
3.5	Asymptotic theory of the QMLE in ARMA(1,1) with an EGARCH(1,1) error . . . . .	104
3.5.1	Consistency of the QMLE in ARMA(1,1)-EGARCH(1,1) process	105
3.5.2	Asymptotic normality of the QMLE in ARMA(1,1) with an EGARCH(1,1) error . . . . .	106
3.6	The asymptotic theory of the QMLE in HY/FIEGARCH processes	110
3.6.1	Asymptotic theory of the QMLE in the HYEGARCH process	111
3.6.2	Investigation of the asymptotic theory of the QMLE in the FIEGARCH process . . . . .	118
3.6.3	Simulations for the HY/FIEGARCH models . . . . .	120
3.7	Conclusion . . . . .	120
3.8	Appendix B . . . . .	125
3.8.1	Proof of Lemma 3.4.1 . . . . .	125



3.8.2	Proof of Lemma 3.5.1 . . . . .	129
3.8.3	Proof of Lemma 3.5.2 . . . . .	130
3.8.4	Proof of Lemma 3.5.3 . . . . .	133
3.8.5	Proof of Theorem 3.5.2 . . . . .	135
3.8.6	Proof of Lemma 3.5.4 . . . . .	136
3.8.7	Proof of Lemma 3.5.5 . . . . .	141
3.8.8	Proof of Lemma 3.5.6 . . . . .	145
3.8.9	Proof of Lemma 3.6.1 . . . . .	151
3.8.10	Proof of Lemma 3.6.2 . . . . .	152
3.8.11	Proof of Lemma 3.6.3 . . . . .	153
3.8.12	Proof of Lemma 3.6.4 . . . . .	156
3.8.13	Proof of Lemma 3.6.5 . . . . .	157
3.8.14	Proof of Lemma 3.6.6 . . . . .	159
3.8.15	Proof of Theorem 3.6.2 . . . . .	160
3.8.16	Proof of Lemma 3.6.7 . . . . .	161
3.8.17	Proof of Lemma 3.6.8 . . . . .	163

**4 Strict stationarity, persistence and volatility forecasting in ARCH( $\infty$ ) processes 166**

4.1	Introduction . . . . .	166
4.2	Literature review . . . . .	170
4.2.1	The stationarity of the ARCH-type models . . . . .	170
4.2.2	Persistence in the ARCH( $\infty$ ) process . . . . .	176
4.3	Stationarity and persistence in ARCH( $\infty$ )-class processes . . . . .	178
4.4	Measuring the persistence of stationary time series . . . . .	186
4.4.1	$J_T$ statistics . . . . .	187
4.4.2	GPH estimator and that for the normalised ranks series . . . . .	189
4.4.3	Persistence measures . . . . .	190
4.5	Some simulation experiments . . . . .	191
4.5.1	Date generation process . . . . .	191
4.5.2	Measurement approach and simulation results . . . . .	193
4.6	Implications for volatility forecasting . . . . .	197
4.7	Conclusion . . . . .	203
4.8	Appendix C . . . . .	205
4.8.1	Proof of Proposition 4.3.1 . . . . .	205
4.8.2	Proof of Proposition 4.3.2 . . . . .	207

**5 Conclusion** **209**

## LIST OF TABLES

2.1	Simulation results for the HYEGARCH and FIEGARCH(DL) processes.	60
3.1	Simulation results for the QMLE in the HYEGARCH( $0, d, 0$ ) process.	121
3.2	Simulation results for the QMLE in the FIEGARCH( $0, d, 0$ ) process.	121
4.1	Persistence measures in a fractional linear time series, . . . . .	191
4.2	Series properties and persistence measures for the covariance stationary GARCH(1, 1) model. . . . .	195
4.3	Series properties and persistence measures for the nonstationary GARCH(1, 1) model. . . . .	196
4.4	Series properties and persistence measures for the covariance stationary HY/FIGARCH model. . . . .	197
4.5	Series properties and persistence measures for the nonstationary HY/FIGARCH model. . . . .	198
4.6	2MAV 2-step forecast error in GARCH(1, 1), against $M$ . . . . .	203
4.7	MAV 2-step forecast error in HY/FIGARCH, against $M$ . . . . .	204

## LIST OF FIGURES

2.1	Simulation plots of the pure HYGARCH and FIEGARCH(DL) models. Note: This figure provides a corresponding plot of the simulation results presented in Table 2.1. The estimated memory parameters in the HYGARCH and FIEGARCH(DL) models are indicated by a blue line and a red line, respectively. It is clear to see that the estimated memory parameters in the HYGARCH model are very close to zero but, in the FIEGARCH(DL) model, the estimated memory parameters increase consistently with the absolute value of $d$ . . . . .	60
4.1	Gaussian GARCH(1,1) model: $(\alpha, \beta)$ pairs where $\zeta = 0$ and the stationarity boundary points of Nelson (1990). Note: this figure provides some numerical experiments with Gaussian shocks showing $\alpha$ -values at which $\zeta \approx 0$ for $\beta = 0, 0.1, 0.2, \dots, 0.9$ . The mean is estimated in each case as the average of 20,000 values of $\ln(\Psi_t)$ . The actual stationarity boundary points from Nelson (1990) are also shown for comparison. . . . .	181
4.2	Simulation of IGARCH(1,1) with Student(3) shocks. Note: This figure provides the simulation results for samples of size 5000, with 10,000 pre-sample steps, for the IGARCH(1,1) model with $\omega = 1$ , $\beta = 0.9$ and Student(3) shocks. It shows the persistence properties of the IGARCH(1,1) process with Student(3) shocks. More interpretations of this figure are shown in Section 4.3. . . . .	184
4.3	Simulation of IGARCH(1,1) with Gaussian shocks. This figure follows same simulation procedure to figure 4.2, except it considers the IGARCH(1,1) process with Gaussian shocks. This figure shows the persistence properties of the IGARCH(1,1) process with Gaussian shocks. More interpretations of this figure are shown in Section 4.3. . . . .	184
4.4	Simulation of IGARCH(1,1) with uniformly distributed shocks. Note: This figure follows same simulation procedure to figure 4.2, except it considers the IGARCH(1,1) process with uniformly distributed shocks. This figure shows the persistence properties of the IGARCH(1,1) process with uniformly distributed shocks. More interpretations about this figure are shown in Section 4.3. . . . .	184

CHAPTER 1  
INTRODUCTION

Volatility modelling and forecasting play an important role in option pricing, dynamic capital theory and asset pricing theory. Over the last few decades, many methods of volatility modelling have been introduced. Among these methods, Engle (1982) first introduced the autoregressive conditional heteroscedastic (ARCH) model to investigate the volatility of UK inflation. The expression of the ARCH( $p$ ) model is shown in Equations (1.1) to (1.3):

$$\xi_t = \sqrt{h_t} z_t; \tag{1.1}$$

$$h_t = \omega + \sum_{j=1}^p \alpha_j \xi_{t-j}^2; \tag{1.2}$$

$$h_t = E[\xi_t^2 | \mathcal{F}_{t-1}], \tag{1.3}$$

where  $\{\xi_t\}$  is a real-valued discrete time stochastic process, which can also be interpreted as a return series; process  $\{z_t\}$  is an independent and identically distributed (i.i.d.) stochastic variable with a mean of zero and a variance of one, such that  $z_t \sim i.i.d.(0, 1)$ , which can be interpreted as an innovation or shock; and  $\sqrt{h_t}$  has a positive value with a probability of one, such that  $\sqrt{h_t} > 0$ , and it can be interpreted as the volatility of the return series. The series of  $\xi_t$  is independent because the underlying series  $z_t$  is an i.i.d. and is independent of  $\sqrt{h_t}$ . In Equations (1.2) and (1.3),  $h_t$  is the conditional variance of  $\xi_t$ , and the  $\mathcal{F}_{t-1}$ -measurable function, where  $\mathcal{F}_{t-1}$  is the  $\sigma$ -algebra generated by  $\{z_{t-1}, z_{t-2}, \dots\}$ , indicating the information up to and including time  $t - 1$ . The well-defined ARCH( $p$ ) processes require that  $\omega > 0$ , and  $\alpha_j \geq 0$  ( $j = 1, 2, \dots, p$ ). Since Engle's study was published, the ARCH processes have been widely used for modelling financial time series because they can capture the stylised facts of financial data, such as volatility clustering and

leptokurtic characteristics. Bollerslev (1986) proposed the generalised ARCH( $p, q$ ) model, which is defined in Equations (1.1) and (1.4)

$$\begin{aligned} h_t &= \omega + \sum_{i=1}^q \alpha_i \xi_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \\ &= \omega + \alpha(L) \xi_t^2 + \beta(L) h_t, \end{aligned} \tag{1.4}$$

where all parameters are non-negative,  $L$  denotes the lag operator, (i.e.  $L^k z_t = z_{t-k}$ ), and  $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q$  and  $\beta(L) = \beta_1 L + \beta_2 L^2 + \dots + \beta_p L^p$  are polynomials in the lag operator. Equation (1.4) is the benchmark GARCH(1,1) model when  $p = q = 1$ .

Although the GARCH ( $p, q$ ) process represents significant progress and is very popular, it has some drawbacks. For example, it is unable to capture some of the characteristics of the volatility, such as asymmetric and long memory (a descriptive definition is provided in Section 1.1). To overcome these drawbacks, several conditional heteroscedasticity models have been proposed. In order to capture the asymmetric properties of the volatility, the exponential GARCH (EGARCH), the Glosten–Jagannathan–Runkle GARCH (GJR-GARCH) and threshold GARCH (TGARCH) models were introduced by Nelson (1991), Glosten et al. (1993) and Zakořan (1994) respectively. To capture the high persistence properties of financial data, several conditional mean models, such as  $I(1)$  and  $I(d)$ , have been extended to conditional variance models based on similar ideas. For instance, Engle and Bollerslev (1986) proposed the integrated GARCH (IGARCH) model, and Baillie et al. (1996) introduced the fractionally integrated GARCH (FIGARCH) model. However, the IGARCH model exhibits short memory or geometric memory, see Ding and Granger (1996) and Davidson (2004), and the stationary FIGARCH fails to capture the long memory property of volatility, see e.g. Giraitis et al. (2009). In addition, Robinson (1991) introduced the ARCH( $\infty$ ) to model long-run depen-

dence in volatility, but the squares  $(\xi_t^2)$  from the stationary ARCH( $\infty$ ) process does not allow the existence of long memory in volatility. Nevertheless, the memory properties of the nonstationary ARCH( $\infty$ ) process, including the IGARCH and FIGARCH models, have not been fully explored. Chapter 4 discusses the persistence in the IGARCH, FIGARCH and ARCH( $\infty$ ) processes further.

A few models have also attempted to capture both the asymmetric and high persistence properties of volatility. For instance, the linear ARCH (LARCH) process and the fractionally integrated EGARCH (henceforth FIEGARCH(BM)) model were introduced by Robinson (1991) and Bollerslev and Mikkelsen (1996), respectively. In addition, in order to distinguish between the different memory properties of the EGARCH-type models, a general hyperbolic/fractional integrated EGARCH (henceforth HY/FIEGARCH (DL)) is proposed in Chapter 2. A stochastic volatility class model was also introduced by Breidt et al. (1998) and Harvey (2002), namely the long memory stochastic volatility (LMSV) model. Ruiz and Veiga (2008) developed the theoretical property of the asymmetric LMSV (A-LMSV) model.

Research on the statistical properties of these conditional heteroscedasticity models has attracted a substantial amount of attention, since these models are widely used in practice. Among these models, the theoretical properties of the ARCH and GARCH models have been extensively explored. The memory property and the asymptotic theory of the estimator of the parametric LARCH process have been well developed by Giraitis et al. (2000b), Giraitis et al. (2004), Schützner (2009), and Beran and Schützner (2009), although some of the theoretical properties remain unresolved. The statistical properties of the LMSV models are worthy of further academic consideration, but these are more complex and are therefore

not included in this thesis. In addition, mystery continues to surround the theoretical properties of EGARCH-type models, even these have been investigated in a few recent studies, such as Surgailis and Viano (2002), Ruiz and Veiga (2008), and Lopes and Prass (2014).

The second and third chapters of this thesis investigate the theoretical properties of EGARCH-type models. More specifically, Chapter 2 considers the moment and memory properties of EGARCH-type models, and near-epoch dependence (NED) is applied to measure the memory property and establish the functional central limit theorem (FCLT) and fractional FCLT of the relevant processes. The main reasons for undertaking this study are as follows. Firstly, this idea was inspired by Davidson's (2004) study, which investigated the moment and memory properties of linear conditional heteroscedasticity models. These properties are the two main statistical features of such models. Research on the moment properties of conditional variance models is useful for measuring how large the effects of shocks to volatility can be, and the existence of moments is a necessary condition for determining the memory properties of conditional heteroscedasticity models, see e.g. He and Teräsvirta (1999a) and Davidson (2004). The memory property shows how long the effect of shocks on conditional variance can persist, see e.g. Giraitis et al. (2000a), Davidson (2004) and Giraitis et al. (2009). Therefore, research on these properties is significant for volatility estimation and forecasting.

Secondly, motivated by Davidson (2004), in Chapter 2, the NED concept is applied to measure the memory property of EGARCH-type models. A descriptive definition of the NED is provided in Section 1.1. The main reason for applying the NED approach is that this method may be easier for parameter estimation and inference in relevant models, especially for nonlinear and long memory conditional



heteroscedasticity models. Davidson (2004) introduced geometric and hyperbolic memory when applying the NED approach to measure the persistence of linear conditional heteroscedasticity models including GARCH, IGARCH, FIGARCH and HYGARCH. The theoretical and empirical results show that ARCH( $\infty$ ) processes may exhibit hyperbolic memory or geometric memory, the FIGARCH process has hyperbolic memory, and both the standard GARCH and IGARCH processes have geometric memory. However, the models in Davidson (2004) did not include the EGARCH-type models. This is discussed further in Chapter 2.

Thirdly, the FCLT and fractional FCLT are vital for statistical inference of financial time series. The descriptive definitions of the FCLT and fractional FCLT are provided in Section 1.1. The concept of NED is applied for the FCLT, as was proposed by McLeish (1975). It satisfies the conditions for the law of large numbers (LLN), the central limit theorem (CLT), FCLT and fractional FCLT, and its restrictions can easily be verified. (see e.g. Davidson, 1992, 1993, 1994; De Jong and Davidson, 2000; Davidson and De Jong, 2000). Thus the first aim of Chapter 2 is to investigate the existence of the finite order moment of the relevant processes and, following a similar procedure to Davidson (2004), to study the NED properties of the EGARCH-type models, and then to measure the memory properties of these models by applying the concept of NED on an independent process. The second aim is to construct the FCLT and fractional FCLT for the relevant EGARCH-type models.

Chapter 2 makes three main contributions. Firstly, it shows that EARCH( $\infty$ ) processes can have hyperbolic memory, geometric memory and long memory, as the EGARCH is a geometric memory process, the FIEGARCH(BM) with a negative memory parameter has a hyperbolic memory, and the FIEGARCH(BM) with

a positive parameter can capture long memory in volatility. Secondly, this study provides a general HY/FIEGARCH(DL) process, that the HYEGARCH process has hyperbolic memory and the FIEGARCH(DL) process exhibits long memory. Thirdly, it establishes the FCLT for the EGARCH and the fractional FCLT for the partial sum of the  $\{\ln h_t - \omega\}$  processes in the FIEGARCH(BM) and FIEGARCH(DL) models.

Chapter 3 focuses on establishing an asymptotic theory of the quasi-maximum-likelihood estimators (QMLE) in the ARMA model with EGARCH-type errors. The QMLE is a popular estimation method for conditional heteroscedasticity models. The ARMA model with EGARCH-type innovations has been widely used in empirical analyses of the financial time series. The asymptotic properties of the QMLE in the ARMA, GARCH and ARMA-GARCH processes have already been examined in detail (Weiss, 1986; Lee and Hansen, 1994; Ling and McAleer, 2003; Francq and Zakoïan, 2004; Straumann, 2005). However, only a few recent studies have focused on the asymptotic theory of the EGARCH process. For example, Straumann (2005) and Straumann and Mikosch (2006) proved the consistency of the QMLE in EGARCH (1, 1) and asymptotic normality for EGARCH(1,0) under strong conditions. Wintenberger (2013) established the strong consistency of the QMLE, and the consistency and asymptotic normality (CAN) of the stable QMLE in the EGARCH(1, 1) process under conditions that are difficult to verify. Martinet and McAleer (2015) investigated the asymptotic theory of the EGARCH( $p, q$ ) model by deriving it from a stochastic process, which made it easier to verify the invertibility conditions. However, some open questions in this area are yet to be answered, especially regarding the asymptotic properties of the QMLE for the ARMA-EGARCH and HY/FIEGARCH(DL) models. Therefore, it is vital to investigate the asymptotic properties of the estimators in the ARMA process with

EGARCH or HY/FIEGARCH(DL) errors.

The main purpose of the research described in Chapter 3 is to establish an asymptotic theory for the ARMA model with EGARCH or HYGARCH errors, and a planned further study will extend these results to the FIEGARCH(DL) and ARMA-FIEGARCH processes.

The main contributions of Chapter 3 are as follows. Firstly, the CAN of the QMLE of the ARMA-EGARCH is established. Secondly, the consistency of the QMLE in the pure HYGARCH process is proved. The asymptotic normality of the pure HYGARCH process and the asymptotic property of the FIEGARCH(DL) are discussed, indicating the difficulty of the investigation.

The fourth chapter is a joint work with Professor James Davidson, the paper is forthcoming in *Journal of Empirical Finance*, see Davidson and Li (2015). This chapter first examines the stationarity property of ARCH-type models. A considerable volume of previous literature has considered the existence of covariance stationarity and strict stationarity in ARCH-type models. The covariance stationarity and strict stationarity are defined in Section 1.1. However, the existence of strict stationarity for IGARCH and FIGARCH models with covariance non-stationary is still an open question.

Moreover, further research needs to be undertaken on the persistence of the covariance non-stationary ARCH( $\infty$ ) processes. In the empirical literature, the FIGARCH (Baillie et al., 1996) and HYGARCH (Davidson, 2004) models are widely used to model the long memory in volatility. However, the persistence properties of the FIGARCH process have not yet been fully investigated (see e.g. Giraitis et al., 2009; Beran et al., 2013). The stationary HYGARCH model is

unable to capture long memory, although it embodies hyperbolic memory, as does the stationary FIGARCH model (Davidson, 2004).

Furthermore, the existing volatility forecasting method for GARCH-type models normally replaces the square of the returns with its conditional expectation (see e.g. Poon, 2005). This may not be an appropriate way to forecast volatility.

Therefore, the main purposes of Chapter 4 are: (1) to study the strictly stationary property of the nonstationary ARCH( $\infty$ ) process; (2) to measure the persistence of the relevant volatility process and to investigate the relation how the persistence of the process depends on the lag coefficients of the parametric ARCH model and the distribution of the driving shocks; and (3) to investigate the optimal method of volatility forecasting.

The main contributions of Chapter 4 are as follows. Firstly, it provides a simple sufficient condition for strict stationarity in the ARCH( $\infty$ ) class of processes by applying the Beveridge-Nelson (BN) decomposition. Secondly, it explores the persistence of these processes and proposes the  $J_T$  statistic to measure the persistence of the ARCH( $\infty$ ) process. Thirdly, it investigates that how persistence depends on both the pattern of the lag coefficients of the ARCH models and the distribution of the driving shocks. Finally, it provides an alternative method for volatility forecasting.

## 1.1 Basic concepts

This section aims to provide descriptive definitions for some terminologies, such as long memory, NED, FCLT, fractional FCLT, strict stationarity and covariance stationarity.

Long memory was first used in the fields of hydrology and climatology areas around 1950. Granger (1980), and Hosking (1981) introduced the use of this concept to investigate economics and financial data. The notion of long memory, is also known as long-range dependence or strong dependence. It has been defined in both the time domain and the frequency domain, see e.g. Beran (1994) and Giraitis et al. (2012). For a covariance stationary process  $\{x_t\}$ , long memory normally means the spectral density is unbounded at the origin, with the condition of the sum of the absolute value of the autocovariance function being infinite. For short memory, the sum of the autocovariance function of  $\{x_t\}$  is finite, which means that  $\{x_t\}$  has a continuous bounded spectral density at the zero frequency, see e.g. Beran (1994) and Giraitis et al. (2012). Davidson (2004) also introduced geometric and hyperbolic memory to investigate the memory properties of conditional variance processes. A considerable body of evidence shows that there is long memory in volatility and the power transformation of the return processes. Long memory in volatility means that the effect of shocks on volatility decays at a hyperbolic rate (see e.g. Beran, 1994; Baillie, 1996) .

The NED was introduced by Ibragimov (1962), and was then developed by Billingsley (1968), Gallant and White (1988), and Davidson (1994), among others. Gallant and White (1988) first introduced the NED on a mixing process to econometrics. The main idea of this concept is that, for the stochastic process  $\{x_t, t \in Z\}$ , we can predict the  $x_t$  exactly if we know all the past information about  $x_t$ . If we predict that it is only dependent on the information relating to the near epoch, such that  $E(x_t | \mathcal{F}_{t-m}^{t+m})$ , and the  $L_p$ -norm of the difference between  $x_t$  and  $E(x_t | \mathcal{F}_{t-m}^{t+m})$  converges to zero at as  $m$  tends to infinity. We can then say that  $\{x_t, t \in Z\}$  satisfies the condition of NED. If the process has geometric decay as  $m$  tends to infinity, then the process has geometric NED. (see e.g. Davidson, 1994)

More details about the concept of NED are introduced in Chapter 2.

The FCLT is a generalisation of the CLT on functional spaces, such as the space  $D_{[0,1]}$ , which is the space of the right continuous function whose left limit exists everywhere on the unit interval, also known as the cadlag function (See e.g. Stock, 1994; Davidson, 1994, 2006). It was first introduced for i.i.d. increments, known as Donsker's Theorem (Donsker, 1951). It is also known as the invariance principle and related to the function-space version of the continuous mapping theorem (see e.g. Davidson 1994, 2000). This current study focuses on the FCLT for partial-sum processes of the relevant return and volatility processes. In a wider definition, a random process holds the FCLT, which means that the partial sum of these processes converges in distribution to standard Brownian motion, which is a continuous time stochastic process having independent Gaussian increments, under certain moment conditions (see, e.g. Baillie, 1996; Davidson, 2002). The fractional FCLT for fractional integrated processes having an i.i.d. innovation was first developed by Davydov (1970). Marinucci and Robinson (2000) extended the fractional FCLT to the fractional integrated process having a class of linear processes under certain moment conditions. Davidson and De Jong (2000) provided a fractional FCLT based on the concept of NED under a much weaker moment condition than previous literature. Johansen and Nielsen (2012) improved the conditions for the fractional FCLT from Davidson and De Jong (2000) and proposed a necessary condition for the fractional FCLT. In a wide sense, a random process holds the fractional FCLT, meaning that the partial sum of the stochastic process converges in distribution to fractional Brownian motion, having dependent increments, under certain regularity conditions, see e.g. Davidson and De Jong (2000). For the discrete time versions of Brownian motion and fractional Brownian motion are known as random walk and fractional differenced white noise, respectively, see e.g.

Baillie (1996).

The stationarity properties of time series include covariance stationarity and strict stationarity. The covariance stationarity (or weak stationarity) means that the time series  $\{x_t, t \in Z\}$  has a constant mean and finite variance, and that the covariance between  $\{x_t\}$  and  $\{x_{t+j}\}$  for  $j > 0$  does not depend on  $t$ , and only depends on the  $j$ . Strict stationarity means that the joint distributions of the time series  $\{x_t, t \in Z\}$  are identical, and the joint distributions of the collections  $(x_t, x_{t+1}, \dots, x_{t+k})$ , for all  $k > 0$ , do not depend on  $t$  (see e.g. Davidson, 1994, 2000). It is important to note that strict stationarity does not require the existence of the second moment. Therefore, strict stationarity does not imply weak stationarity, and neither does weak stationarity imply strict stationarity. In the Gaussian case, the terms "covariance stationary" and "strictly stationary" are equivalent (see e.g. Davidson, 1994, 2000).

CHAPTER 2  
MOMENT AND MEMORY PROPERTIES OF THE  
EXPONENTIAL-TYPE CONDITIONAL HETEROSCEDASTICITY  
MODELS

## 2.1 Introduction

Over the last few decades, EGARCH-type models, especially the FIEGARCH(BM) model, have been successfully applied to model the volatility of financial data. Recent research has increasingly focused on the theoretical properties of conditional variance models. The moment properties of ARCH-class models have been extensively studied in the literature (e.g. He and Teräsvirta, 1999a,b; He et al., 2002; Davidson, 2004). Studying the moment properties of these volatility models is important for several reasons. Firstly, it provides a better understanding of the relationship between future data and past information. Secondly, it is useful for determining the size of any shocks to volatility (e.g. Davidson, 2004). Thirdly, the existence of moments provides necessary conditions for the limited memory property of the processes (e.g. Davidson, 2004). Finally, moment properties play an important role in investigating the stationary property and establishing the FCLT of time series.

A growing body of literature has focused on the memory properties of conditional heteroscedasticity models. Several methods have been applied to measure the persistence of ARCH-type models, such as the rate of decay of autocorrelation, computing autocorrelation, mixing processes and NED (see e.g. Beran, 1994; Davidson, 2004). If we compare the NED approach with other methods, NED may be an easier way to determine the memory properties of ARCH-type models,



especially for models which capture the long-run dependence in volatility. Firstly, the uncorrelatedness may not satisfy the requirements of the LLN or CLT for long memory processes (see e.g. Beran, 1994; Davidson, 2004). It may cause problems associated with estimation and inference in the models. However, Gallant and White (1988) emphasised that the NED concept is significant in constructing the uniform LLN and establishing an asymptotic theory of the estimators. Moreover, although the mixing process also satisfies the restrictions of the CLT, it is difficult to verify the mixing properties of the processes. In addition, some financial time series do not satisfy the conditions of mixing processes, especially the infinite order of stochastic processes, but they mostly satisfy the conditions of NED on a mixing process. For instance, Andrews (1984) argued that some AR(1) processes might not to be a mixing process. It is evident in Davidson (1994) that these processes can be NED on a mixing process. Furthermore, the concept of NED can be applied to measure the limited memory of both linear and nonlinear models (see e.g. Davidson, 1994, 2004). Further information on the NED concept can be found, for example, in Gallant and White (1988), Wooldridge and White (1988), and Davidson (2002).

Therefore, the concept of NED plays a crucial role in determining the memory properties of time series models, and researchers are increasingly paying attention to the NED properties of linear and nonlinear models. For example, Davidson (2002) showed that the innovations of several nonlinear models are near-epoch dependent, including the GARCH model. Davidson (2004) investigated the NED properties of linear conditional heteroscedasticity models including GARCH, IGARCH, and FIGARCH, and found that the ARCH( $\infty$ ) models may exhibit either geometric memory or hyperbolic memory under certain conditions, that the FIGARCH model can capture the hyperbolic memory in volatility, and that both

the standard GARCH and IGARCH models can capture the geometric memory in volatility. The results of Davidson (2004) provide evidence against the opinions of some of researchers, who claimed that the IGARCH process has the highest persistence and that there is hyperbolic memory in the FIGARCH process. Chapter 4 of this thesis discusses the persistence of the IGARCH and FIGARCH models further.

Consequently, motivated by the advantages of NED on a mixing process, one of the main purposes of this study is to investigate the moment and memory properties of the EGARCH-class models by applying the NED on an independent process concept. The findings of this study show that  $\text{EARCH}(\infty)$  processes may exhibit hyperbolic memory, geometric memory or long memory. More specifically, the EGARCH model is an example of capturing geometric memory; the FIEGARCH(BM) can capture long memory in volatility when it has a positive memory parameter and hyperbolic memory in volatility when it has a negative memory parameter. Considering these results, this study proposes that the HYEGARCH model can capture the hyperbolic memory and the FIEGARCH(DL) exhibits long memory in volatility.

Moreover, the NED concept is widely applied to prove limit theorems, such as the LLN, CLT, FCLT and fractional FCLT, for time series processes. For instance, Wooldridge and White (1988) showed the usefulness of the NED concept to prove the CLT and FCLT, and Hansen (1991) demonstrated the limit theorems for GARCH(1, 1) processes by applying the NED approach. Meanwhile, these asymptotic convergence results are vital for the statistical inference of time series models. The FCLT and fractional FCLT play an especially important role in the statistical property of integrated processes (see Davidson, 2002). Lee (2014) provided appli-

cations to change point analysis and obtained the asymptotic distribution of the least-squares estimator for a unit root process with a GARCH error by applying the FCLT. The NED concept can be applied to derive the FCLT and fractional FCLT for the time series. In terms of this discussion, it is worth investigating these limit theorems for EGARCH-type processes by using NED. Another purpose of this chapter is to apply the NED approach to establish the FCLT or fractional FCLT for the relevant processes in EGARCH-type models.

The rest of this chapter is organised as follows. Section 2.2 briefly reviews some of the main literature on the statistical properties and empirical results of the relevant volatility models. Section 2.3 introduces the exponential-type conditional heteroscedasticity models, including the EGARCH and FIEGARCH(BM) models. Section 2.4 focuses on obtaining restrictions for the existence of moments with finite order in the exponential-type conditional heteroscedasticity models by applying the results of Nelson (1991). Section 2.5 considers whether or not the return series has NED, applies NED to measure the limited memory properties of shocks to volatility, and introduces the HY/FIEGARCH(DL) models. Section 2.6 investigates the FCLT or fractional FCLT for the EGARCH, FIEGARCH(BM) and HY/FIEGARCH(DL) models. The simulations for the memory properties of the HY/FIEGARCH(DL) process are provided in Section 2.7. A conclusion and suggestions for further research are presented in the final section. All proofs of this chapter are presented in Appendix A.

## 2.2 Literature review

This section critically reviews the theoretical and empirical literature on conditional heteroscedasticity models.

### 2.2.1 Relevant volatility models

As discussed in Chapter 1, the ARCH model was first introduced by Engle (1982) and extended to the GARCH model by Bollerslev (1986). Although the ARCH and GARCH models have been widely used in empirical applications, they have some drawbacks (see e.g. Nelson, 1991; Zakoïan, 1994; Bollerslev and Mikkelsen, 1996). These weaknesses were highlighted by Nelson (1991). First, GARCH processes require all coefficients to have non-negative values. Second, GARCH models are not able to capture the asymmetric properties of volatility. However, this is inconsistent with Black's (1976) demonstration that stock returns are negatively correlated with the volatility of returns, meaning that bad news leads to greater volatility than good news. Third, the GARCH models cannot capture the long-run dependence properties of financial data.

To overcome some of the shortcomings of GARCH processes, Nelson (1991) proposed the EGARCH model. For example, the EGARCH model can capture the asymmetric properties of shocks to volatility under certain conditions. The empirical applications of the GARCH and EGARCH models were reviewed by Bollerslev et al. (1992). Additional asymmetric models have subsequently been introduced. Glosten et al. (1993) found that conditional expected monthly returns were negatively correlated with the volatility of the monthly returns by applying a modified GARCH-M model. Their model is known as the GJR-GARCH model. Unlike other

ARCH models, in the threshold heteroscedastic model (TGARCH) introduced by Zakořian (1994), conditional variance is replaced by the conditional standard deviation. The reasons for including the conditional standard deviation are as follows: first, estimating the absolute residuals is more efficient than the squared residuals when the innovation has a non-normal distribution; second, the conditional standard deviation does not require all coefficients to be non-negative. This model is therefore capable of capturing the asymmetric properties of shocks to volatility, since both positive and negative shocks are included in the TGARCH model. However, Zakořian's study still assumed that all parameters are non-negative because the negative case is too difficult to analyse.

In order to investigate the high persistence property of volatility, Engle and Bollerslev (1986) proposed the IGARCH model and considered it to capture the effect of a shock on volatility remained forever. Researchers have suggested that the GARCH and IGARCH processes for conditional variance are similar to those seen in the  $I(0)$  and  $I(1)$  processes for the conditional mean. However, a puzzle has been identified regarding the persistence of volatility in the IGARCH model. Davidson (2004) showed that the IGARCH process has a geometric memory.

Taking the properties of the  $I(d)$  process in conditional mean models into account, the FIGARCH model was introduced by Baillie et al. (1996). They claimed that the effects of shocks on conditional variance dissipate at a hyperbolic rate in the FIGARCH model, which may be able to capture the long-run dependence properties of volatility. However, there is a debate over the memory properties of the FIGARCH model (see e.g. Giraitis et al., 2007). Bollerslev and Mikkelsen (1996) proposed the FIEGARCH(BM) model to analyse long memory in the volatility of financial time series, and the empirical results showed long-run dependence in the

volatility of the US stock market. The FIEGARCH(BM) model is more advanced than the FIGARCH model. It can capture the asymmetric properties and the long memory in volatility (see e.g. Bollerslev and Mikkelsen, 1996; Surgailis and Viano, 2002; Lopes and Prass, 2014). Subsequent empirical studies have demonstrated these properties of the FIEGARCH(BM) process. For instance, the stationary FIEGARCH(BM) process can be the best-fit model for the stock market when using the daily returns of the Tunisian stock market (Saadi et al., 2006); and Lopes and Prass (2014) demonstrated long-run dependence features in volatility applied to the Brazilian stock markets exchange index.

## 2.2.2 Literature on the moment property of conditional heteroscedasticity models

This section focuses on the theoretical properties of conditional heteroscedasticity models. Nelson (1990) investigated the stationary properties of GARCH-type models. He and Teräsvirta (1999a, b) also examined the moment conditions of conditional heteroscedasticity models. He and Teräsvirta (1999a) concentrated on the fourth moment structure of a family of GARCH(1, 1) models and obtained expressions for the fourth moment, kurtosis and the autocorrelation function of squared observations for these models. They investigated a unified framework of theoretical properties for seven different GARCH models<sup>1</sup>. He and Teräsvirta (1999b) demonstrated the conditions for the existence of the unconditional fourth moment of the higher-order GARCH process. However, they did not obtain the strictly stationary

---

<sup>1</sup>These models include the standard GARCH(1,1), absolute value GARCH (AVGARCH(1,1)), nonlinear GARCH(1,1), volatility switching GARCH(1,1), TGARCH(1,1), fourth-order nonlinear generalised moving-average conditional heteroscedasticity (4NLGMACH(1,1)), and generalised quadratic ARCH(GQARCH(1,1)) models.

condition for the GARCH-class models (see Ling and McAleer, 2002). Ling and McAleer (2002) investigated the structural features of GARCH-class models, and provided a sufficient condition for the existence of strictly stationary and ergodic processes and a necessary and sufficient condition for the existence of moments. He et al. (2008) showed how the autocorrelation function of squared and logarithmic observations may be obtained as a limiting case from the asymmetric power ARCH model. Their work proved that the autocorrelation function decays exponentially from the first lag.

Research has been carried out on the conditions for the existence of higher-order moments in GARCH models (e.g. Giraitis et al., 2000a; Carrasco and Chen, 2002). Carrasco and Chen (2002) pointed out that most GARCH-class models can be considered as generalised hidden Markov models. They focused on eight GARCH(1,1)-type models and derived sufficient conditions for the existence of higher-order moments; they also provided sufficient conditions for the existence of finite higher-order moments and a  $\beta$ -mixing process.

With regard to the EGARCH-type models, Nelson (1991) derived the autocorrelation function of the logarithm of the conditional variance and provided some conditions for the strict stationarity of the EGARCH model when the logarithmic conditional variance has an infinite moving average representation. Breidt et al. (1998) obtained the autocorrelation function of squared and logarithmic observations following the EGARCH model. He et al. (2002) also investigated the moment properties of the first-order standard EGARCH model and the symmetric and asymmetric logarithmic GARCH models. Their work derived conditions for the existence of moments, and expressions for kurtosis and the autocorrelation of positive powers of absolute-valued observations without assuming normal errors.

Ruiz and Veiga (2008) investigated the statistical properties of a new stochastic volatility model (A-LMSV) and the FIEGARCH(BM) model, and derived the kurtosis and autocorrelation function for each model. They also found that the kurtosis and correlation properties of absolute and squared returns are different, but they have similar features for cross-correlations between the returns and the power of absolute returns.

### **2.2.3 The memory property of conditional heteroscedasticity models**

The memory properties of conditional heteroscedasticity models have been considered in the literature. Firstly, the application of NED to measure the persistence of time series has been examined. The idea of NED was first introduced by Ibragimov (1962), and then developed by Gallant and White (1988), Andrews (1988) and Davidson (1994), among others. NED on a mixing process covers a large number of the time series; for example, Gallant and White (1988) showed that an autoregressive regression (AR)(1) process and ARMA with a finite order can be near-epoch dependent on shocks, and that the infinite ARCH process may be near-epoch dependent, based on some additional conditions. Hansen (1991) showed that GARCH (1, 1) processes are near-epoch dependent without assuming strict stationarity. Davidson (2002) also demonstrated that some nonlinear processes satisfy the condition of NED. Davidson (2004) proved that linear conditional heteroscedasticity processes are near-epoch dependent. He introduced hyperbolic memory and geometric memory based on the properties of the concept of NED, and showed that no long memory appears in the stationary FIGARCH and IGARCH processes.



Some of the previous studies indicated that the memory properties of conditional variance and conditional mean models are not parallel. The memory properties of the latter models have been well established. In the  $I(0)$  process, shocks die out at an exponential rate and the  $I(1)$  process has the longest persistence, whereas shocks decay at a hyperbolic rate in the  $I(d)$  process with  $0 < d < 1$ , see Granger (1980), Granger and Joyeux (1980), and Hosking (1981). However, in contrast to the  $I(1)$  process, the effect of shocks to volatility decay at an exponential rate in the IGARCH process, more details are presented in Ding and Granger (1996) and Davidson (2004). Ding and Granger (1996) found higher persistence in the fractional ordered model than in the integrated model for foreign exchange rate returns. This differs from Ding et al. (1993), who found that there is the strongest persistence with an integrated order of 1. They also proved that the theoretical autocorrelation functions display exponential decay properties in various GARCH (1, 1) processes, including the IGARCH model. Davidson (2004) also proved that the IGARCH process has geometric memory by using the  $L_0$ -approximable method. Therefore, the memory property of the conditional mean equation might be different from that of the corresponding conditional variance equation (see e.g. Ding et al., 1993; Davidson, 2004).

#### **2.2.4 The FCLT and fractional FCLT of conditional heteroscedasticity models**

Several methods have been introduced to derive the FCLT for time series models. For example, the mixing condition, weak dependence, association and NED can be used to derive the FCLT (see e.g. Hansen, 1991; Hörmann, 2008; Davidson, 2002; Lee, 2013, 2014).

NED is useful for establishing the FCLT and fractional FCLT for time series. Andrews (1988) introduced the idea that NED processes satisfy the conditions of the weak LLN. Hansen (1991) then applied the NED concept to show that GARCH(1,1) models satisfy the conditions for the weak and strong LLN, and CLT, and established invariance principles for the GARCH(1,1) process. Davidson (2002) derived sufficient conditions for the CLT and FCLT to be satisfied in nonlinear processes and semiparametric linear processes. Taking advantage of NED on a mixing process, he established the FCLT for a group of nonlinear processes, including GARCH, bilinear and threshold autoregressive models, satisfying the conditions of NED on an independent process. He also proved a new FCLT for semi-parametric linear processes, which belongs to the concept of NED on a mixing process. Lee (2013) showed that both of the stationary GARCH process under a second moment condition and the stationary ARMA-GARCH under the finite second moment conditions are geometrically  $L_2$ -NED. He also established the FCLT for both models under weaker conditions, leading to no further restrictions on the distributional assumptions of errors and higher-order moments

Berkes et al. (2008) derived the FCLT for an augmented GARCH model under a finite second-moment condition using Theorem 21.1 from Billingsley (1968) (see Lee, 2013). Lee (2014) also derived the FCLT for the augmented GARCH( $p, q$ ) process by applying the NED approach, and provided the FCLT for the EGARCH process as a special case of the augmented GARCH( $p, q$ ) process. However, the results from Lee (2014) did not consider the hyperbolic memory case of the EGARCH-type models. Surgailis and Viano (2002) also considered a general stochastic volatility model, including the EGARCH and FIEGARCH(BM) models, in which they investigated the covariance structure and dependence properties of these models. The FCLT for the short and moderate memory of EGARCH models was estab-

lished in Surgailis and Viano’s (2002) study.

Davydov (1970) introduced the fractional FCLT for partial sums of the fractionally integrated process, such that  $x_t = (1 - L)^{-d}u_t$ , where  $u_t$  is an i.i.d. process with a mean of zero, and provided a moment condition of  $E|u_t|^q < \infty$  with  $q \geq 4$  and  $q > -4d/(d + 1/2)$  for the fractional FCLT of this process; Taqqu (1975) provided a weaker moment condition, such that,  $E|u_t|^q < \infty$  with  $q \geq 2$  and  $q > (d + 1/2)^{-1}$ , see e.g. Johansen and Nielsen (2012). Davidson and De Jong (2000) proved the weaker moment conditions for the fractional FCLT for some NED processes in their Theorem 3.1. These previous studies were reviewed by Johansen and Nielsen (2012), who pointed out that the weaker moments conditions may not be sufficient to support the fractional FCLT in some cases. They also provided a necessary moment condition for the fractional FCLT. Wu and Shao (2006) investigated the FCLT for a class of fractionally integrated nonlinear processes. Lee (2014) discussed the fractional FCLT for the FIGARCH processes by applying the NED approach.

## 2.3 Exponential-type conditional heteroscedasticity models

Nelson (1991) introduced the exponential ARCH ( $\infty$ ) (EARCH ( $\infty$ )) model, which is defined by Equations (1.1) and (2.1):

$$\ln h_t = \omega + \sum_{j=1}^{\infty} \beta_j g(z_{t-j}), \quad \beta_1 = 1, \quad (2.1)$$

$$g(z_t) = \theta z_t + \gamma[|z_t| - E|z_t|], \quad (2.2)$$

where the lag coefficients  $\{\beta_j\}$ , for all  $j \geq 1$ , are lag coefficients and real, non-stochastic sequences, and can be negative or positive values, except  $j = 1$ . The properties of the  $g(z_t)$  function decide whether the EARCH( $\infty$ ) model can capture the asymmetric feature of shocks to volatility or not. In other words, the well-defined process  $g(z_t)$  needs to be a function that includes the magnitude and the sign of  $z_t$ , such as Equation (2.2) (Nelson, 1991). The  $g(z_t)$  function is composed of two terms:  $\theta z_t$  and  $\gamma[|z_t| - E|z_t|]$ ; these are orthogonal when  $z_t$  is symmetrically distributed.

$$g(z_t) = \begin{cases} (\theta - \gamma)z_t - \gamma[|E|z_t|] & z_t < 0 \\ (\theta + \gamma)z_t - \gamma[|E|z_t|] & z_t \geq 0. \end{cases} \quad (2.3)$$

It is clear that the function  $g(z_t)$  has the slope  $\theta - \gamma$  when the sign of  $z_t$  is negative, and the slope  $\theta + \gamma$  when the sign of  $z_t$  is non-negative. The sequence of  $g(z_t)$  is a linear combination of  $\{z_t\}$ , and the expectation value of  $g(z_t)$  is zero. The variance of the process  $g(z_t)$  depends on the distribution of  $\{z_t\}$  as follows:

$$\begin{aligned} E[g(z_t)^2] &= E[(\theta z_t + \gamma[|z_t| - E|z_t|])^2] \\ &= \theta^2 E(z_t)^2 + \gamma^2 E|z_t|^2 + \gamma^2 (E|z_t|)^2 - 2\gamma^2 (E|z_t|)^2 + 2\theta\gamma E(z_t|z_t|) \\ &= \theta^2 + \gamma^2 - \gamma^2 (E|z_t|)^2 + 2\theta\gamma E(z_t|z_t|), \end{aligned}$$

and since:

$$E(z_t|z_t|) \leq E(z_t^2) = 1 \text{ for all } t \in Z.$$

In this case,  $E[g(z_t)^2]$  is finite. For similar results, see Prass (2008). Nelson (1991) also derived a stationarity condition and a condition for the existence of moments of the innovation processes in his Theorem 2.1.

Nelson (1991) also defined the EGARCH( $p, q$ ) model, which is:

$$\ln h_t = \omega + \frac{\psi(L)}{\varphi(L)} g(z_{t-1}),$$

where  $L$  denotes the back-shift operator, and  $\psi(L) = 1 + \psi_1 L + \dots + \psi_q L^q$  and  $\varphi(L) = 1 - \varphi_1 L - \dots - \varphi_p L^p$  are polynomials in the lag operator. Let us now suppose that the EGARCH( $p, q$ ) process satisfies the following assumption.

**Assumption 2.3.1** *Assume that  $|\psi(z)| > 0$  and  $|\varphi(z)| > 0$  for any  $|z| \leq 1$ , and that  $\psi(L)$  and  $\varphi(L)$  have no common roots.*

Based on a similar idea to Theorem 7.2.3(i) in Giraitis et al. (2012) for the ARMA( $p, q$ ) model, we can rewrite:

$$\frac{\psi(L)}{\varphi(L)} = \sum_{j=0}^{\infty} \beta_j L^j, \quad (2.4)$$

with  $\beta_0 = 0$ . Under Assumption 2.3.1,  $|z| \leq 1$ , then for some  $\epsilon > 0$ , we have:

$$\left| \frac{\psi(z)}{\varphi(z)} \right| < \infty \text{ for any } |z| \leq 1 + \epsilon.$$

The convergence of  $\psi(z)/\varphi(z)$  means that:

$$|\beta_j z^j| \leq |\beta_j| (1 + \epsilon)^j \rightarrow 0,$$

as  $j$  tends to infinity. This also means that:

$$|\beta_j| \leq C(1 + \epsilon)^{-j}, \text{ where } C \text{ is a constant and } j > 1. \quad (2.5)$$

If we set  $1 + \epsilon = \rho$ , then the EGARCH( $p, q$ ) process can be written as the EARCH( $\infty$ ) model:

$$\ln h_t = \omega + \sum_{j=0}^{\infty} \beta_j g(z_{t-j}),$$

with  $|\beta_j| \leq C\rho^{-j}$ , where  $C$  is a positive constant,  $j > 1$ , and  $\rho > 1$ . Nelson (1991) also proposed the EGARCH(1, 1) model, which can be written as Equations (1.1) and (2.6):

$$\ln h_t = w + g(z_{t-1}) + \beta \ln h_{t-1}. \quad (2.6)$$

This can be reorganised as the EARCH( $\infty$ ) model, such that:

$$\ln h_t = \omega + \sum_{j=1}^{\infty} \beta^{j-1} g(z_{t-j}), \quad (2.7)$$

where  $|\beta| < 1$  and  $\omega = w/(1 - \beta)$ . Bollerslev and Mikkelsen (1996) extended the EGARCH model to the FIEGARCH( $p, d, q$ ) model, which is defined as:

$$\begin{aligned} \ln h_t &= \omega + \frac{\psi(L)}{\varphi(L)} (1 - L)^{-d} g(z_{t-1}) \\ &= \omega + \frac{\sum_{i=1}^q \psi_i L^i}{1 - \sum_{i=1}^p \varphi_i L^i} (1 - L)^{-d} g(z_{t-1}) \\ &= \omega + \sum_{j=0}^{\infty} \beta_{-d,j} L^j g(z_{t-1}), \end{aligned} \quad (2.8)$$

where:

$$(1 - L)^{-d} = \sum_{j=0}^{\infty} \frac{\Gamma(j + d)}{\Gamma(d)\Gamma(j + 1)}, \quad (2.9)$$

where  $p, q$  are non-negative integers and  $(1 - L)^{-d}$  is the fractional differencing operator. The FIEGARCH(BM) model becomes an EGARCH model when  $d$  equals 0. The pure FIEGARCH(BM) model is defined as:

$$\begin{aligned} \ln h_t &= \omega + (1 - L)^{-d} g(z_{t-1}) \\ &= \omega + \sum_{j=0}^{\infty} \beta_{-d,j} L^j g(z_{t-1}). \end{aligned} \quad (2.10)$$

The property of the lag coefficients in the FIEGARCH(BM) model can be investigated, similar to Giraitis et al.'s (2012) Theorem 7.2.3 (ii) for the AR fractional integrated MA (ARFIMA( $p, d, q$ )) model. In the FIEGARCH(BM) model, the memory parameter  $d$  may be either positive or negative, and  $|\beta_{-d,j}| \leq Cj^{d-1}$  with  $C > 0$ .

## 2.4 Moment properties of the exponential-type conditional heteroscedasticity models

This section first considers the finite-order moments of the general EARCH( $\infty$ ) processes. Davidson (2004) emphasised that the amplitude property is a significant feature of volatility models. Amplitude is used to show how large the fluctuations in conditional variance and shocks to volatility may be. Amplitude is measured by the sum of the lag coefficients:

$$S = \sum_{j=1}^{\infty} \beta_j.$$

For the general EARCH( $\infty$ ) processes, Nelson (1991) derived that  $\sum_{j=1}^{\infty} \beta_j^2 < \infty$  is both a covariance stationary condition and a strictly stationary condition for the process  $\{\ln(h_t) - \omega\}$ , but it is only a strictly stationary and ergodic condition for  $\{\exp(-\omega)h_t\}$ . In the simplest EGARCH (1, 1) model, when  $|\beta| < 1$ , the sum of the lag coefficients can be defined as:

$$S = \frac{1}{1 - \beta}. \quad (2.11)$$

The investigation of the moment property of the EARCH( $\infty$ ) model is as follows. First, set:

$$M_p = E[\xi_t^p], \quad (2.12)$$

where  $M_p$  does not depend on  $t$ , and  $p$  denotes the order of the moment. Next, we consider the second moment of  $\xi_t$ :

$$M_2 = E[\xi_t^2] = E[h_t z_t^2], \quad (2.13)$$

when  $p = 2$ , and the processes of  $\sqrt{h_t}$  and  $z_t$  are independent. Then:

$$E[h_t z_t^2] = E[h_t]E[z_t^2] = E[h_t], \text{ since } z_t \sim i.i.d.(0, 1). \quad (2.14)$$

This equation means that the second moment of  $\xi_t$  is equal to the expectation value of the conditional variance  $h_t$ . The second moment of  $\xi_t$  can now be obtained by:

$$\begin{aligned}
E[\xi_t^2] &= E[h_t] = E \left[ \exp \left( \omega + \sum_{j=1}^{\infty} \beta_j g(z_{t-j}) \right) \right] & (2.15) \\
&= E \left[ \exp(\omega) \cdot \exp \left( \sum_{j=1}^{\infty} \beta_j g(z_{t-j}) \right) \right] \\
&= \exp(\omega) E \left[ \exp \left( \sum_{j=1}^{\infty} \beta_j g(z_{t-j}) \right) \right] \\
&= \exp(\omega) \prod_{j=1}^{\infty} E[\exp(\beta_j g(z_{t-j}))].
\end{aligned}$$

The final equation is obtained because  $g(z_t)$  is an i.i.d. process. An equivalent expression can be written as:

$$E[\exp(-\omega)h_t] = \prod_{j=1}^{\infty} E[\exp(\beta_j g(z_{t-j}))]. \quad (2.16)$$

Nelson (1991) showed that the condition of strict stationarity and ergodicity is  $\sum_{j=1}^{\infty} \beta_j^2 < \infty$  for the process  $\{\exp(-\omega)h_t\}$ . However, this does not mean that the processes are covariance stationary. In other words, a finite order of the unconditional moment may not exist. Thus, a further assumption about the distribution of  $\{z_t\}$  is required to derive the unconditional expectation of  $\exp(\beta_j g(z_{t-j}))$ .

In this chapter, the general generalised error distribution (GED( $v$ )) will be considered as the distribution of  $\{z_t\}$ . Following Nelson (1991), the density function of the GED, with a mean of zero and a variance of one, is defined as:

$$f(z) = \frac{v \exp[-(\frac{1}{2}|z/\lambda|^v)]}{\lambda 2^{(1+1/v)} \Gamma(1/v)}, \quad (2.17)$$

where  $-\infty < z < +\infty$ ,  $0 < v < \infty$ ,  $v$  denotes the the tail-thickness parameter,  $\Gamma(\cdot)$  is the gamma function and:

$$\lambda = \left[ \frac{2^{(-2/v)} \Gamma(1/v)}{\Gamma(3/v)} \right]^{1/2}. \quad (2.18)$$



The GED can exhibit a different distribution with a different value of  $v$ . The normal distribution is a special case of the GED when the tail-thickness parameter  $v$  equals 2. The tail of the distribution of  $z_t$  is thicker than the normal distribution when  $v$  is less than 2 and is thinner than the normal distribution when  $v$  is greater than 2.

In the first place, let us consider the special case of GED(2), assuming that  $\{z_t\}$  is an i.i.d.  $N(0, 1)$  and setting  $g(z_t)$  as Equation (2.2). It can be seen that  $g(z_t)$  is a linear combination of  $z_t$  and  $|z_t|$ , and can include both the magnitude and the sign of the shocks depending on the parameters  $\theta$  and  $\gamma$ . Therefore, the leverage effect of the process can be performed by  $g(z_t)$ . It can also be seen that the process  $\{g(z_t)\}$  is an i.i.d. random sequence with a mean of zero. According to Nelson (1991, Theorem A1.1), we can derive the following:

$$E[\exp(g(z_{t-j})b)] < \infty, \quad (2.19)$$

where  $b$  is one of the lag coefficients  $\beta_j$ . We then have:

$$\prod_{j=1}^{\infty} E[\exp(\beta_j g(z_{t-j}))] < \infty. \quad (2.20)$$

and then:

$$E[\xi_t^2] = E[h_t] = \exp(\omega) \prod_{j=1}^{\infty} E[\exp(\beta_j g(z_{t-j}))] < \infty. \quad (2.21)$$

Thus the second moment of  $\xi_t$  exists. Secondly, considering the fourth moment case, let:

$$E[z_t^4] = \mu_4. \quad (2.22)$$

Since the processes  $z_t$  and  $\sqrt{h_t}$  are independent, the fourth moment is:

$$E[\xi_t^4] = E[z_t^4 h_t^2] = E[z_t^4] E[h_t^2] = \mu_4 \cdot E[h_t^2], \quad (2.23)$$

where:

$$\begin{aligned} E[h_t^2] &= E \left[ \exp \left( \omega + \sum_{j=1}^{\infty} \beta_j g(z_{t-j}) \right) \right]^2 \\ &= \exp(2\omega) \left\{ \prod_{j=1}^{\infty} E[\exp(\beta_j g(z_{t-j}))] \right\}^2. \end{aligned} \quad (2.24)$$

Similarly, according to Nelson (1991), the inequality (2.20) holds, then:

$$\left\{ \prod_{j=1}^{\infty} E[\exp(\beta_j g(z_{t-j}))] \right\}^2 < \infty, \quad (2.25)$$

and since it is assumed that  $\mu_4 < \infty$ . These ensure that the fourth moment of  $\xi_t$  is finite. Equation (2.24) can be extended to the  $n^{\text{th}}$  ( $n < \infty$ ) moment of  $h_t$ :

$$\begin{aligned} E[h_t^n] &= E \left[ \exp \left( \omega + \sum_{j=1}^{\infty} \beta_j g(z_{t-j}) \right) \right]^n \\ &= \exp(n\omega) \left[ \prod_{j=1}^{\infty} E(\exp(\beta_j g(z_{t-j}))) \right]^n \\ &< \infty, \end{aligned} \quad (2.26)$$

assuming that:

$$\mu_n < \infty. \quad (2.27)$$

The  $n^{\text{th}}$  moment also exists for EARCH( $\infty$ ) models. This condition also supports the NED of the EARCH( $\infty$ ) processes, which will be explained in the next section. The discussion above is based on the assumption that the process  $\{z_t\}$  has a normal distribution. For the general GED case, Nelson (1991, A 1.2) showed that if we assume the process  $\{z_t\} \sim \text{i.i.d.GED}(v)$  with a mean of zero and a variance of one, then:

$$E[z_t^n \exp(g(z_t)b)] < \infty, \quad (2.28)$$

when  $v$  is greater than 1 and without other restrictions. However, when  $v$  is less than 1, it must satisfy  $b\gamma + |b\theta| \leq 0$ , which is the requirement for the existence of a finite order of moment. If  $v$  equals 1, it needs  $b\gamma + |b\theta| \leq \sqrt{2}$  to meet the

requirement for the existence of a moment. For the EGARCH(1,1) model, He et al. (2002) derived conditions for the existence of the second and fourth moments.

## 2.5 Memory properties of the exponential-type conditional heteroscedasticity models

Research on the memory property of conditional heteroscedasticity models is crucial for volatility forecasting. Davidson (2004) investigated the memory properties of linear conditional heteroscedasticity models by applying the NED concept. With regard to the memory characteristic of ARCH-type models, he introduced two kinds of memory: hyperbolic and geometric memory. Motivated by Davidson (2004), this section aims to derive conditions for the process  $\{\xi_t\}$  following the EARCH( $\infty$ ) model is  $L_p$ -NED on  $\{z_t\}$ , for either  $p = 1$  or  $p = 2$ , and to investigate the memory properties of EGARCH-type models by applying the NED approach. Moreover, considering the memory properties of these models, this section also introduces a general expression of the exponential-type conditional heteroscedasticity model – HY/FIEGARCH(DL).

### 2.5.1 Memory properties of the EGARCH-type models

The notation used in this section is as follows. The parameter  $a$  is used to measure the hyperbolic memory, where:

$$\beta_j = O(j^{-1-a}). \quad (2.29)$$

The parameter  $\rho$  is used to measure the geometric memory, such that:

$$\beta_j = O(\rho^{-j}). \quad (2.30)$$

Moreover, we denote the  $L_p$  norm of a random variable  $X_t$  as  $\|X_t\|_p = (E|X_t|^p)^{1/p}$ .

In order to study the memory properties of these models, two moment inequalities in Lemma 2.5.2 of Giraitis et al. (2012) are essential to be recalled.

**Lemma 2.5.1** (*Lemma 2.5.2 of Giraitis et al., 2012*) *Let  $p \geq 1$  and  $\{Y_t, F_j, 1 \leq j \leq n\}$  be a martingale difference sequence with  $E|Y_j|^p < \infty$ . For every  $n \geq 1$ , when  $1 \leq p \leq 2$ :*

$$E \left| \sum_{j=1}^n Y_j \right|^p \leq 2 \sum_{j=1}^n E |Y_j|^p.$$

When  $p > 2$ ,

$$E \left| \sum_{j=1}^n Y_j \right|^p \leq C_p \left( \sum_{j=1}^n (E |Y_j|^p)^{2/p} \right)^{p/2},$$

with a constant  $C_p > 0$  depending only on  $p$ . These inequalities are valid for  $n = \infty$  as long as the series on the right-hand side of them are convergent.

The idea of how to show this lemma is based on the von Bahr and Esseen (1965) inequality for  $1 \leq p \leq 2$  and Rosenthal's inequality (see Hall and Heyde, 1980). The proof of this lemma is provided in Giraitis et al. (2012).

And following, the definition of NED is given in Definition 2.5.1.

**Definition 2.5.1** *Assume that  $\{\xi_t\}$  is a function of the whole past and future history of process  $\{z_t, -\infty < t < +\infty\}$ . Set  $\mathcal{F}_{t-m}^{t+m} = \sigma\{z_{t-m}, \dots, z_{t+m}\}$ , the  $\sigma$ -field generated by the near epoch of the process  $\{z_t\}$ . The process  $\{\xi_t\}$  is then said to*

be near-epoch dependent in the  $L_p$ -norm ( $L_p$ -NED) on  $\{z_t\}$  of size<sup>2</sup>  $-a_0$  if:

$$\|\xi_t - E(\xi_t | \mathcal{F}_{t-m}^{t+m})\|_p \leq d_{pt} v_m(a) \quad \text{for } p > 0, \quad (2.31)$$

where  $\{d_{pt}\}$  is a sequence of positive constants,  $v_m(a) = m^{-a}$  for  $a > a_0$ , and  $m^{-a} \rightarrow 0$  as  $m \rightarrow \infty$ .

This definition is drawn from Davidson (1994, Definition 17.1). Definition 2.5.1 defines hyperbolic NED. The process can be considered as geometric NED if the term  $v_m(a) = m^{-a}$  is replaced by  $v_m(\rho) = \rho^{-m}$  (see e.g. Andrews, 1988; Davidson, 2004). The terms  $m^{-a}$  and  $\rho^{-m}$  both converge to 0 but at different decay rates, as  $m$  tends to infinity. Davidson (2002, 2004) verified that the  $\xi_t$  process is  $L_p$ -NED on  $\{z_t\}$ , for either  $p = 1$  or  $p = 2$ , for GARCH-class models. Following a similar procedure, this section shows the NED properties of EARCH-type models.

In the EARCH( $\infty$ ) model, because  $\xi_t = \sqrt{h_t} z_t$ , and  $z_t \sim \text{i.i.d.}(0, 1)$ , then:

$$\begin{aligned} \|\xi_t - E[\xi_t | \mathcal{F}_{t-m}^{t+m}]\|_p &= \left\| \sqrt{h_t} z_t - E\left[\sqrt{h_t} z_t | \mathcal{F}_{t-m}^{t+m}\right] \right\|_p \\ &= \|z_t\|_p \left\| \sqrt{h_t} - E\left[\sqrt{h_t} | \mathcal{F}_{t-m}^{t+m}\right] \right\|_p \\ &\leq \left\| \sqrt{h_t} - E\left[\sqrt{h_t} | \mathcal{F}_{t-m}^{t+m}\right] \right\|_p. \end{aligned} \quad (2.32)$$

The above inequality is obtained by:

$$\|z_t\|_p \leq 1, \quad (2.33)$$

where  $p = 1$  or  $p = 2$ . Therefore, if:

$$\left\| \sqrt{h_t} - E\left[\sqrt{h_t} | \mathcal{F}_{t-m}^{t+m}\right] \right\|_p \leq d_{pt} v_m(a) \quad (\text{or } d_{pt} v_m(\rho)), \quad (2.34)$$

---

<sup>2</sup>The terminology "size" can be used to denote the decay rate of the mixing numbers. It is also applied to describe a sequence which is  $\alpha$ -mixing of size  $-\varphi_0$  if  $\alpha_m = O(m^{-\varphi})$  for some  $\varphi > \varphi_0$ . (see e.g. White, 1984; Davidson, 1994)

then:

$$\|\xi_t - E[\xi_t | \mathcal{F}_{t-m}^{t+m}]\|_p \leq d_{pt} v_m(a) \text{ (or } d_{pt} v_m(\rho)), \quad (2.35)$$

respectively. By Equation (2.1), the conditional variance can be written as:

$$h_t = \exp\left(\omega + \sum_{j=1}^{\infty} \beta_j g(z_{t-j})\right), \beta_1 = 1. \quad (2.36)$$

Then,

$$\sqrt{h_t} = \left[ \exp\left(\frac{1}{2}\omega + \frac{1}{2} \sum_{j=1}^{\infty} \beta_j g(z_{t-j})\right) \right]. \quad (2.37)$$

Based on Equations (2.36) and (2.37), it is straightforward to see that if:

$$\|h_t - E[h_t | \mathcal{F}_{t-m}^{t+m}]\|_p \leq d_{pt} v_m(a) \text{ (or } d_{pt} v_m(\rho)), \quad (2.38)$$

where  $d_{pt} < \infty$ , for  $p = 1$  or  $p = 2$ , then Equation (2.34) holds, and vice versa. Furthermore, Equation (2.35) also holds. This means that if the processes  $\{h_t\}$  or  $\{\sqrt{h_t}\}$  are  $L_p$ -NED on  $\{z_t\}$ , then the process  $\{\xi_t\}$  is also  $L_p$ -NED on  $\{z_t\}$ , for either  $p = 1$  or  $p = 2$ . In other words, to obtain the NED properties of  $\{\xi_t\}$ , it is essential to prove the inequalities of (2.38) or (2.34).

The general EARCH( $\infty$ ) process is given in Equation (2.36). First, denote:

$$G_t = \sum_{j=1}^{\infty} \beta_j g(z_{t-j}). \quad (2.39)$$

Here, the expression of the  $L_p$ -NED of the conditional variance is:

$$\|h_t^k - E[h_t^k | \mathcal{F}_{t-m}^{t+m}]\|_p = \|\exp(k(\omega + G_t)) - E[\exp(k(\omega + G_t)) | \mathcal{F}_{t-m}^{t+m}]\|_p,$$

where  $p = 1, 2$ ,  $k = 1/2$  or  $k = 1$ . Since, by the Liapunov's inequality<sup>3</sup>,

$$E|h_t^k - E[h_t^k | \mathcal{F}_{t-m}^{t+m}]| \leq \|h_t^k - E[h_t^k | \mathcal{F}_{t-m}^{t+m}]\|_2,$$

---

<sup>3</sup>This is also called norm inequality: If  $a > b > 0$ , then  $\|X\|_a \geq \|X\|_b$ . See e.g. Davidson (1994).

and the  $L_2$ -NED plays a more important role for the limit theorem of the processes, thus this study mainly focuses on the  $L_2$ -NED of  $\{h_t^k\}$ .

Based on the moment conditions and the properties of EGARCH-type models, the memory properties of the EGARCH-type models can be seen from Theorems 2.5.2 and 2.5.3. Firstly, the hyperbolic property of the EARCH( $\infty$ ) model can be derived.

**Theorem 2.5.2** *If  $|\beta_j| \leq Cj^{-1-a}$  for  $j \geq 1$ ,  $C > 0$ ,  $a > 0$ ,  $\|g(z_t)\|_q < \infty$ , and  $\|h_t^{2k}\|_{q/(q-1)} < \infty$ , then:*

$$\|h_t^k - E(h_t^k | \mathcal{F}_{t-m}^{t+m})\|_2 \leq d_{2t} v_m(a), \quad (2.40)$$

where  $k = 1/2$  or  $k = 1$  and  $\{d_{2t}\}$  is a sequence of positive constants. Therefore, if  $1 < q \leq 2$ ,  $\{h_t^k\}$  and  $\xi_t$  are  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2} \left( \frac{1}{q} - 1 - a_0 \right)$ , where ( $a > a_0 > 0$ ); if  $q > 2$ ,  $\{h_t^k\}$  and  $\{\xi_t\}$  are  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2} \left( -a_0 - \frac{1}{2} \right)$ , where ( $a > a_0 > 0$ ). When  $-1/2 < a < 0$ ,  $\{h_t^k\}$  and  $\{\xi_t\}$  are  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2} \left( \frac{1}{q} - 1 - a_0 \right)$ , where  $\left( \frac{1}{2} \left( 1 + a - \frac{1}{q} \right) > a_0 > 0 \right)$  if  $1 < q \leq 2$ ; and of size  $\frac{1}{2} \left( -a_0 - \frac{1}{2} \right)$ , where  $\left( \frac{1}{2} \left( a + \frac{1}{2} \right) > a_0 > 0 \right)$  if  $q > 2$ .

According to the properties of the lag coefficients, the FIEGARCH(BM) model is a case of hyperbolic processes when  $-1/2 < d < 0$ , where  $d$  is same as  $-a$ . However, the memory parameter may also have a positive value in the FIEGARCH(BM) model. When the memory parameter is  $0 < d < 1/2$ , this model is able to capture long memory because of the non-summable properties of the absolute lag coefficients.

The following theorem shows that the EARCH( $\infty$ ) model has geometric memory.

**Theorem 2.5.3** *If  $|\beta_j| \leq C\rho^{-j}$  for  $j \geq 1$ ,  $C \geq \rho > 1$ ,  $\|g(z_t)\|_q < \infty$ , and  $\|h_t^{2k}\|_{q/(q-1)} < \infty$ , for  $q > 1$ , then:*

$$\|h_t^k - E[h_t^k | \mathcal{F}_{t-m}^{t+m}]\|_2 \leq \acute{d}_{2t} v_m(\rho), \quad (2.41)$$

where  $k = 1/2$  or  $k = 1$ , and  $\{\acute{d}_{2t}\}$  is a sequence of positive constants. Therefore,  $\{h_t^k\}$  and  $\{\xi_t\}$  are geometrically  $L_2$ -NED.

According to this theorem, the process  $\{\xi_t\}$  in EARCH( $\infty$ ) processes can be geometrically  $L_2$ -NED. The geometric NED means that the impacts of shocks on current or future volatility decay at a geometric rate. Therefore, EARCH( $\infty$ ) models can capture geometric memory. The simplest example is the EGARCH(1, 1) model, since  $\beta_j = \beta^{j-1}$  and  $|\beta| < 1$ , then it has geometric memory. Similar results can also be held for the EGARCH( $p, q$ ) process, by the condition (2.5).

## 2.5.2 The HYGARCH and FIEGARCH(DL) models

The EARCH( $\infty$ ) processes can capture geometric, hyperbolic and long memory in volatility. In order to distinguish the memory properties of EGARCH-type models, a general framework of the HY/FIEGARCH(DL)( $p, d_2, q$ ) model is used, which is defined as:

$$\ln h_t = \omega + \theta(L)g_1(z_t), \quad (2.42)$$

where<sup>4</sup>  $g_1(z_t) = |z_t| - E|z_t| + \theta z_t$  and:

$$\theta(L) = \left(1 - \frac{\delta(L)(1 + \alpha((1-L)^{d_2} - 1))}{\varphi(L)}\right), \quad (2.43)$$

where  $\delta(L) = 1 - \delta_1 L + \dots + \delta_p L^p$ , and  $\varphi(L) = 1 - \varphi_1 L - \dots - \varphi_p L^p$ , and:

$$(1-L)^{d_2} = 1 - \sum_{j=1}^{\infty} b_j L^j,$$

---

<sup>4</sup>This is because of the parameters in  $\theta(L)$ , the coefficient of  $|z_t| - E|z_t|$ , is defined as 1.



with:

$$b_j = \begin{cases} \frac{d_2 \Gamma(j-d_2)}{\Gamma(1-d_2)\Gamma(j+1)} & d_2 \geq 0 \\ \frac{-\Gamma(j-d_2)}{\Gamma(-d_2)\Gamma(j+1)} & d_2 < 0 \end{cases}, \quad j \geq 1.$$

And the Stirling's approximation of  $b_j$  is given as:

$$b_j = O(j^{-1-d_2}).$$

The HY/FIEGARCH(DL) process can be also written as:

$$\ln h_t = \omega + \left(1 - \frac{\delta(L)}{\varphi(L)} (1 + \alpha ((1-L)^{d_2} - 1))\right) g_1(z_t).$$

It is an EGARCH( $p, q$ ) process with  $\gamma = 1$  when  $d_2 = 0$ . The pure HY/FIEGARCH(DL) is:

$$\begin{aligned} \ln h_t &= \omega + (1 - (1 + \alpha ((1-L)^{d_2} - 1)))g_1(z_t) & (2.44) \\ &= \omega + \alpha(1 - (1-L)^{d_2})g_1(z_t) \\ &= \omega + \alpha \left(1 - \left(1 - \sum_{j=1}^{\infty} b_j L^j\right)\right) g_1(z_t) \\ &= \omega + \alpha \sum_{j=1}^{\infty} b_j g_1(z_{t-j}). \end{aligned}$$

Two distinct cases exist:

**Case 1** : *HYEGARCH occurs when  $0 < d_2 < \frac{1}{2}$  and  $\alpha > 0$ , and thus  $\alpha b_j > 0$ ,  $j \geq 1$ , with the lag coefficients being summable.*

**Case 2** : *FIEGARCH(DL) occurs when  $-\frac{1}{2} < d_2 < 0$  and  $\alpha < 0$ , and thus  $\alpha b_j > 0$ ,  $j \geq 1$ , with the lag coefficients not being summable.*

It is worth noting that  $d_2$  in the HY/FIEGARCH(DL) process is different from  $d$  in the FIEGARCH(BM) process, and  $d_2$  has the same sign as  $a$  in Theorem 2.5.2. This means that HYEGARCH processes have hyperbolic memory and FIEGARCH(DL) processes can capture long memory in volatility.

## 2.6 FCLT and fractional FCLT for the EGARCH-type models

To derive the main results of the FCLT and fractional FCLT for the relevant processes in the EGARCH-type models, several important theorems must first be recalled. Let the stochastic process  $X_n : [0, 1] \rightarrow R$  be defined as:

$$X_n(\kappa) = \sigma_n^{-1} \sum_{t=1}^{[n\kappa]} (x_t - Ex_t), 0 < \kappa \leq 1, \quad (2.45)$$

where  $\sigma_n^2 = Var(\sum_{t=1}^n x_t)$  and  $[x]$  denotes the largest integer not exceeding  $x$ . If  $X_n$  of Equation (2.45) converges weakly to Brownian motion, then the process  $x_t$  holds the FCLT. And this yields a CLT when  $\kappa = 1$ . Recall that Davidson's (2002) Theorem 1.2, which proposes Assumptions A1-A3:

- A1.  $x_t$  is  $L_2$ -NED of size  $-1/2$  on the underlying i.i.d. process  $\{z_s\}$  with respect to the constants  $d_t \leq \|x_t\|_r$ ;
- A2.  $\sup_t E|x_t - Ex_t|^r < \infty$ , for  $r \geq 2$ ; if  $r = 2$ , then  $\{(x_t - Ex_t)^2\}$  is uniformly integrable;
- A3.  $\sigma_n^2/n \rightarrow \sigma^2 > 0$  as  $n \rightarrow \infty$ .

Let  $\rightarrow^d$  denote the convergence in distribution and let  $B$  denote standard Brownian motion on  $[0, 1]$ .

**Theorem 2.6.1** *(Theorem 1.2 in Davidson (2002)) If Conditions A1–A3 in Assumption A hold, then  $X_n \rightarrow^d B$ . This means that  $X_n$  converges in distribution to standard Brownian motion.*

To derive Condition A3, we can use Theorem 21.1 of Billingsley (1968), which is recalled as follows.

**Theorem 2.6.2** (*Theorem 21.1 in Billingsley (1968)*) Suppose that  $\{z_t\}$  is  $\varphi$ -mixing with  $\sum \varphi_t^{\frac{1}{2}} < \infty$  and that the  $\eta_t$ , as defined by  $\eta_t = f(\dots, z_{t-1}, z_t, z_{t+1}, \dots)$ ,  $n = 0, \pm 1, \pm 2, \dots$ , (where  $z_t$  occupies the  $0^{\text{th}}$  place in the argument of  $f$ ) have mean 0 and finite variance. Suppose further that there are random variables of the form:

$$\eta_{m,t} = f_m(z_{t-m}, \dots, z_t, \dots, z_{t+m}) = E\{\eta_t | \mathcal{F}_{t-m}^{t+m}\},$$

where the subscript  $m, t$  indicates a pair, not a product, such that:

$$\sum_{m=1}^{\infty} (E\{|\eta_0 - \eta_{m,0}|^2\})^{\frac{1}{2}} < \infty.$$

In this case, the series:

$$\sigma^2 = E\{\eta_0^2\} + 2 \sum_{k=1}^{\infty} E\{\eta_0 \eta_k\}$$

converges absolutely. If  $\sigma^2 > 0$  and  $X_n$  is defined by Equation (21.10), then:

$$X_n \rightarrow^d W.$$

The fractional FCLT is defined as follows, if  $X_n$  is given by:

$$X_n(\kappa) = \sigma_n^{-1} \sum_{t=1}^{[n\kappa]} (x_t), 0 < \kappa \leq 1, \quad (2.46)$$

and if we let:

$$x_t = (1 - L)^{-d} u_t, \quad (2.47)$$

where  $\sigma_n^2 = E[\sum_{t=1}^n x_t]^2$ . For the case where  $0 < d < 1/2$ , Theorem 3.1 of Davidson and De Jong (2000) can be applied to derive the fractional FCLT of  $x_t$  under Assumptions B1-B4 for the sequence  $\{u_t, -\infty < t < \infty\}$ :

B1. It has a zero mean.

B2. It is uniformly  $L_r$ -bounded for  $r > 2$ .

B3. It is  $L_2$ -NED of size  $-1/2$  on  $z_t$  with  $d_t = 1$ , where  $z_t$  is either an  $\alpha$ -mixing sequence of size  $-r/(r-2)$  or a  $\phi$ -mixing sequence of size  $-r/(2(r-1))$ .

B4. It is covariance stationary and  $0 < \sigma_u^2 < \infty$ , where:

$$\sigma_u^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^{\infty} \sum_{s=1}^{\infty} E(u_t u_s).$$

**Theorem 2.6.3** (*Theorem 3.1 in Davidson and De Jong, 2000*)  $X_n$  is defined by (2.46) and (2.47), where  $0 < d < 1/2$ , and satisfy Assumptions B1–B4, then  $X_n \rightarrow^d X$ , where  $X$  is fractional Brownian motion.

For the case where  $-1/2 < d < 0$ , Johansen and Nielsen (2012) provide a necessary moment condition of  $\{u_t\}$  in their Theorem 1, which means that Assumption B2 requires  $\{u_t\}$  to be uniformly  $L_r$ -bounded for  $r \geq 2$  and  $r \geq (d + 1/2)^{-1}$ .

### 2.6.1 The FCLT for the EGARCH( $p, q$ ) process

In this subsection, Theorem 2.6.1 is used to establish the FCLT for the EARCH( $\infty$ ) representation of EGARCH( $p, q$ ) processes. By Theorem 2.5.3, EGARCH( $p, q$ ) processes are geometrically  $L_2$ -NED on the underlying i.i.d. process  $\{z_t\}$ . Therefore, Condition A1 in Assumption A can be satisfied. For Condition A2, because  $\{h_t\}$  is covariance stationary with  $E[h_t]^2 < \infty$  in Section (2.4),  $\{h_t\}$  is uniformly

integrable, such that:

$$\begin{aligned}
& \sup_t E|h_t - E[h_t]|^2 \\
&= \sup_t E|h_t^2 - 2h_tE[h_t] + (E[h_t])^2| \\
&\leq \sup_t E[h_t^2] - (E[h_t])^2 \\
&< \infty.
\end{aligned}$$

For Condition A3, using Billingsley's (1968) Theorem 21.1, let  $\eta_t = \ln h_t - \omega$  and  $\eta_{mt} = E\{(\ln h_t - \omega) | \mathcal{F}_{t-m}^{t+m}\}$ . According to the  $L_2$ -NED properties of the  $\{G_t\}$  shown in Equation (2.59),  $\{\ln h_t - \omega\}$  is also  $L_2$ -NED and thus  $\sum_{m=1}^{\infty} (E\{|\eta_0 - \eta_{m0}|^2\})^{\frac{1}{2}} < \infty$ . Therefore,  $n^{-1}\sigma^2 = n^{-1}\text{var}(\sum_{t=1}^n (\ln h_t - \omega))$  converges. By the inequality (2.50), let  $q = 2$ , that is:

$$\begin{aligned}
& E|\sqrt{h_t} - E(\sqrt{h_t} | \mathcal{F}_{t-m}^{t+m})| \tag{2.48} \\
&\leq \left( \|\exp(\omega) \exp[G_t - \lambda(G_t - E(G_t | \mathcal{F}_{t-m}^{t+m}))]\|_2 \|G_t - E(G_t | \mathcal{F}_{t-m}^{t+m})\|_2 \right)^{\frac{1}{2}},
\end{aligned}$$

where  $0 < \lambda < 1$ , and  $\|\exp(\omega) \exp[G_t - \lambda(G_t - E(G_t | \mathcal{F}_{t-m}^{t+m}))]\|_2 < \infty$ . Thus  $\{\sqrt{h_t}\}$  satisfies Condition A3 as well. Therefore, following Theorem 1.2 in Davidson (2002), the EGARCH( $p, q$ ) process satisfies the conditions for the FCLT, as well as the  $\{\xi_t\}$  process.

Lee (2014) considered the EGARCH model as an example of an augmented GARCH( $p, q$ ) process applying Box–Cox transformation, and also proved the FCLT for the EGARCH process. In order to show that  $|h_t|^v$  and  $|\xi_t|^v$  ( $v > 0$ ) satisfy the FCLT, Lee's (2014) paper provides similar conditions:  $\sum_{j=1}^q |\varphi_j| \leq \rho^{-1} < 1$ ,

$E \left[ \exp \left( 4v \sum_{j=1}^q |\omega/p + g(z_t)|^2 \right) \right] < \infty$  and  $E|z_0|^{2v} < \infty$ . However, Lee's paper did not cover the limit theorem of the hyperbolic memory and long memory cases of EGARCH-type models. The following of this section investigates the limit theorem of these processes.

## 2.6.2 Investigation of the FCLT and fractional FCLT for the FIEGARCH(BM) process

This subsection aims to derive the FCLT or fractional FCLT for the FIEGARCH (BM) process under different signs of the memory parameter  $d$ . The FIEGARCH (BM) process is defined in Equations (2.2) and (2.8), and the pure FIEGARCH(BM) process is given in Equation (2.10). To establish the fractional FCLT for the pure FIEGARCH(BM) with  $0 < d < 1/2$ , the Theorem 3.1 in Davidson and De Jong (2000) can be applied. Here, let us assume  $u_t = g(z_{t-1})$ , according to the properties of the  $g(z_t)$  function, which is a white noise. It is then straightforward to determine that  $g(z_{t-1})$  satisfies all the conditions in Assumption B. Therefore, it is easier to verify that the  $\{\ln h_t - \omega\}$  process satisfies the fractional FCLT. However, unlike the EGARCH models, the limit theorem of  $h_t$  cannot be directly obtained from that of  $\{\ln h_t - \omega\}$  because of the fractional difference term.

The fractional FCLT for the FIEGARCH(BM)( $p, d, q$ ) process, with  $0 < d < 1/2$ , which is defined in Equation (2.8), is provided in Theorem 2.6.4. Let us set the  $X_{n(BM)}$  as

$$X_{n(BM)}(\kappa) = \sigma_n^{-1} \sum_{t=1}^{[n\kappa]} \{\ln h_t - \omega\}, \quad 0 < \kappa < 1$$

where

$$\ln h_t - \omega = (1 - L)^{-d} \frac{\psi(L)}{\varphi(L)} g(z_{t-1}).$$

Then, by Theorem 3.1 in Davidson and De Jong (2000), the Theorem 2.6.4 is established as follows.

**Theorem 2.6.4** *Suppose that  $u_t = (\psi(L)/\varphi(L)) g(z_{t-1})$  and that  $0 < d < 1/2$  satisfy the Assumption B, and  $\sup_t \|g(z_{t-1})\|_r < \infty$  for  $r > 2$  and  $\rho > 1$ . Then,*

$X_{n(BM)} \rightarrow^d X$ , where  $X$  is a fractional Brownian motion.

However, in the FIEGARCH(BM) process, the limit theorem of the  $\{\xi_t\}$  process cannot be directly obtained by the fractional FCLT of the partial sum of  $\{\ln h_t - \omega\}$ . It would be worth investigating the FCLT or fractional FCLT for  $\{\xi_t\}$  in further studies. In addition, for the limit theorem for the case when  $d$  is negative, Theorem 3.1 of Davidson and De Jong (2000) cannot be applied, but the method from Johansen and Nielsen (2012) might be useful for deriving the limit properties of the FIEGARCH(BM) process. This would also be an interesting avenue for further research.

### 2.6.3 Investigation of the FCLT and fractional FCLT for the HYEGARCH and FIEGARCH(DL) processes

This subsection aims to establish the FCLT and fractional FCLT for the pure HY/FIEGARCH(DL)(0,  $d_2$ , 0) processes, and for the higher-order HYEGARCH and FIEGARCH(DL)( $p$ ,  $d_2$ ,  $q$ ) processes, which can be similarly derived.

First, consider the HYEGARCH(0,  $d_2$ , 0) model with  $0 < d_2 < \frac{1}{2}$  and set  $\alpha = 1 > 0$  without loss of generality. The HYEGARCH model is now:

$$\begin{aligned} \ln h_t &= \omega + (1 - (1 + ((1 - L)^{d_2} - 1)))g_1(z_t) \\ &= \omega + (1 - (1 - L)^{d_2})g_1(z_t) \\ &= \omega + \sum_{j=1}^{\infty} b_j g_1(z_{t-j}), \end{aligned}$$

where  $b_j > 0$ , and  $b_j = \frac{d_2 \Gamma(j-d_2)}{\Gamma(1-d_2)\Gamma(j+1)}$ . Let us set that  $X_{n(HY)}$  as

$$X_{n(HY)}(\kappa) = \sigma_n^{-1} \sum_{t=1}^{[n\kappa]} \{\ln h_t - \omega\}, \quad 0 < \kappa < 1$$

where

$$\ln h_t - \omega = (1 - (1 - L)^{d_2}) g_1(z_t).$$

Then, by Theorem 1.2 in Davidson (2002), the Theorem 2.6.5 is established as follows.

**Theorem 2.6.5** *In the HYEGARCH process, under the Assumption A, and suppose that  $\sup_t \|g_1(z_{t-1})\|_r < \infty$  for  $r > 2$ . Then  $X_{n(HY)} \rightarrow^d X$ , where  $X$  is Brownian motion.*

For the pure FIEGARCH(DL) process with  $-\frac{1}{2} < d_2 < 0$ , without loss of generality, let  $\alpha = -1 < 0$ , then:

$$\ln h_t - \omega = ((1 - L)^{d_2} - 1)g_1(z_t). \quad (2.49)$$

Let us set that  $X_{n(FI)}$  as

$$X_{n(FI)}(\kappa) = \sigma_n^{-1} \sum_{t=1}^{[n\kappa]} \{\ln h_t - \omega\}, \quad 0 < \kappa < 1.$$

Since the Equation (2.49) can be written as follows:

$$\begin{aligned} \ln h_t - \omega &= (1 - (1 - L)^{d_2}) g_1(z_t) \\ &= (1 - L)^{d_2} (1 - (1 - L)^{-d_2}) g_1(z_t) \end{aligned}$$

where  $-1/2 < d_2 < 0$ . Then, by Theorem 3.1 of Davidson and De Jong (2000), the fractional FCLT can be established for the FIEGARCH(DL) as follows.

**Theorem 2.6.6** *In the FIEGARCH(DL) process, let  $u_t = (1 - (1 - L)^{-d_2})g_1(z_t)$ , satisfying Assumption B, and  $\sup_t \|g_1(z_{t-1})\|_r < \infty$  for  $r > 2$ . Then*

$$X_{n(FI)} \rightarrow^d X,$$

where  $X$  is fractional Brownian motion.



Theorems 2.6.5 and 2.6.6 show that  $\{\ln h_t - \omega\}$  obeys the FCLT and fractional FCLT in the HYGARCH and FIEGARCH(DL) processes, respectively. Similar to the FIEGARCH(BM) process, the limit theorem of the  $\{\xi_t\}$  process cannot be obtained by the limit property of the partial sum of  $\{\ln h_t - \omega\}$ . Theorem 2.5.2 shows that  $\{\xi_t\}$  is  $L_2$ -NED of size  $\frac{1}{2}(1/q - 1 - a_0)$  when  $1 < q \leq 2$ , and of size  $\frac{1}{2}(-a_0 - \frac{1}{2})$  when  $q > 2$  where  $0 < a_0 < a < 1/2$ . However, in the HYGARCH process, the NED size of  $\{\xi_t\}$  cannot be  $-1/2$ , and therefore, the Theorem 1.2 of Davidson (2002) cannot be applied here, this is also the case for the FIEGARCH(DL). Some other methods may be more appropriate, which may be researched in future studies.

## 2.7 Simulation for the memory properties of HYGARCH and FIEGARCH(DL) processes

In this section, Monte Carlo experiments are used to simulate the memory properties for pure HYGARCH and FIEGARCH processes.

### 2.7.1 Data generating process

The processes generated using the HY/ FIEGARCH(DL)(0,  $d$ , 0)<sup>5</sup> models are defined as:

$$\ln h_t = \omega + \alpha(1 - (1 - L)^d)g_1(z_t),$$

---

<sup>5</sup>For the simple notation in these simulations,  $d$  in this section denotes the  $d_2$  in the HY/FIEGARCH(DL) processes, which is defined in Section 2.5.2.

where:

$$g_1(z_t) = [|z_t| - \sqrt{2/\pi}].$$

Here, the memory parameter  $d$  can either be positive or negative, and the parameter  $\alpha$  has the same sign as  $d$ . In this experiment, for the FIEGARCH(DL) process, fix  $\alpha = -1$  when  $d < 0$ , and set  $d = -0.1, -0.2, \dots, -0.5$ ; for the HYEGARCH process, fix  $\alpha = 1$  when  $d > 0$ , such that  $d = 0.1, 0.2, \dots, 0.5$ . The intercept  $\omega$  is set as zero. Similar simulation results have also been obtained when setting  $\omega$  as 1 but these are not presented in this thesis.

## 2.7.2 Estimation procedure and simulation results

The GPH estimation is used to estimate the memory parameters of the  $|\xi_t|$  generated by the HYEGARCH and FIEGARCH(DL) processes.

Tables 2.1<sup>6</sup> and Figure 2.1 show the results for 10,000 observations. The reported values are the averages of N=1000 Monte Carlo replications of the generation process. It is evident that these processes show that the GPH estimators for  $|\xi_t|$  are very close to zero and even the GPH estimator of  $d = 0.2$  is slightly larger than that of  $d = 0.1$  in the HYEGARCH model, whereas these estimators are larger than 0 and less than 1/2 in the FIEGARCH(DL) process. This shows that although the HYEGARCH process has a hyperbolic memory property, it is not long memory, whereas the FIEGARCH(DL) model can capture the long memory property when it exists.

---

<sup>6</sup>These simulation results are obtained by using TSM.

## 2.8 Conclusion

This chapter has investigated the moment and memory properties of EGARCH-type models. The moment and memory properties of the EARCH( $\infty$ ) process depend on the distributions of the underlying process ( $z_t$ ). A finite-order moment of the conditional variance exists when  $z_t$  is normally distributed and when  $z_t$  has a non-normal distribution under appropriate conditions. The NED concept has been used to measure the memory properties of EARCH-class models, and the results show that general EARCH( $\infty$ ) processes can have hyperbolic memory, long memory or geometric memory. The EGARCH(1,1) model has geometric memory. The FIEGARCH(BM) model has hyperbolic memory when the memory parameter  $d < 0$  and has a long memory with a positive memory parameter  $d$ . A similar property can be shown in the HYGARCH process (with hyperbolic memory) and FIEGARCH(DL) embodies the long memory property.

The FCLT and fractional FCLT have been established for EGARCH-type models under NED on an independent process. The FCLT and fractional FCLT are important in studying the asymptotic property of the estimator in parametric time series models. This chapter constructed the FCLT for the partial sum of the return process following EGARCH models, and showed that the partial sums of the  $\{\ln h_t - \omega\}$  processes in both the FIEGARCH(BM) model with positive  $d$  and the FIEGARCH(DL) model hold the fractional FCLT, by applying the NED approach. However, the limit theorem of the return process  $\{\xi_t\}$  and the square of the volatility  $\{h_t\}$  merit further research. Moreover, it would also be worth investigating how to apply the FCLT or fractional FCLT for the score function of the QMLE in EGARCH-type models.

True $d_{HYE}$	$T^{0.5}$		True $d_{FIE}$	$T^{0.5}$	
	$\widehat{d}_{GPH}(\text{abs})$			$\widehat{d}_{GPH}(\text{abs})$	
	Mean	SD		Mean	SD
0.1	0.023	0.070	-0.1	0.052	0.067
0.2	0.025	0.069	-0.2	0.122	0.068
0.3	0.021	0.071	-0.3	0.214	0.072
0.4	0.016	0.068	-0.4	0.314	0.071
0.5	0.012	0.068	-0.5	0.417	0.075

Table 2.1: Simulation results for the HYGARCH and FIEGARCH(DL) processes. Note: This table provides the simulation results for the HYGARCH(0,  $d$ , 0) and FIEGARCH(DL)(0,  $d$ , 0) processes. These simulations were run with 10,000 observations. For the DGP of the HYGARCH processes, set  $\alpha = 1$  when  $d > 0$ , such that  $d = 0.1, 0.2, 0.3, 0.4, 0.5$ . That of FIEGARCH(DL) processes, set  $\alpha = -1$ , when  $d < 0$ , and set  $d = -0.1, -0.2, -0.3, -0.4, -0.5$ . The intercepts in both models were set as zero. The GPH estimators for the memory parameters are reported as the averages of 1000 Monte Carlo replications. It can be seen from this table that the estimated memory parameters in HYGARCH processes are very close to zero, and the persistence increased with the decreasing of  $d$  in FIEGARCH(DL) processes.

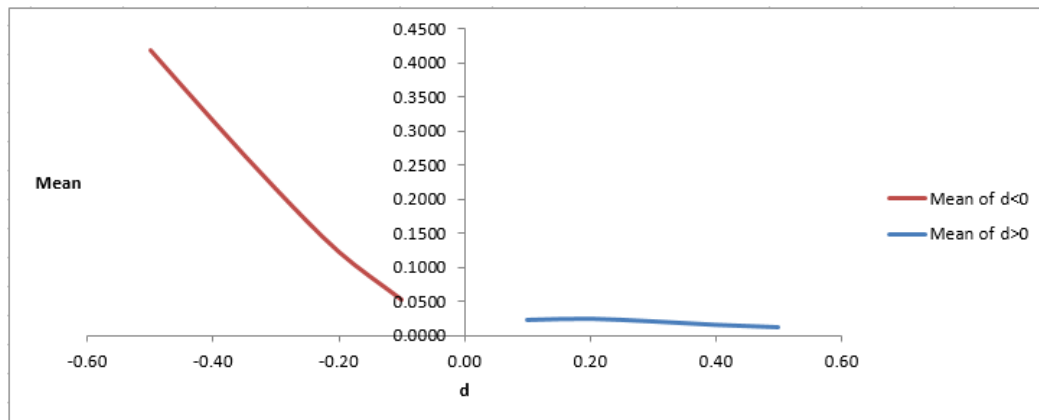


Figure 2.1: Simulation plots of the pure HYGARCH and FIEGARCH(DL) models. Note: This figure provides a corresponding plot of the simulation results presented in Table 2.1. The estimated memory parameters in the HYGARCH and FIEGARCH(DL) models are indicated by a blue line and a red line, respectively. It is clear to see that the estimated memory parameters in the HYGARCH model are very close to zero but, in the FIEGARCH(DL) model, the estimated memory parameters increase consistently with the absolute value of  $d$ .

## 2.9 Appendix A

### 2.9.1 Proof of Theorem 2.5.2

**Proof of Theorem 2.5.2.** The expression of  $L_2$ -NED is:

$$\begin{aligned}
& \left\| h_t^k - E \left[ h_t^k | \mathcal{F}_{t-m}^{t+m} \right] \right\|_2 & (2.50) \\
&= \left\| \exp(k(\omega + G_t)) - E[\exp(k(\omega + G_t)) | \mathcal{F}_{t-m}^{t+m}] \right\|_2 \\
&\leq \left\| \exp(k(\omega + G_t)) - \exp(E[k(\omega + G_t) | \mathcal{F}_{t-m}^{t+m}]) \right\|_2 \\
&= \left( E \left| \exp[k(\omega + G_t)] - \exp(E[k(\omega + G_t) | \mathcal{F}_{t-m}^{t+m}]) \right|^2 \right)^{\frac{1}{2}} \\
&\leq \left( E \left| \exp[2k(\omega + G_t)] - \exp(2kE[(\omega + G_t) | \mathcal{F}_{t-m}^{t+m}]) \right| \right)^{\frac{1}{2}} \\
&\leq \left( E \left| 2k \exp(2k\omega) \exp[2k(1 - \lambda)G_t + 2k\lambda E[G_t | \mathcal{F}_{t-m}^{t+m}]] \right| \left| G_t - E[G_t | \mathcal{F}_{t-m}^{t+m}] \right| \right)^{\frac{1}{2}} & (2.51) \\
&\leq \sqrt{2k} \left( \left\| \left[ \exp(\omega) \exp[G_t - \lambda(G_t - E[G_t | \mathcal{F}_{t-m}^{t+m}])] \right]^{2k} \right\|_{q/(q-1)} \left\| G_t - E[G_t | \mathcal{F}_{t-m}^{t+m}] \right\|_q \right)^{\frac{1}{2}},
\end{aligned}$$

where  $0 < \lambda < 1$ ,  $q > 1$ . The inequalities are obtained by the Theorem 10.12<sup>7</sup> from Davidson (1994),  $|a - b|^2 \leq |a^2 - b^2|$ , where  $a > 0$ ,  $b > 0$ , the mean value theorem, and Hölder inequality respectively. Next, set:

$$R_t = \left[ \exp(\omega) \exp[G_t - \lambda(G_t - E[G_t | \mathcal{F}_{t-m}^{t+m}])] \right]^{2k}, \quad (2.52)$$

and:

$$P_t = \left| G_t - E[G_t | \mathcal{F}_{t-m}^{t+m}] \right|, \quad (2.53)$$

then:

$$\left\| h_t^k - E \left[ h_t^k | \mathcal{F}_{t-m}^{t+m} \right] \right\|_2 \leq \sqrt{2k} \left( \|R_t\|_{q/(q-1)} \|P_t\|_q \right)^{\frac{1}{2}}. \quad (2.54)$$

---

<sup>7</sup>Let  $\hat{Y}$  denote any  $\mathcal{F}$ -measurable approximation to  $Y$ . Then

$$\|Y - E(Y|\mathcal{F})\|_2 \leq \|Y - \hat{Y}\|_2.$$

To prove that  $\{h_t\}$  is NED, it requires to bound the right-hand side of inequality (2.54). Firstly, to bound  $\|P_t\|_q$ , since  $\{G_t\}$  is a linear process, by Theorem 1 in Davidson (2004) and Davidson (1994, Chapter 17), it is easy to show the NED properties of the  $\{G_t\}$  process. Accordingly, the  $\{G_t\}$  process can be written as:

$$G_t = \sum_{j=1}^m \beta_j g(z_{t-j}) + \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}), \quad (2.55)$$

and:

$$\begin{aligned} E [G_t | \mathcal{F}_{t-m}^{t+m}] &= E \left[ \left( \sum_{j=1}^m \beta_j g(z_{t-j}) + \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right) \middle| \mathcal{F}_{t-m}^{t+m} \right] \\ &= \sum_{j=1}^m \beta_j g(z_{t-j}). \end{aligned} \quad (2.56)$$

The second equality is obtained because the process  $\{g(z_t)\}$  is the linear combination of  $z_t$  and  $|z_t|$  and with mean zero, since the process  $\{z_t\}$  is an i.i.d. with mean zero and variance one. Substituting (2.55) and (2.56) into  $P_t$  yields:

$$\|P_t\|_q = \left\| \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right\|_q. \quad (2.57)$$

Thus, for the term  $\|R_t\|_{q/(q-1)}$ :

$$\begin{aligned} \|R_t\|_{q/(q-1)} &= \left\| \left( \exp(\omega) \exp \left[ G_t - \lambda \left( G_t - E [G_t | \mathcal{F}_{t-m}^{t+m}] \right) \right] \right)^{2k} \right\|_{q/(q-1)} \\ &= \left\| \left[ \exp \left( \omega + \sum_{j=1}^{\infty} \beta_j g(z_{t-j}) - \lambda \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right) \right]^{2k} \right\|_{q/(q-1)} \\ &= \left\| \left[ \exp \left( \omega + \sum_{j=1}^{\infty} \beta_j g(z_{t-j}) \right) \exp \left( -\lambda \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right) \right]^{2k} \right\|_{q/(q-1)} \\ &= \left\| h_t^{2k} \left[ \exp \left( -\lambda \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right) \right]^{2k} \right\|_{q/(q-1)}. \end{aligned} \quad (2.58)$$

By Taylor expansion  $\exp(x) = 1 + O(x)$ , as  $x \rightarrow 0$ , then:

$$\exp \left[ -\lambda \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right] = 1 + O_p \left[ -\lambda \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right].$$

Therefore,

$$\begin{aligned} \|R_t\|_{q/(q-1)} &= \left\| h_t^{2k} \left[ 1 + O_p \left( -\lambda \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right) \right]^{2k} \right\|_{q/(q-1)} \\ &\leq \|h_t^{2k}\|_{q/(q-1)} + \left\| h_t^{2k} \cdot O_p \left( -\lambda \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right)^{2k} \right\|_{q/(q-1)}. \end{aligned}$$

This inequality is obtained by applying the Minkowski inequality. However, to obtain the upper bounds for  $\|P_t\|_q$  and  $\|R_t\|_{q/(q-1)}$ , further conditions on the lag coefficients  $\beta_j$  are required. If we assume that:

$$|\beta_j| \leq Cj^{-1-a}, \quad (a > 0),$$

and by Equation (2.57) and Lemma 2.5.1 (because  $g(z_{t-j})$  is i.i.d. with mean zero and finite variance and  $Y_j = \beta_j g(z_{t-j})$  satisfy the condition of martingale difference), and suppose  $\|g(z_t)\|_q < \infty$ , if  $1 < q \leq 2$ :

$$\begin{aligned} \|P_t\|_q &= \left[ E \left| \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right|^q \right]^{\frac{1}{q}} \tag{2.59} \\ &\leq \left[ 2 \sum_{j=m+1}^{\infty} E |\beta_j g(z_{t-j})|^q \right]^{\frac{1}{q}} \\ &\leq \left[ 2E |g(z_t)|^q \sum_{j=m+1}^{\infty} |\beta_j|^q \right]^{\frac{1}{q}} \\ &\leq \left[ 2E |g(z_t)|^q \sum_{j=m+1}^{\infty} |\beta_j|^q \right]^{\frac{1}{q}} \\ &= [2E |g(z_t)|^q]^{\frac{1}{q}} O(m^{1/q-1-a}) \\ &= d_t v_m, \end{aligned}$$

where  $d_t = [2E |g(z_t)|^q]^{\frac{1}{q}} < \infty$  and  $v_m = O(m^{1/q-1-a})$ . Therefore, the process  $\{G_t\}$  is  $L_q$ -NED on  $\{z_t\}$ , of size  $(1/q - 1 - a_0)$  ( $a > a_0 > 0$ ). Then, for:

$$\|R_t\|_{q/(q-1)} \leq \|h_t^{2k}\|_{q/(q-1)} + \|h_t^{2k} \cdot O_p(m^{(1/q-1-a)})^{2k}\|_{q/(q-1)}, \tag{2.60}$$

by the moment condition (2.26),  $\|h_t^{2k}\|_{q/(q-1)}$  in inequality (2.60) is bounded. Then, the second term is also bounded. Thus,  $\|R_t\|_{q/(q-1)}$  is finite. Thus, the process  $\{h_t^k\}$ , for  $k = 1/2$  or  $1$ , holds  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2}(1/q - 1 - a_0)$ , where  $(a > a_0 > 0)$ , and the  $\{\xi_t\}$  is also a  $L_2$ -NED. For the simple case, if let  $q = 2$ , under the conditions of  $\|h_t^{2k}\|_2 < \infty$  and  $\|g(z_t)\|_2 < \infty$ , it can be obtained that  $\|R_t\|_2 < \infty$ ,  $\|P_t\|_2 < \infty$ ,  $d_{2t} = \sqrt{2k} \left( \|h_t^{2k}\|_2 + \|h_t^{2k} O_p(m^{2k(-a-\frac{1}{2})})\|_2 \right)^{\frac{1}{2}} [2E|g(z_t)|^2]^{\frac{1}{4}} < \infty$ . Thus, the process  $\{h_t^k\}$ , for  $k = 1/2$  or  $1$ , holds  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2}(-a_0 - \frac{1}{2})$ , where  $(a > a_0 > 0)$ , and the  $\{\xi_t\}$  is also a  $L_2$ -NED. If  $q > 2$ :

$$\begin{aligned}
\|P_t\|_q &\leq \left[ E \left| \sum_{j=m+1}^{\infty} \beta_j g(z_{t-j}) \right|^q \right]^{\frac{1}{q}} \\
&\leq C_q \left[ \sum_{j=m+1}^{\infty} (E |\beta_j g(z_{t-j})|^q)^{2/q} \right]^{\frac{1}{2}} \\
&= C_q \left[ (E |g(z_{t-j})|^q)^{\frac{2}{q}} \sum_{j=m+1}^{\infty} (|\beta_j|^q)^{\frac{2}{q}} \right]^{\frac{1}{2}} \\
&\leq C_q \left[ (E |g(z_{t-j})|^q)^{\frac{2}{q}} \sum_{j=m+1}^{\infty} (|\beta_j|^q)^{\frac{2}{q}} \right]^{\frac{1}{2}} \\
&\leq C_q (E |g(z_{t-j})|^q)^{\frac{1}{q}} O \left( m^{-a-\frac{1}{2}} \right),
\end{aligned}$$

where  $C_q > 0$ ,  $(E |g(z_{t-j})|^q)^{\frac{1}{q}} < \infty$ , and then:

$$\|R_t\|_{q/(q-1)} \leq \|h_t^{2k}\|_{q/(q-1)} + \left\| h_t^{2k} \cdot O_p \left( m^{-a-\frac{1}{2}} \right)^{2k} \right\|_{q/(q-1)}. \quad (2.61)$$

Then similarly by the moment conditions,  $\|h_t^{2k}\|_{q/(q-1)} < \infty$ , and  $\|R_t\|_{q/(q-1)}$  is also bounded. Thus:

$$\begin{aligned}
&\|h_t^k - E [h_t^k | \mathcal{F}_{t-m}^{t+m}]\|_2 \\
&\leq C_q \left( \|h_t^{2k}\|_{q/(q-1)} + \left\| h_t^{2k} \cdot O_p \left( m^{-a-\frac{1}{2}} \right)^{2k} \right\|_{q/(q-1)} \right)^{\frac{1}{2}} (E |g(z_{t-j})|^q)^{\frac{1}{2q}} O \left( m^{\frac{1}{2}(-a-\frac{1}{2})} \right) \\
&= d_{2t} O \left( m^{\frac{1}{2}(-a-\frac{1}{2})} \right),
\end{aligned}$$



and  $d_{2t} < \infty$ , then  $\{h_t^k\}$  is  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2}(-a_0 - \frac{1}{2})$ , where  $(a > a_0 > 0)$ . And this also holds for  $\{\xi_t\}$ . If  $-1/2 < a < 0$ , it can be seen that the  $\{h_t^k\}$  and  $\{\xi_t\}$  are  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2}\left(\frac{1}{q} - 1 - a_0\right)$ , where  $\left(\frac{1}{2}\left(1 + a - \frac{1}{q}\right) > a_0 > 0\right)$  if  $1 < q \leq 2$ ; and  $L_2$ -NED on  $\{z_t\}$ , of size  $\frac{1}{2}\left(-a_0 - \frac{1}{2}\right)$ , where  $\left(\frac{1}{2}\left(a + \frac{1}{2}\right) > a_0 > 0\right)$  if  $q > 2$ . This completes the proof. ■

## 2.9.2 Proof of Theorem 2.5.3

**Proof of Theorem 2.5.3.** Following similar procedures to those in the proof of Theorem 2.5.2, assuming the lag coefficients  $|\beta_j| \leq C\rho^{-j}$  where  $\rho > 1$ , and  $\sup_{j \geq m+1} \|g(z_{t-j})\|_q < \infty$ . Then, it is able to show that:

$$\begin{aligned}
& \|P_t\|_q & (2.62) \\
& \leq C \sup_{j \geq m+1} \|g(z_{t-j})\|_q \sum_{j=m+1}^{\infty} |\beta_j| \\
& = C \sup_{j \geq m+1} \|g(z_{t-j})\|_q \sum_{j=m+1}^{\infty} |\rho^{-j}| \\
& = \sup_{j \geq m+1} \|g(z_{t-j})\|_q O(\rho^{-m}) \\
& = \acute{d}_{2t} \rho^{-m},
\end{aligned}$$

where  $\acute{d}_{2t} = \sup_{j \geq m+1} \|g(z_{t-j})\|_q < \infty$ . Therefore,  $\{G_t\}$  is geometrically  $L_q$ -NED on  $\{z_t\}$ .

$$\begin{aligned}
\|R_t\|_{q/(q-1)} & \leq \left\| \left[ h_t \left( 1 + O_p \left( -\lambda \sup_{j \geq m+1} g(z_{t-j}) \sum_{j=m+1}^{\infty} (C\rho^{-j}) \right) \right) \right]^{2k} \right\|_{q/(q-1)} & (2.63) \\
& \leq \|h_t^{2k}\|_{q/(q-1)} + \left\| h_t^{2k} \cdot O_p \left( -\lambda \sup_{j \geq m+1} g(z_{t-k}) \sum_{j=m+1}^{\infty} (C\rho^{-m}) \right) \right\|_{q/(q-1)}^{2k} \\
& = \|h_t^{2k}\|_{q/(q-1)} + \|(h_t O_p(\rho^{-m}))^{2k}\|_{q/(q-1)} \\
& < \infty.
\end{aligned}$$

Substituting  $\|R_t\|_{q/(q-1)}$  and  $\|P_t\|_q$  into the following equation:

$$\|h_t^k - E[h_t^k | \mathcal{F}_{t-m}^{t+m}]\|_2 \quad (2.64)$$

$$\leq \sqrt{2k} \left( \|R_t\|_{q/(q-1)} \|P_t\|_q \right)^{\frac{1}{2}} \quad (2.65)$$

$$\begin{aligned} &\leq \left[ \left( \|h_t^{2k}\|_{q/(q-1)} + \| (h_t O_p(\rho^{-m}))^{2k} \|_{q/(q-1)} \right) \sup_{j \geq m+1} \|g(z_{t-j})\|_q \right]^{\frac{1}{2}} O(\rho^{-\frac{m}{2}}) \\ &= \acute{d}_{2t} O(\rho^{-\frac{m}{2}}), \end{aligned}$$

where  $\acute{d}_{2t} = \left[ \left( \|h_t^{2k}\|_{q/(q-1)} + \| (h_t O_p(\rho^{-m}))^{2k} \|_{q/(q-1)} \right) \sup_{j \geq m+1} \|g(z_{t-j})\|_q \right]^{\frac{1}{2}} < \infty$ . Therefore, the processes  $\{h_t^k\}$  and  $\{\xi_t\}$  are geometrically  $L_2$ -NED. The EARCH( $\infty$ ) models can be geometric memory processes as well. This completes the proof. ■

### 2.9.3 Proof of Theorem 2.6.4

**Proof of Theorem 2.6.4.** First, let:

$$u_t = \frac{\psi(L)}{\varphi(L)} g(z_{t-1}),$$

by the moment properties of the  $g(z_{t-1})$  function, which is a white noise process, and thus  $(\ln h_t - \omega)$  has zero mean. This satisfies Assumption B1. To satisfy the condition B2, consider the following:

$$\begin{aligned} \|u_t\|_r &= \left\| \frac{\psi(L)}{\varphi(L)} g(z_{t-1}) \right\|_r \\ &= \left\| \sum_{j=0}^{\infty} \beta_j g(z_{t-1-j}) \right\|_r \\ &\leq \sup_t \|g(z_{t-1})\|_r \left\| \sum_{j=0}^{\infty} \beta_j \right\|_r \\ &\leq C \sup_t \|g(z_{t-1})\|_r \left\| \sum_{j=0}^{\infty} \rho^{-j} \right\|_r. \end{aligned}$$

Suppose that  $\sup_t \|g(z_{t-1})\|_r < \infty$  for  $r > 2$  and  $\rho > 1$ , and  $\|u_t\|_r < \infty$ . Thus  $\{\ln h_t - \omega\}$  is uniformly  $L_r$ -bounded for  $r > 2$ . For the condition B3, since:

$$\|u_t - E(u_t | \mathcal{F}_{t-m}^{t+m})\|_2 = \|G_t - E(G_t | \mathcal{F}_{t-m}^{t+m})\|_2 \leq \sup_{k \geq m+1} \|g(z_{t-k})\|_2 O(\rho^{-m})$$

as  $m \rightarrow \infty$ , where  $\rho > 1$ . Thus  $\{u_t\}$  is a geometrically  $L_2$ -NED. For the condition B4, since  $\{u_t\}$  is a geometrically  $L_2$ -NED, then because the  $\{\ln h_t - \omega\}$  is covariance stationary, it satisfies the covariance stationary condition, and  $0 < \sigma_u^2 < \infty$ , where:

$$\sigma_u^2 = \lim_{n \rightarrow \infty} n^{-1} \sum_{t=1}^{\infty} \sum_{s=1}^{\infty} E(u_t u_s).$$

Therefore, by Theorem 3.1 of Davidson and De Jong (2000),  $\{\ln h_t - \omega\}$  holds the fractional FCLT. This completes the proof. ■

## 2.9.4 Proof of Theorem 2.6.5

**Proof of Theorem 2.6.5.** According to Theorem 1.2 in Davidson (2002):

$$\begin{aligned} & \left\| (1 - (1 - L)^{d_2}) g_1(z_t) - E \left[ ((1 - (1 - L)^{d_2}) g_1(z_t)) | \mathcal{F}_{t-m}^{t+m} \right] \right\|_2 \\ &= \left\| \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) - E \left[ \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) \middle| \mathcal{F}_{t-m}^{t+m} \right] \right\|_2 \\ &= \left\| \sum_{j=m+1}^{\infty} b_j g_1(z_{t-j}) \right\|_2 \\ &\leq \left( \sum_{j=m+1}^{\infty} E (b_j g_1(z_{t-j}))^2 \right)^{\frac{1}{2}} \\ &\leq \left( \sup_{j \geq m+1} E |g_1(z_{t-j})|^2 \sum_{j=m+1}^{\infty} b_j^2 \right)^{\frac{1}{2}} \\ &\leq C \sup_{j \geq m+1} \|g_1(z_{t-k})\|_2 O(m^{-d_2 - \frac{1}{2}}). \end{aligned}$$

Thus the  $\{\ln h_t - \omega\}$  process in the HYGARCH(0,  $d_2$ , 0) is the hyperbolic  $L_2$ -NED of size  $-1/2$ , because  $0 < d_2 < 1/2$ , on the underlying i.i.d. process  $\{z_t\}$ . Thus

Assumption A1 holds. For the Assumption A2, since:

$$\begin{aligned}
& \sup_t E |\ln h_t - \omega - E(\ln h_t - \omega)|^r \\
&= \sup_t E \left| \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) - E \left( \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) \right) \right|^r \\
&= \sup_t E \left| \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) \right|^r \\
&= \sup_t E |g_1(z_t)|^r \left| \sum_{j=1}^{\infty} b_j \right|^r \\
&= \sup_t E |g_1(z_t)|^r \left| \sum_{j=1}^m b_j + \sum_{j=m+1}^{\infty} b_j \right|^r \\
&\leq \sup_t E |g_1(z_t)|^r \left( \left| \sum_{j=1}^m b_j \right|^r + O(m^{-d_2})^r \right).
\end{aligned}$$

Thus if  $\sup_t E |g(z_t)|^r < \infty$  ( $r > 2$ ), then Assumption A2 holds. For Assumption A3, by the Theorem 21.1 of Billingsley (1968),  $n^{-1}\sigma_n^2 = n^{-1}\text{var}(\sum_{t=1}^n (\ln h_t - \omega))$  converges. Therefore,  $\{\ln h_t - \omega\}$  process holds the FCLT. This completes the proof. ■

### 2.9.5 Proof of Theorem 2.6.6

**Proof of Theorem 2.6.6.** This theorem is proved by using Theorem 3.1 of Davidson and De Jong (2000). First, consider that the FIEGARCH(DL) process can be rewritten as:

$$\begin{aligned}
\ln h_t - \omega &= ((1 - L)^{d_2} - 1)g_1(z_t) \\
&= (1 - L)^{d_2}(1 - (1 - L)^{-d_2})g_1(z_t),
\end{aligned}$$

where  $-1/2 < d_2 < 0$ . Let  $u_t = (1 - (1 - L)^{-d_2})g_1(z_t)$ , since the  $g_1(z_t)$  function has mean zero, Assumption B1 holds. For Assumption B2, since:

$$\begin{aligned}
\|u_t\|_r &= \|(1 - (1 - L)^{-d_2})g_1(z_t)\|_r \\
&= \left\| \sum_{j=1}^{\infty} b_j g_1(z_t) \right\|_r \\
&\leq C \left\| \sum_{j=1}^{\infty} j^{d_2-1} g_1(z_{t-j}) \right\|_r \\
&\leq C \sup_t \|g_1(z_{t-1})\|_r \left\| \sum_{j=1}^{\infty} j^{d_2-1} \right\|_r \\
&\leq C \sup_t \|g_1(z_{t-1})\|_r \left( \left\| \sum_{j=1}^m j^{d_2-1} \right\|_r + \left\| \sum_{j=m+1}^{\infty} j^{d_2-1} \right\|_r \right) \\
&= C \sup_t \|g_1(z_{t-1})\|_r \left( \sum_{j=1}^m j^{d_2-1} \right) + O(m^{d_2}).
\end{aligned}$$

Suppose  $\sup_t \|g_1(z_{t-1})\|_r < \infty$ , then  $\|u_t\|_r < \infty$ . For Assumption B3:

$$\begin{aligned}
&\|(1 - (1 - L)^{-d_2})g_1(z_t) - E[(1 - (1 - L)^{-d_2})g_1(z_t) | \mathcal{F}_{t-m}^{t+m}]\|_2 \\
&= 2 \left\| \sum_{j=1}^{\infty} j^{d_2-1} g_1(z_{t-j}) + E \left[ \sum_{j=1}^{\infty} j^{d_2-1} g_1(z_{t-j}) \middle| \mathcal{F}_{t-m}^{t+m} \right] \right\|_2 \\
&= \left\| \sum_{j=m+1}^{\infty} j^{d_2-1} g_1(z_{t-j}) \right\|_2 \\
&\leq \sup_{j \geq m+1} \|g_1(z_{t-j})\|_2 \left\| \sum_{j=m+1}^{\infty} j^{d_2-1} \right\|_2 \\
&= \sup_{j \geq m+1} \|g_1(z_{t-j})\|_2 \left| \sum_{j=m+1}^{\infty} j^{d_2-1} \right| \\
&= O(m^{d_2}).
\end{aligned}$$

Because  $-\frac{1}{2} < d_2 < 0$ , Assumption B3 holds. For Assumption B4:

$$\begin{aligned}
\sigma_u^2 &= \lim_{n \rightarrow \infty} n^{-1} \sum_{t=1}^n \sum_{s=1}^n E(u_t u_s) \\
&= \lim_{n \rightarrow \infty} n^{-1} \sum_{t=1}^n \sum_{s=1}^n E \left( \sum_{j=1}^{\infty} j^{d_2-1} g_1(z_{t-j}) \sum_{j=1}^{\infty} j^{d_2-1} g_1(z_{s-j}) \right) \\
&< \infty.
\end{aligned}$$

Thus, the  $\{\ln h_t - \omega\}$  process in FIEGARCH(DL) satisfies Assumption B. This completes the proof. ■

CHAPTER 3  
ASYMPTOTIC THEORY OF THE QMLE IN ARMA MODELS  
WITH EGARCH AND HY/FIEGARCH ERRORS

### 3.1 Introduction

The QMLE is one of the most common methods used for estimating the parameters of conditional heteroscedasticity models. Engle (1982) mentioned that this estimation method can be applied when he introduced the ARCH model. However, the asymptotic properties of the estimators were not constructed in Engle's study. Weiss (1986) first established the CAN of the maximum likelihood estimator (MLE) in the ARCH model. Subsequently, a large number of studies focused on establishing the asymptotic theory for the ARCH and GARCH models, such as Lee and Hansen (1994) and Lumsdaine (1996). In fact, EGARCH-type models are more reliable in empirical applications because this class of model can overcome some of the shortcomings of GARCH models, especially the FIEGARCH<sup>1</sup> process, which can capture both the long memory and the leverage effect of volatility. Research on the asymptotic properties of the EGARCH models (Nelson, 1991) has attracted considerable attention recently (e.g. Straumann, 2005; Straumann and Mikosch, 2006; Wintenberger, 2013; Martinet and McAleer, 2015; Kyriakopoulou, 2015)<sup>2</sup>. However, the asymptotic properties of the QMLE in the EGARCH-type models still have not been fully explored, especially for the HYGARCH and FIEGARCH models.

---

<sup>1</sup>In this chapter, the FIEGARCH model means the FIEGARCH(DL) model, which is interested in Chapter 2.

<sup>2</sup>An earlier version of Kyriakopoulou (2015) is Demos and Kyriakopoulou (2013) which was first drafted in 2011.

Meanwhile, autoregressive moving-average (ARMA) models with error sequences driven by GARCH-type models (usually referred to as ARMA-GARCH models), or by EGARCH-type models (usually referred as ARMA-EGARCH models) are becoming more popular in practice. In other words, it is more reliable to assume that conditional heteroscedasticity processes as being errors of the ARMA process, and use these to model return series in financial applications, because it may be too restrictive to assume that an observed process is a pure conditional heteroscedastic model. The asymptotic theory of the QMLE in the ARMA-GARCH models was established by Francq and Zakoïan (2004), but that of the QMLE in the ARMA model with EGARCH-type errors has not been investigated.

Motivated by the advantages of EGARCH-type models and the mystery surrounding the asymptotic properties of these models, the main purposes of this chapter are: (1) to study the invertibility of EGARCH-type models, (2) to establish the asymptotic theory in the ARMA(1,1)-EGARCH(1,1) model by extending previous asymptotic results for the ARMA process with a conditional heteroscedasticity error, and (3) to investigate the CAN of the QMLE in the HYEGARCH and FIEGARCH processes. The asymptotic properties of ARMA models with HY/FIEGARCH errors are also of interest. By Section 3.4 and 3.5, however, it is clear that, if we can establish the asymptotic properties of the HY/FIEGARCH models, it will be easier to establish these of the ARMA with HY/FIEGARCH errors following a similar procedure to that shown in Section 3.5. Therefore, for the simplification and easy understanding of the theory, the third purpose of this chapter is mainly to consider the asymptotic properties of the pure HY/FIEGARCH processes.

The rest of this chapter is organised as follows. Section 3.2 briefly reviews



the relevant asymptotic properties for the QMLE in conditional heteroscedasticity models. Section 3.3 presents the relevant model expressions and the quasi-maximum-likelihood function, and Section 3.4 reviews the existing results for the invertibility property of the EGARCH process and extends these results to that of the ARMA-EGARCH and HYGARCH processes. Section 3.5 derives the conditions for the CAN of the QMLE in the ARMA-EGARCH model. Section 3.6 discusses the CAN of the QMLE in the HYGARCH and FIEGARCH processes, and provides Monte Carlo simulations to demonstrate the properties of the QMLE in the HY/FIEGARCH models. The final section is the conclusion and provides the contributions of this study and information regarding further research. All proofs are presented in Appendix B.

The following notation will be used in this chapter.  $|\cdot|$  denotes the absolute value of a univariate variable.  $\|\cdot\|$  denotes the Euclidean norm of a matrix or vector.  $O(1)$  (or  $o(1)$ ) denotes a series of non-stochastic variables that are (at most) of the order of magnitude of 1 (or of a smaller order of magnitude than 1);  $O_p(1)$  (or  $o_p(1)$ ) denotes a series of random variables that are bounded (or that converge to zero) in probability.

## 3.2 Literature review

This section briefly reviews the literature on the stationarity, ergodicity and invertibility properties of volatility models; the asymptotic theory of the estimators in the ARCH family models and the ARMA process with conditional heteroscedasticity errors; and the asymptotic properties of the estimators in the EGARCH-type models.

### 3.2.1 Stationarity and ergodicity properties of the relevant volatility models

This subsection reviews some of the main studies on the stationarity and ergodicity properties of GARCH-class models. Nelson (1990) first provided the conditions for the strict stationarity of the GARCH(1,1) process. Bougerol and Picard (1992) extended Nelson's study to investigate the strict stationarity of GARCH( $p, q$ ) processes. Ling and McAleer (2003) derived the conditions for the existence of strict stationarity, ergodicity and higher-order moments in the vector ARMA-GARCH models.

For EGARCH-type models, Nelson (1991) introduced the EGARCH model and derived the conditions for the existence of strict stationarity in EARCH( $\infty$ ) processes. However, the literature on the theoretical properties of EGARCH-type models remains limited. He et al. (2002) investigated the moment conditions for first-order EGARCH models without assuming errors with a normal distribution. Karanasos and Kim (2003) discussed the moment structure of general ARMA-EGARCH models. Straumann (2005), and Straumann and Mikosch (2006) provided conditions for the stationarity and ergodicity properties of the conditional heteroscedasticity models using a stochastic recurrence equation (SRE) approach. They also considered the general augmented GARCH (AGARCH) and EGARCH(1,1) models as examples. Kyriakopoulou (2015) summarised the conditions for the existence of stationarity and ergodicity from previous literature: Straumann and Mikosch (2006) and Aue et al. (2006). More literature on the stationarity of the conditional heteroscedasticity type volatility models is also reviewed in Chapters 2 and 4.

### 3.2.2 Asymptotic theory of the QMLE in ARCH/GARCH models

The asymptotic theory of the QMLE in ARCH/GARCH models has been studied extensively. Weiss (1986) first established the asymptotic theory for ARCH models and derived the CAN of the MLE in these model. However, this theory does not cover the GARCH and IGARCH processes (see e.g. Lee and Hansen, 1994). Subsequently, Lee and Hansen (1994), and Lumsdaine (1996) relaxed the conditions in Weiss (1986) and extended the investigation to a wider class of conditional heteroscedasticity models. Lee and Hansen (1994) introduced the asymptotic theory for the GARCH (1, 1) and IGARCH processes. The existence of the local consistency of the QMLE has been shown for both models given the existence of the conditional  $2 + \delta$  moment of the innovations. They also proved the global consistency of the QMLE and asymptotic normality for both models under the condition of a uniformly bounded conditional fourth moment of the innovations. Lumsdaine (1996) considered the asymptotic property for the QMLE in GARCH-class models that also included the IGARCH case, and derived the CAN of the QMLE of the parameters under the condition of the  $32^{nd}$ -order moment of the innovation. A stronger moment condition than Lee and Hansen's (1994) is required.

With regard to higher-order GARCH models, Berkes et al. (2003) established the CAN for the QMLE in GARCH  $(p, q)$  models, and Francq and Zakoïan (2004) derived the asymptotic properties of the QMLE for the pure GARCH model under weaker conditions. In Francq and Zakoïan's (2004) work, only a strict stationarity condition is required for the consistency of QMLE in GARCH  $(p, q)$  processes. Straumann (2005), and Straumann and Mikosch (2006) also investigated the asymptotic properties of the QMLE in conditionally heteroscedastic models and es-

tablished them for AGARCH  $(p, q)$  models by applying the SRE approach. They relaxed some of Berkes et al.'s (2003) conditions and derived the consistency of the QMLE under stationary conditions, and thus the asymptotic normality only requires the existence of the fourth moment of the underlying processes.

### **3.2.3 Asymptotic theory of the QMLE in ARMA-type models with GARCH-type errors**

Several studies have focused on the asymptotic theory of ARMA-type processes with GARCH-type errors. Ling and Li (1997) investigated the theoretical properties of the local MLE and QMLE in fractionally integrated ARMA models driven by the GARCH innovations (ARFIMA-GARCH). Ling and McAleer (2003) established the CAN of the QMLE in the vector ARMA-GARCH models under weaker conditions, such that the consistency of the QMLE only required a second-order moment condition, and the second-order moment of unconditional errors and the finite fourth-order moment of the conditional errors supported the requirements for the asymptotic normality of the QMLE. Francq and Zakoïan (2004) established the strong consistency and asymptotic normality of the QMLE of ARMA-GARCH models under weaker conditions. Francq and Zakoïan (2009) also summarised the asymptotic theory of the QMLE in GARCH-type models. Halunga and Orme (2009) investigated the asymptotic properties of the regression model with GARCH errors and misspecification tests for GARCH  $(p, q)$  models, and introduced a new test for asymmetry and nonlinearity properties.

### 3.2.4 Asymptotic theory in the EGARCH-type models

Many researchers have commented that it is difficult to establish the asymptotic theory of the QMLE in EGARCH-type models, especially the asymptotic normality of the estimators (see Straumann, 2005; Wintenberger, 2013). Some researchers have focused on the asymptotic theory for the modified EGARCH model or/and other estimation methods. Dahl and Iglesias (2008) established the asymptotic properties of the modified EGARCH process by changing the specifications of the conditional variance processes. Zaffaroni (2009) established the strong consistency and asymptotic normality of the Whittle estimator in exponential class volatility models. He considered one-shock volatility models, such as the EGARCH and GJR-GARCH models, and two-shock models, including the stochastic volatility (SV) model. He also focused on short and long memory conditional variance models to determine how long shocks to volatility will persist. His study extended the Whittle estimation to estimate one-shock exponential models and estimate long memory models, such as the FIEGARCH(BM) model and long memory stochastic volatility models.

However, the CAN of the QMLE for the EGARCH-type models still has not yet been fully developed. One of the key issues for establishing the asymptotic theory of the EGARCH-type models is invertibility. This was earlier investigated by Straumann (2005), Straumann and Mikosch (2006), and Wintenberger (2013). The most recent study of Martinet and McAleer (2015) also investigated the invertibility property of the EARCH( $\infty$ ) and EGARCH( $p, q$ ) processes, and provided a simple and explicit way to derive the invertibility conditions. More details on the invertibility are included in Section 3.4.

Some of the asymptotic properties of the QMLE in EGARCH models have

been partly established in previous literature. Straumann and Mikosch (2006) established the consistency of the QMLE in EGARCH models, and Straumann (2005) established the asymptotic normality of the MLE in the EGARCH (1,0) model under strict conditions. Demos and Kyriakopoulou (2010) investigated the asymptotic properties of the MLE and QMLE in EGARCH(1,1) parameters and demonstrated the theoretical results by using simulations. They also considered the bias between the MLE and QMLE and corrected the bias for all estimators.

However, Wintenberger (2013) critically reviewed the invertibility conditions of Straumann and Mikosch (2006) and argued that the invertibility notion cannot satisfy the conditions for the asymptotic theory of the EGARCH(1,1) models in some cases. He proposed the notion of continuous invertibility (CI) for investigating the invertibility of the EGARCH model and provided some sufficient conditions for the strong consistency of some of the QMLE in the general model when the maximisation procedure is conducted on a continuously invertible domain. He also proved the strong consistency of the volatility forecast for the EGARCH(1,1) model based on a continuously invertible domain. The approach of Wintenberger (2013) provided a weaker restriction than the approach of Straumann and Mikosch (2006), which applied the uniform Lipschitz coefficients.

For the asymptotic properties of the QMLE in EGARCH(1,1) models, Wintenberger (2013) proposed the SQMLE to some compact sets satisfying the empirical version of the condition, since the invertibility conditions are not observable. The SQMLE provided reliable volatility forecasts in which the initial value could be asymptotically ignored. To establish the asymptotic normality of the SQMLE in the EGARCH(1,1) processes, he extended theorem 5.7.9 of Straumann (2005) and showed the asymptotic normality of the QMLE because the SQMLE and

QMLE are asymptotically almost surely equivalent. However, unlike Straumann (2005), he assumed that the volatility process constitutes a geometrically ergodic Markov chain rather than assuming uniform moments on the compact set for the likelihood function and its derivatives. He also provided an additional moment condition (MM) which provides a sufficient condition for the existence of the asymptotic covariance matrix, and proved the invertibility of the first and second derivatives of the conditional variances by applying the methodologies of Theorem 3.1 of Bougerol (1993), and Straumann and Mikosch (2006). These ensure the existence of the asymptotic normality of the (S)QMLE in the EGARCH(1,1) process.

Kyriakopoulou (2015) focused on establishing the asymptotic normality of the QMLE in the EGARCH(1,1) process by applying some of the results from the studies of Straumann (2005), Demos and Kyriakopoulou (2010) and Wintenberger (2013), such as the existence of the stationarity and ergodicity of the first and second derivatives of log-variance processes and the (continuous) invertibility conditions. To establish the asymptotic normality of the QMLEs in the EGARCH(1,1) model, Kyriakopoulou (2015) first provided the conditions for the existence of the stationarity and ergodicity of the second order derivatives of the conditional variance to ensure that the Taylor expansion could be applied to the first order derivative of the log-likelihood function. Kyriakopoulou (2015) then established the CLT for the score functions by using Theorem 18.3 of Billingsley (1999) and proved, in her Lemma 1, the asymptotic normal distribution for the standardised first order derivative of the log-likelihood function. Following, motivated by Straumann (2005) and Straumann and Mikosch (2006), Kyriakopoulou (2015) applied the ergodic theorem for continuous-valued sequences of random functions to establish the uniform convergence of the second-order derivative of the log-likelihood func-

tion. This is more advanced than some of the previous literature, which applied the classical method, which normally requires the boundness of third order derivatives of the log-variance functions. By the ergodic theorem here, only the second order derivatives need to be bounded. To show the uniform convergence, Kyriakopoulou (2015) provided the sufficient conditions for the existence of the expectation of the supremum norm of the first order and second order derivative of the log-variance function and proved the existence of the moment estimates. This is one of the main contributions of their work. The existence of these moment conditions supports the establishment of the asymptotic normality of the QMLE in the EGARCH(1, 1) model.

### **3.2.5 Finite sample properties of the EGARCH-type models**

Research on the finite sample properties of EGARCH models has also attracted the attention of scholars. Deb (1996) performed the finite sample properties of the MLE and QMLE in the EGARCH(1, 1) model by applying the Monte Carlo method. Perez and Zaffaroni (2008) examined the finite sample properties of the MLE and Whittle estimators in EGARCH and FIEGARCH(BM) models, and compared both estimation methods using Monte Carlo simulations. Lopes and Prass (2014) considered the theoretical properties of the FIEGARCH(BM) process, and applied Monte Carlo simulations to show how to generate, estimate and forecast FIEGARCH(BM) models.



### 3.3 The models and the QMLE

This section introduces the specifications of the ARMA(1, 1)-EGARCH(1, 1) model, and the HY/FIEGARCH process, as well as the quasi-maximum likelihood functions.

#### 3.3.1 The ARMA(1,1)-EGARCH(1,1) model

Suppose the observations  $\{y_1, \dots, y_n\}$  follow the ARMA(1, 1)-EGARCH (1, 1) process, which is defined as:

$$y_t - \mu_0 = \phi_{01}(y_{t-1} - \mu_0) + \xi_{0t} + \psi_{01}\xi_{0t-1}; \quad (3.1)$$

$$\xi_{0t} = \sqrt{h_{0t}}z_{0t}; \quad (3.2)$$

$$\ln h_{0t} = w_0 + g(z_{0t-1}) + \beta_0 \ln h_{0t-1}; \quad (3.3)$$

$$g(z_{0t-1}) = \theta_0 z_{0t-1} + \gamma_0 |z_{0t-1}|, \quad (3.4)$$

where  $w_0, \theta_0, \gamma_0 \in R$ , and  $\beta_0$  can be positive or negative values. In this chapter, for the convenience of establishing the asymptotic theory of the relevant EGARCH-type models, similar to Straumann (2005), the standard EGARCH(1, 1) process is considered to be a causal AR(1) process with a mean of  $((w + \gamma E|z_0|) / (1 - \beta))$  and the error term  $(\theta z_{t-1} + \gamma (|z_{t-1}| - E|z_0|))$ . In this chapter, the  $g(z_{0t-1})$  function is defined as (3.4) and can capture the leverage effect of volatility,  $\{z_{0t}\}$  is a sequence of i.i.d. random variables with a mean of zero and a variance of one, such that  $z_{0t} \sim iid(0, 1)$ ; and  $\sqrt{h_{0t}} > 0$  is the volatility, where the square volatility process  $h_{0t} = E[\xi_t^2 | \mathcal{F}_{t-1}]$  and where the  $\mathcal{F}_{t-1}$ -measurable is a  $\sigma$ -field, and follows an EGARCH(1, 1) process; and assume that the  $y_t$  process exists and is a stationary process. The value of the true parameter vector is unknown and is denoted by

$\lambda_0 = (\varphi'_0, \varsigma'_0)'$ , where  $\varphi_0 = (\mu_0, \phi_{01}, \psi_{01})'$  in the conditional mean equation, and  $\varsigma_0 = (w_0, \beta_0, \theta_0, \gamma_0)'$  in the conditional variance equation.

The model for the generic parameter vector  $\lambda = (\varphi', \varsigma)'$ , where  $\varphi = (\mu, \phi_1, \psi_1)'$ , and  $\varsigma = (w, \beta, \theta, \gamma)'$ , is as follows:

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \tilde{\xi}_t + \psi_1\tilde{\xi}_{t-1}; \quad (3.5)$$

$$\tilde{\xi}_t = \sqrt{\tilde{h}_t}\tilde{z}_t; \quad (3.6)$$

$$\ln \tilde{h}_t = w + g(\tilde{z}_{t-1}) + \beta \ln \tilde{h}_{t-1}; \quad (3.7)$$

$$g(\tilde{z}_{t-1}) = \theta\tilde{z}_{t-1} + \gamma|\tilde{z}_{t-1}|, \quad (3.8)$$

The error term  $\tilde{\xi}_t$  and  $\ln \tilde{h}_t$  are defined as the processes which depend on finite past information, where  $\tilde{\xi}_t$  is expressed as:

$$\tilde{\xi}_t = \sum_{i=0}^{t-1} \gamma_i(y_{t-i} - \mu),$$

and  $(1 + \psi_1 L)^{-1}(1 - \phi_1 L) = \sum_{i=0}^{\infty} \gamma_i L^i$ . For the definition of  $\ln \tilde{h}_t$  refers to Section 3.4.

In order to establish the asymptotic theory of the QMLE in the ARMA-EGARCH model, the following assumptions need to be satisfied.

**Assumption 3.3.1** *The parameter space  $\Theta$  is a compact subspace of Euclidean space and  $\lambda_0$  is an interior point in  $\Theta$  for all  $\lambda \in \Theta$ .*

**Assumption 3.3.2** *In the ARMA(1,1) process,  $|\phi_1| < 1$  and  $|\psi_1| < 1$ , and  $\phi(z) = 1 - \phi_1 z$  and  $\psi_1(z) = 1 + \psi_1 z$  have no common zeros.*

**Assumption 3.3.3** *In the EGARCH(1,1) process, the lag coefficient of  $|\beta| < 1$ .*

**Assumption 3.3.4** *In the EGARCH(1,1) process, the parameters  $(\theta, \gamma) \neq (0, 0)$  to avoid the identification problem for the asymmetric function  $g(z_{t-1})$ .*

**Assumption 3.3.5** *The conditional variance  $h_t$  has a lower bound uniformly over  $\Theta$  in the EGARCH (1,1) process.*

Assumption 3.3.1 requires the true parameter vector  $\lambda_0$  to be in the interior of  $\Theta$ . The stationarity, ergodicity and invertibility properties are very crucial for deriving the asymptotic properties of the QMLE in the ARMA(1,1)-EGARCH (1,1) process. Based on previous literature, Assumptions 3.3.1 and 3.3.2 ensure the stationarity and invertibility of the ARMA(1,1) process. Assumption 3.3.2 provides an identifiability assumption for the ARMA(1,1) process (see e.g. Straumann, 2005). Assumption 3.3.3 ensures the existence of a stationary solution to the SRE of the EGARCH(1,1) process. This satisfies the stationary condition that was introduced by Straumann (2005) and Wintenberger (2013).

### 3.3.2 The HY/FIEGARCH models

The general framework of HY/FIEGARCH models was defined in the previous chapter. Suppose that the observations  $\{\xi_1, \dots, \xi_n\}$  follow the HY/FIEGARCH(0,  $d$ , 0) models<sup>3</sup>, which are defined in (3.2) and:

$$\ln h_{0t} = \omega_0 + \alpha \sum_{j=1}^{\infty} c_0 j^{-d_0-1} g_1(z_{0t-j}), \quad (3.9)$$

and:

$$g_1(z_{0t-j}) = (\theta_0 z_{0t-j} + |z_{0t-j}|),$$

---

<sup>3</sup>The memory parameter  $d$  for the HY/FIEGARCH(0,  $d$ , 0) models in this chapter is the same as  $d_2$  in the previous chapter.

where  $\alpha$  is fixed as 1 or  $-1$  in the HYGARCH and FIEGARCH models in this chapter, respectively. The true parameter vector is denoted as  $\vartheta_0 = (\omega_0, \theta_0, c_0, d_0)'$ . The parameter  $\hat{\vartheta} = (\omega, \theta, c, d)'$ , as follows:

$$\xi_t = \sqrt{\tilde{h}_t} \tilde{z}_t, \quad (3.10)$$

and:

$$\ln \tilde{h}_t = \omega + \alpha \sum_{j=1}^{\infty} c j^{-d-1} g_1 \left( \frac{\xi_{t-j}}{\sqrt{\tilde{h}_{t-j}}} \right). \quad (3.11)$$

The conditional variance process  $\{\tilde{h}_t\}$  depends on finite past information. In order to establish the CAN of the QMLE in the HY/FIEGARCH processes, the following assumptions are proposed.

**Assumption 3.3.6** *The parameter space  $\Theta$  is a compact subspace of Euclidean space, and  $\vartheta_0$  is an interior point in  $\Theta$ , for all  $\vartheta \in \Theta$ .*

**Assumption 3.3.7** *In the HYGARCH model,  $d \in [\underline{d}_h, \bar{d}_h]$ , where  $0 < \underline{d}_h < \bar{d}_h < 1/2$ .*

**Assumption 3.3.8** *In the FIEGARCH model,  $d \in [\underline{d}_f, \bar{d}_f]$ , where  $-1/2 < \underline{d}_f < \bar{d}_f < 0$ .*

**Assumption 3.3.9** *The parameter  $\omega \in [\underline{\omega}, \bar{\omega}]$ , where  $-\infty < \underline{\omega} < \bar{\omega} < +\infty$ .*

Assumption 3.3.6 is similar to Assumption 3.3.1 for the ARMA-EGARCH process in requiring the true parameter  $\vartheta_0$  to be in the interior of  $\Theta$ . Assumptions 3.3.7 and 3.3.8 define the lower and upper values for each parameter and the sign of the memory parameter  $d$  in the HYGARCH and FIEGARCH processes, respectively.

### 3.3.3 Quasi-maximum-likelihood estimator

For estimation of the ARMA(1,1)-EGARCH(1,1) model, the log-likelihood function of the QML estimator (ignoring the constant term) is expressed as:

$$\tilde{L}_n(\lambda) = \frac{1}{n} \sum_{t=1}^n \tilde{l}_t(\lambda), \quad (3.12)$$

where:

$$\tilde{l}_t(\lambda) = -\ln \tilde{h}_t(\lambda) - \frac{\tilde{\xi}_t^2(\lambda)}{\tilde{h}_t(\lambda)}, \quad (3.13)$$

where  $\tilde{L}_n(\lambda)$  denotes the Gaussian log-likelihood. The estimators in these equations are the QMLEs depending on the finite past information. It is worth noting that  $\tilde{\xi}_t$  and  $\tilde{h}_t$  are unobserved in the ARMA-EGARCH process.

A QMLE of  $\lambda_0$  is defined as:

$$\hat{\lambda}_n = \arg \max_{\lambda \in \Theta} \tilde{L}_n(\lambda), \quad (3.14)$$

which means that the QMLE  $\hat{\lambda}_n$  maximises  $\tilde{L}_n(\lambda)$  on the parameter space  $\Theta$ . The variance processes  $\{(\xi_t, \ln h_t) : t = 0, \pm 1, \pm 2, \dots\}$  depend on infinite past information in the ARMA(1,1) -EGARCH(1,1) process is:

$$y_t - \mu = c + \phi_1(y_{t-1} - \mu) + \xi_t + \psi_1 \xi_{t-1}; \quad (3.15)$$

$$\xi_t = \sqrt{h_t} z_t; \quad (3.16)$$

$$\ln h_t = w + g(z_{t-1}) + \beta \ln h_{t-1}; \quad (3.17)$$

$$g(z_{t-1}) = \theta z_{t-1} + \gamma |z_{t-1}|. \quad (3.18)$$

Equation (3.17) can be rewritten as:

$$\ln h_t = \frac{1}{1 - \beta L} (w + g(z_{t-1})). \quad (3.19)$$

When  $\lambda = \lambda_0$ , we are able to obtain  $\xi_t = \xi_{0t}$ , and  $\ln h_t = \ln h_{0t}$ . The unobserved log-likelihood function, which is conditional on infinite past observations, can be expressed as (ignoring the constant term):

$$L_n(\lambda) = \frac{1}{n} \sum_{t=1}^n l_t(\lambda); \quad (3.20)$$

$$l_t(\lambda) = -\ln h_t - \frac{\xi_t^2}{h_t}. \quad (3.21)$$

The difference between these two log-likelihood functions ( $\tilde{L}_n(\lambda)$  and  $L_n(\lambda)$ ) is that the former is conditional on any initial values and the latter is conditional on the infinite past observations. In addition, the  $\tilde{L}_n(\lambda)$  is normally used in practice because it is not possible to obtain all of the past information for  $L_n(\lambda)$ .

With regard to the HY/FIEGARCH(0,  $d$ , 0), the unobserved variance processes, given all the past information, is defined as:

$$\xi_t = \sqrt{h_t} z_t, \quad (3.22)$$

and

$$\ln h_t = \omega + \alpha(1 - (1 - L)^d) g_1 \left( \frac{\xi_t}{\sqrt{h_t}} \right). \quad (3.23)$$

The unobserved log-likelihood function (ignoring the constants), which is conditional on infinite past information, can be expressed the same way as the likelihood functions in (3.20) and (3.21), replacing  $\lambda$  with  $\vartheta$ . The log-likelihood function (ignoring the constant term) depends on finite past information, which can be expressed in the same way as the likelihood function in (3.12) and (3.13), replacing  $\lambda$  with  $\vartheta$ .

### 3.4 The invertibility of the EGARCH-type processes

This section reviews the existing results for the invertibility<sup>4</sup> of the EGARCH-type processes and extends these results to investigate that of the ARMA-EGARCH and the HYGARCH processes. Invertibility is an important issue for establishing the asymptotic theory of nonlinear processes. It was proposed by Granger and Andersen (1978) and implies that the variance of the error between the correct generation formula and the estimation in terms of a finite number of past observations converges to zero as the sample size tends to infinity. In other words, the starting up values can be asymptotically ignored. This notion of invertibility was developed by Tong (1993), Straumann (2005), and Straumann and Mikosch (2006), among others. However, it is still difficult to derive the invertibility of the EGARCH-type models (see e.g. Straumann, 2005; Straumann and Mikosch, 2006; Wintenberger, 2013; Martinet and McAleer, 2015).

#### 3.4.1 Invertibility of the EGARCH( $p, q$ ) process

Straumann (2005), and Straumann and Mikosch (2006) made a great contribution towards investigations into the invertibility of conditional heteroscedastic models.

They defined these models as:

$$\left\{ \begin{array}{l} \xi_t = \sqrt{h_t} z_t \\ h_t = f_\zeta(\xi_t, \dots, \xi_{t-p}, h_{t-1}, \dots, h_{t-q}) \end{array} \right. \quad (t \in Z), \quad (3.24)$$

where the volatility process  $\sqrt{h_t}$  is non-negative and  $\{z_t\}$  is an i.i.d. with a mean of 0 and a variance of 1. Straumann and Mikosch (2006) explained that, in the

---

<sup>4</sup>There is an alternative interpretation for the notion of invertibility in the ARMA process (see e.g. Straumann, 2005).

context of nonlinear models, the unique stationary ergodic solution  $(\xi_t, \sqrt{h_t})$  to the model (3.24) is invertible if:

$$\left| h_t - \tilde{h}_t \right| = o_p(1), \quad t \rightarrow \infty, \quad (3.25)$$

where  $h_t$  indicates that the conditional variance depends on infinite past information and  $\tilde{h}_t$  indicates that the conditional variance depends on finite past observations. This means that invertibility ensures that the difference between  $h_t$  and  $\tilde{h}_t$  converges to zero in probability (see Straumann, 2005). Consequently, the invertibility of the process plays a vital role in establishing the asymptotic theory of the QMLE in the relevant conditional heteroscedasticity processes.

Straumann (2005), and Straumann and Mikosch (2006) considered the SRE<sup>5</sup> technique to be an important approach to investigate conditionally heteroscedastic time series models. According to Theorem 2.8 of Straumann and Mikosch (2006), the EGARCH(1,1) process is transformed as a SRE, such that:

$$\ln h_{t+1} = \phi_t^\ddagger(\ln h_t), \quad t \in Z, \quad (3.27)$$

where:

$$\phi_t^\ddagger(\ln h_{t+1}) = w + \theta z_t + \gamma |z_t| + \beta \ln h_t. \quad (3.28)$$

---

<sup>5</sup>With reference to Straumann and Mikosch (2006), the general SRE can be expressed as a case of a complete separable metric space. Let us set  $(E, D)$  to be a Polish space equipped with its Borel  $\sigma$ -field  $B(E)$ . A map  $\phi^\ddagger : E \rightarrow E$  is called a Lipschitz, if:

$$\Lambda(\phi^\ddagger) := \sup_{x, y \in E, x \neq y} \frac{D(\phi^\ddagger(x), \phi^\ddagger(y))}{D(x, y)}$$

is finite and is called a contraction if  $\Lambda(\phi^\ddagger) < 1$ . The  $\Lambda$  is submultiplication (i.e., if  $\phi^\ddagger$  and  $\psi^\ddagger$  are Lipschitz maps  $E \rightarrow E$ , then  $\Lambda(\phi^\ddagger \circ \psi^\ddagger) \leq \Lambda(\phi^\ddagger)\Lambda(\psi^\ddagger)$ ). Consider a process  $(\phi_t^\ddagger)$  of random Lipschitz maps  $E \rightarrow E$  with  $\phi_t^\ddagger(x)$  being  $B(E)$ -measurable for every fixed  $x \in E$  and  $t \in Z$ . If for a stochastic process  $(X_t)_{t \in T}$  with values in  $E$ ,

$$X_{t-1} = \phi_t^\ddagger(X_t) \text{ a.s., } t \in T \quad (3.26)$$

Then, the  $(X_t)_{t \in T}$  obeys the SRE associated with  $\phi_t^\ddagger$ . It is also referred to as a solution to SRE (3.26).



Straumann and Mikosch (2006) derived sufficient conditions for the existence of a unique stationary solution for the EGARCH process:  $|\beta| < 1$  together with  $E[\ln^+(w + \theta z_0 + \gamma|z_t|)] < \infty$ , since  $Ez_t^2 = 1$ . The SRE form of the EGARCH process can also be reorganised as:

$$\ln h_t = \frac{w}{1 - \beta} + \sum_{j=1}^{\infty} \beta^{j-1} (\theta z_{t-j} + \gamma|z_{t-j}|), \quad t \in Z. \quad (3.29)$$

Straumann (2005), and Straumann and Mikosch (2006) investigated the invertibility of AGARCH and EGARCH processes. Both studies derived conditions for the existence of the invertibility for the general AGARCH models. Nevertheless, the conditions for the EGARCH model are very complex. Since  $z_t = \xi_t/\sqrt{h_t}$ , the  $\ln h_t$  is written as:

$$\begin{aligned} \ln h_t &= w + \theta \frac{\xi_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left| \frac{\xi_{t-1}}{\sqrt{h_{t-1}}} \right| + \beta \ln h_{t-1} \\ &= w + (\theta \xi_{t-1} + \gamma |\xi_{t-1}|) \exp\left(-\frac{1}{2} \ln h_{t-1}\right) + \beta \ln h_{t-1}. \end{aligned} \quad (3.30)$$

Theorem 3.1 of Bougerol (1993), which is used to prove the invertibility of the model under the condition that  $\phi_t^\dagger(\cdot, \varsigma_0)$  is a Lipschitz map, is rewritten as Theorem 2.6.1 in Straumann (2005).

**Theorem 3.4.1** (*Theorem 2.6.1 in Straumann (2005)*) *Let  $(\phi_t^\dagger)$  be a stationary ergodic sequence of Lipschitz maps from  $E$  into  $E$ . Suppose that the following conditions hold:*

$$S1: \text{ There is a } y \in E \text{ such that } E \left[ \ln^+ D(\phi_0^\dagger(y), y) \right] < \infty.$$

$$S2: E \left[ \ln^+ \Lambda(\phi_0^\dagger) \right] < \infty \text{ and for some integer } r \geq 1 :$$

$$E \left[ \ln \Lambda(\phi_0^\dagger \circ \dots \circ \phi_{-r+1}^\dagger) \right] < 0 \quad (3.31)$$

The SRE admits a stationary ergodic solution  $(Y_t)_{t \in T}$ , which has the stochastic representation:

$$Y_t = \lim_{m \rightarrow \infty} \phi_0^\dagger \circ \cdots \circ \phi_{t-m}^\dagger(y), t \in T. \quad (3.32)$$

The random elements  $Y_t$  are measurable with respect to the  $\sigma$ -field generated by  $\{\phi_{t-k}^\dagger \mid k \geq 1\}$ . If  $(\tilde{Y}_t)_{t \in T}$  is any other solution to the SRE, then:

$$D(\tilde{Y}_t, Y_t) \xrightarrow{e.a.s} 0, t \rightarrow \infty.$$

Moreover, in the case where  $T = Z$ , the stationary solution to the SRE is unique.

Proposition 5.2.9 from Straumann (2005) then provided a sufficient condition for the invertibility of the EGARCH process. This supposes that if  $\xi_t$  is a stationary ergodic EGARCH process and  $\gamma \geq |\theta|$ , then the inequality in (3.33) is sufficient for the invertibility of the conditional variance process.

$$E \left[ \ln \left( \max \left\{ \beta, \frac{1}{2}(\theta z_0 + \gamma |z_0|) \exp(f) - \beta \right\} \right) \right] < 0, \quad (3.33)$$

with:

$$f = \frac{1}{2} \sum_{j=0}^{\infty} \beta^j (\theta z_{-j-1} + \gamma |z_{-j-1}|).$$

However, Straumann (2005) also mentioned that it is difficult to obtain an explicit representation of  $h_t$  based on past observations, except when  $\beta = 0$ . If  $\beta = 0$ , then the condition (3.33) becomes:

$$-\ln 2 + \frac{\gamma}{2} E|z_0| + E[\ln(\theta z_0 + \gamma |z_0|)] < 0. \quad (3.34)$$

However, this uniform invertibility condition can only establish the asymptotic normality of the estimator for a degenerate case with  $\beta = 0$ .

Wintenberger (2013) extended the previous literature and investigated the asymptotic property of the QMLE in the EGARCH(1,1) process under the (ST) condition:

(ST): The SRE admits a unique stationary solution denoted  $h_t$  that is non-anticipative that is,  $h_t$  is independent of  $(z_t, z_{t+1}, z_{t+2}, \dots)$  for all  $t \in Z$  and has finite log moments:  $E [\ln^+ h_{0t}] < \infty$ .

Under this condition, Wintenberger (2013) provided a definition of invertibility following Tong (1993):

**Definition 3.4.2** *Under the (ST) condition, the model is invertible if the sequence of volatilities  $h_t$  is adapted to the filtration generated by  $(\xi_{t-1}, \xi_{t-2}, \dots)$ .*

He also defined the notion of continuous invertibility as:

**Definition 3.4.3** *The model,*

$$\begin{aligned} f_{t+1}(\varsigma) &= \phi_t^\dagger(f_t(\varsigma), \varsigma) \\ &= w + (\theta \xi_{t-1} + \gamma |\xi_{t-1}|) \exp\left(-\frac{1}{2} f_t(\varsigma)\right) + \beta \ln h_{t-1}, \end{aligned}$$

$t \in Z$  and  $\varsigma \in \Theta$ , is continuously invertible on  $\Theta$  if and only if  $\left\| \widehat{f}_t(\varsigma) - f_t(\varsigma) \right\|_{\Theta} \xrightarrow{a.s.} 0$  as  $t \rightarrow \infty$ , where  $\|\cdot\|_{\Theta}$  denotes the uniform norm.

Wintenberger (2013) proved the QMLE on  $\Theta$  is strongly consistent on the continuous invertible domain and provided sufficient conditions for the continuous invertibility. The sufficient condition is that if the conditions (ST) and (CI) hold on some compact set  $\Theta$ , then the model is continuously invertible on  $\Theta$ , where the (CI) condition is as follows:

(CI):  $E \left[ \ln^+ \left\| \phi_0^\dagger(x, \cdot) \right\|_{\Theta} \right] < \infty$  for some  $x \in E$ ,  $E \left[ \ln^+ \left\| \Lambda_{\phi_0^\dagger} \right\|_{\Theta} \right] < \infty$  and  $E \left[ \ln \Lambda_{\phi_0^\dagger(r)}(\varsigma) \right] < 0$  for any  $\varsigma \in \Theta$ .

Wintenberger (2013) claimed that the QMLE is not reliable for the EGARCH(1,1) model when the model is non-invertible and proved the strong consistency and asymptotic normality of the SQMLE. He also showed that the volatility forecasting of using SQMLE asymptotically ignores the initial values even when the model is non-invertible. Thus, he believed that the SQMLE is more reliable than the QMLE for volatility forecasting in some cases.

Martinet and McAleer (2015) established the invertibility conditions for the EGARCH process via a simple and easy procedure. They rewrote the EARCH( $\infty$ ) process, which was introduced by Nelson (1991), as the error  $\xi_t$  that is driven by a stochastic process such that:

$$\xi_t = z_t \cdot \exp \left( \frac{\omega}{2} + \sum_{j=1}^{\infty} \beta_j \left[ \frac{\theta}{2} z_{t-j} + \frac{\gamma}{2} |z_{t-j}| \right] \right), \quad (3.35)$$

where  $\sum_{j=1}^{\infty} |\beta_j| < \infty$ ,  $z_t \sim i.i.d.(0, 1)$  and  $z_t \in L^2$ , and  $(\theta, \gamma) \in R^2$ ,  $\omega \in R$ . In the EGARCH(1,1) process,  $\omega = w / (1 - \beta)$  and  $\beta_j = \beta^{j-1}$ . In order to derive the invertibility conditions, Martinet and McAleer (2015) defined the normalised shocks  $z_t$  in terms of the past observed information, which means that  $z_t$  is  $\sigma(\xi_t, \xi_{t-1}, \dots)$ -adapted. They emphasised that the way to define the normalised shocks of  $z_t$  that is similar to that in Straumann and Mikosch (2006), and Wintenberger (2013), who stated that  $\sqrt{h_t}$  is  $\sigma(\xi_t, \xi_{t-1}, \dots)$ -adapted. Equation (3.35) can be reorganised as:

$$\ln |z_t| = \ln |\xi_t| - \frac{\omega}{2} + \sum_{j=1}^{\infty} \beta_j \cdot g_{\theta, \gamma}(\ln |z_{t-j}|, \xi_{t-j}), \quad (3.36)$$

where:

$$g_{\theta, \gamma}(\ln |z_t|, \xi_t) = -\frac{|z_t|}{2} (\gamma + \theta \cdot \text{sign}(\xi_t)),$$

as  $\xi_t = z_t \sqrt{h_t}$  and  $\sqrt{h_t} > 0$ , and thus  $\text{sign}(z_t) = \text{sign}(\xi_t)$ . It is important to understand the variability of the  $g_{\theta, \gamma}(\ln |z_t|, \xi_t)$  function. However, since the function (3.36) is not Lipschitz, Martinet and McAleer (2015) used a similar method to obtain the bound for Lyapunov exponents or Lipschitz coefficients, and also provided

its variability in their Lemma 1.1.<sup>6</sup> They also proved that EGARCH-type models might capture the leverage effect of shocks to volatility under three cases: (i)  $\gamma \geq |\theta|$ , (ii)  $|\gamma| < |\theta|$  and  $\gamma < 0$ , (iii)  $\gamma \leq -|\theta|$ . Martinet and McAleer (2015) proposed that a recursion among the  $\ln |z_t|$  which is defined in Equation (3.36). By this equation, if we have a fixed  $t$  and an independent shock  $z_t$ , and when  $\{z_s, s \leq t - n\}$  are known for any positive integer  $n$ , then, from the observed shocks  $(\xi_t)$ , we will be able to get the exact value of  $u_t$ . Therefore, the following series  $u_k^{(n)}$  can be defined by extending the 'exact' recursion  $n$  steps backward. First, we can derive one further step by knowing  $\{z_s, s \leq t - n\}$ , where  $u_1^{(n)}$  is defined as:

$$\begin{aligned} u_1^{(n)} &= \ln |\xi_{t-n+1}| - \frac{\omega}{2} + \sum_{j=1}^{\infty} \beta_j \cdot g_{\theta, \gamma} (\ln |z_{t-n+1-j}|, \xi_{t-n+1-j}) \\ &= \ln |\xi_{t-n+1}| - \frac{\omega}{2} + \sum_{j=0}^{\infty} \beta_{j+1} \cdot g_{\theta, \gamma} (\ln |z_{t-n-j}|, \xi_{t-n-j}). \end{aligned}$$

Since all the past information of  $|z_{t-n-j}|$  for all  $j \geq 0$  is known, we are able to obtain  $u_1^{(n)}$ , and thus:

$$\begin{aligned} u_2^{(n)} &= \ln |\xi_{t-n+2}| - \frac{\omega}{2} + \sum_{j=1}^{\infty} \beta_j \cdot g_{\theta, \gamma} (\ln |z_{t-n+2-j}|, \xi_{t-n+2-j}) \\ &= \ln |\xi_{t-n+2}| - \frac{\omega}{2} + \beta_1 \cdot g_{\theta, \gamma} (u_1^{(n)}, \xi_{t-n+1}) + \sum_{j=0}^{\infty} \beta_{j+2} \cdot g_{\theta, \gamma} (\ln |z_{t-n-j}|, \xi_{t-n-j}), \end{aligned}$$

and:

$$\begin{aligned} u_3^{(n)} &= \ln |\xi_{t-n+3}| - \frac{\omega}{2} + \sum_{j=1}^{\infty} \beta_j \cdot g_{\theta, \gamma} (\ln |z_{t-n+3-j}|, \xi_{t-n+3-j}) \\ &= \ln |\xi_{t-n+3}| - \frac{\omega}{2} + \sum_{i=1}^2 \beta_i \cdot g_{\theta, \gamma} (u_{3-i}^{(n)}, \xi_{t-n+3-i}) + \sum_{j=0}^{\infty} \beta_{j+3} \cdot g_{\theta, \gamma} (\ln |z_{t-n-j}|, \xi_{t-n-j}). \end{aligned}$$

---

<sup>6</sup>Lemma 1.1 of Martinet and McAleer (2015), (1)

$$|g_{\theta, \gamma}(x_1, y) - g_{\theta, \gamma}(x_2, y)| \leq \left| \frac{\gamma + \theta \cdot \text{sign}(y)}{2} \right| \exp(\max(x_1, x_2)) |x_1 - x_2|,$$

(2)

$$|g_{\theta, \gamma}(x_1, y) - g_{\theta, \gamma}(x_2, y)| \geq \left| \frac{\gamma + \theta \cdot \text{sign}(y)}{2} \right| \exp\left(\frac{x_1 + x_2}{2}\right) |x_1 - x_2|.$$

Therefore, a more general expression of  $u_{k+1}^{(n)}$  is:

$$u_{k+1}^{(n)} = \ln |\xi_{t-n+k+1}| - \frac{\omega}{2} + \sum_{i=1}^k \beta_i \cdot g_{\theta, \gamma} \left( u_{k+1-i}^{(n)}, \xi_{t-n+k+1-i} \right) + \sum_{j=0}^{\infty} \beta_{j+1+k} \cdot g_{\theta, \gamma} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right).$$

Martinet and McAleer (2015), in their Lemma 2.1, showed that  $u_k^{(n)} = \ln |z_{t-n+k}|$  for any  $n, k \in N$ . This Lemma makes these series represent an 'exact' recursion more obviously. It can be seen that the value of  $\ln |z_t|$  depends more on the observed shocks ( $\xi_t$ ) and the past values  $\{z_s, s \leq t-n\}$  have less effect on  $z_t$  as  $n$  tends to infinity, when following the above 'exact' recursion. This means that the past information  $\{z_s, s \leq t-n\}$  can be asymptotically ignored. Then, Martinet and McAleer (2015) assumed all the past values  $\{z_s, s \leq t-n\}$  to be zero and the observed shocks ( $\xi_t$ ) were used to derive the value of  $z_t$  when  $n$  is large enough. This shows the invertibility of the model. A  $\sigma(\xi_t, \xi_{t-1}, \dots)$ -adapted processes  $v_k^{(n)}$ , which is identical to  $u_k^{(n)}$ , as all the past values  $\{z_s, s \leq t-n\}$  are equal to zero, is defined as:

$$\begin{cases} v_1^{(n)} = \ln |\xi_{t-n+1}| - \frac{\omega}{2} \\ v_{k+1}^{(n)} = \ln |\xi_{t-n+k+1}| - \frac{\omega}{2} + \sum_{i=1}^k \beta_i \cdot g_{\theta, \gamma} \left( v_{k+1-i}^{(n)}, \xi_{t-n+k+1-i} \right) \end{cases}.$$

In addition, by Lemma 2.1 of Martinet and McAleer (2015),  $u_n^{(n)} = \ln |z_t|$  when  $k = n$ . Therefore, to demonstrate the invertibility of the processes, it is essential to show that:

$$|v_n^{(n)} - \ln |z_t|| = |v_n^{(n)} - u_n^{(n)}| \xrightarrow[n \rightarrow \infty]{a.s.} 0. \quad (3.37)$$

If we connect the  $u_n^{(n)}$  and  $v_n^{(n)}$  to the notation for the EGARCH process in the last section, the  $v_n^{(n)}$  can be seen to be the same as the  $\ln |\tilde{z}_t|$ . Thus if (3.37) holds, that means:

$$|\ln |\tilde{z}_t| - \ln |z_t|| \xrightarrow[n \rightarrow \infty]{a.s.} 0,$$

and:

$$\left| \ln \left| \tilde{h}_t \right| - \ln |h_t| \right| \xrightarrow[n \rightarrow \infty]{a.s.} 0,$$

in the EGARCH processes. Martinet and McAleer (2015) provided an upper bound for a general EARCH( $\infty$ ) process in their Proposition 2.1, and invertibility conditions for the EGARCH( $p, q$ ) process in cases (i) and (iii) under certain conditions in their Propositions 4.1 and 4.2.

### 3.4.2 Invertibility of the ARMA-EGARCH process

Based on a similar idea, the approach of Martinet and McAleer (2015) for deriving the invertibility of the EARCH( $\infty$ ) can be used for the invertibility of the ARMA-EGARCH process, and the results are used in Section 3.5. Firstly, it is essential to understand the invertibility property of the process. This study considers the case where  $\gamma \geq |\theta|$ , in the EARCH( $\infty$ ) process as follows:

$$\begin{aligned} & \left| v_n^{(n)} - u_n^{(n)} \right| \\ \leq & \left| \ln |\xi_t| - \frac{\omega}{2} + \sum_{i=1}^{n-1} \beta_i \cdot g_{\gamma, \theta} \left( |v_{n-i}|^{(n)}, \xi_{n-i} \right) - \right. \\ & \left. \ln |\xi_t| + \frac{\omega}{2} - \sum_{i=1}^{n-1} \beta_i \cdot g_{\gamma, \theta} \left( \ln |z_{n-i}|, \xi_{n-i} \right) - \sum_{j=0}^{\infty} \beta_{j+n} \cdot g_{\gamma, \theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \right| \\ \leq & \sum_{j=0}^{\infty} \beta_{j+n} \left| g_{\gamma, \theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \right| \\ & + \sum_{i=1}^{n-1} \beta_i \left| \left( g_{\gamma, \theta} \left( |v_{n-i}|^{(n)}, \xi_{n-i} \right) - g_{\gamma, \theta} \left( \ln |z_{n-i}|, \xi_{n-i} \right) \right) \right| \\ \leq & \sum_{j=0}^{\infty} \beta_{j+n} \left| g_{\gamma, \theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \right| + \\ & \sum_{i=1}^{n-1} \beta_i \cdot \delta_{t-i} \left| \exp \left( \max \left( v_{n-j}^{(n)}, u_{n-j}^{(n)} \right) \right) \right| \left| v_{n-i}^{(n)} - u_{n-i}^{(n)} \right|, \end{aligned}$$

where  $\delta_t = \frac{1}{2}(\gamma + \theta \text{sign}(\xi_t))$  and  $g_\theta(\ln |z_t|, \xi_t) = -\delta_t |z_t|$ , the third inequality is obtained via Lemma 1.1 of Martinet and McAleer (2015) and Equation (3.37).

Accordingly, for the EGARCH(1,1) part in the ARMA-EGARCH process:

$$\begin{aligned}
& \left| v_n^{(n)} - u_n^{(n)} \right| \\
\leq & \left| \ln |\tilde{\xi}_t| - \frac{\omega}{2} + \sum_{i=1}^{n-1} \beta_i \cdot g_{\gamma, \theta} \left( |v_{n-i}|^{(n)}, \tilde{\xi}_{n-i} \right) - \right. \\
& \left. \ln |\xi_t| + \frac{\omega}{2} - \sum_{i=1}^{n-1} \beta_i \cdot g_{\gamma, \theta} \left( \ln |z_{n-i}|, \xi_{n-i} \right) - \sum_{j=0}^{\infty} \beta_{j+n} \cdot g_{\gamma, \theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \right| \\
\leq & \left| \ln |\tilde{\xi}_t| - \ln |\xi_t| \right| + \sum_{i=1}^{n-1} \beta_i \left| g_{\gamma, \theta} \left( |v_{n-i}|^{(n)}, \tilde{\xi}_{n-i} \right) - g_{\gamma, \theta} \left( \ln |z_{n-i}|, \xi_{n-i} \right) \right| \\
& + \sum_{j=0}^{\infty} \beta_{j+n} \left| g_{\gamma, \theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \right| \\
\leq & \left| \ln |\tilde{\xi}_t| - \ln |\xi_t| \right| + \sum_{j=0}^{\infty} \beta_{j+n} \left| g_{\gamma, \theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \right| \\
& + \sum_{i=1}^{n-1} \beta_i \left| \delta_{t-i} \left| \exp \left( \max \left( v_{n-i}^{(n)}, u_{n-i}^{(n)} \right) \right) \right| \left| v_{n-i}^{(n)} - u_{n-i}^{(n)} \right| + \left| \frac{\theta \left( \text{sign}(\tilde{\xi}_{n-i}) - \text{sign}(\xi_{n-i}) \right)}{2} \right| |z_{n-i}| \right|.
\end{aligned}$$

Since  $\tilde{\xi}_t$  converges to  $\xi_t$  as  $t$  tends to infinity, as shown in Lemma 3.5.3. By the proof of Lemma 3.5.3, suppose that  $\tilde{\xi}_t$  and  $\xi_t$  are different from zero and have the same sign, in which case:

$$\left| \ln |\tilde{\xi}_t| - \ln |\xi_t| \right| = \left| \ln \frac{|\tilde{\xi}_t|}{|\xi_t|} \right| \rightarrow 0, \text{ as } t \rightarrow \infty,$$

and:

$$\begin{aligned}
\left| v_n^{(n)} - u_n^{(n)} \right| & \leq \sum_{j=0}^{\infty} \beta_{j+n} \left| g_{\gamma, \theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \right| \\
& + \sum_{i=1}^{n-1} \beta_i \left| \delta_{t-i} \left| \exp \left( \max \left( v_{n-i}^{(n)}, u_{n-i}^{(n)} \right) \right) \right| \left| v_{n-i}^{(n)} - u_{n-i}^{(n)} \right| \right| + o_p(1),
\end{aligned}$$



where<sup>7</sup>:

$$\max \left( v_{n-i}^{(n)}, u_{n-i}^{(n)} \right) \leq \Delta = \ln |z_{t-i}| + \sum_{j=1}^{\infty} \beta_j \delta_{t-j-i} |z_{t-j-i}|.$$

This upper bound for  $\max \left( v_{n-i}^{(n)}, u_{n-i}^{(n)} \right)$  is provided in Martinet and McAleer (2015).

By their Proposition 2.1, we are able to get:

$$\begin{aligned} & \left| v_n^{(n)} - u_n^{(n)} \right| \\ & \leq \sum_{j=0}^{\infty} |\delta_{t-n-j}| |z_{t-n-j}| \\ & \quad \left( \beta_{j+n} + \sum_{p=1}^{n-1} \sum_{i_1, \dots, i_p \in A_p^{(n)}} \prod_{j=1}^p \beta_{i_j} \prod_{j=1}^p \delta_{t-\hat{S}_j} \exp \left( \sum_{j=1}^p \Delta_{t-\hat{S}_j} \right) \beta_{j+n-\hat{S}_j} \right) + o_p(1), \end{aligned}$$

with:

$$\hat{S}_j = \sum_{j=1}^p i_j; \quad A_p^{(n)} = \left\{ i_1 \geq 1, \dots, i_p \geq 1 : \hat{S}_p \leq n-1 \right\}.$$

---

<sup>7</sup>For the  $\max \left( v_{n-i}^{(n)}, u_{n-i}^{(n)} \right)$ , since:

$$\begin{aligned} v_{n-i}^{(n)} &= \ln |\xi_{t-i}| - \frac{\omega}{2} + \sum_{j=1}^{n-i-1} \beta_j \cdot g_{\gamma, \theta} \left( |v_{n-i-j}|^{(n)}, \xi_{n-i-j} \right) \\ &= \ln |z_{t-i}| + \sum_{j=1}^{\infty} \beta_j \cdot \delta_{t-i-j} |z_{t-i-j}| + \sum_{j=1}^{n-i-1} \beta_j \cdot g_{\gamma, \theta} \left( |v_{n-i-j}|^{(n)}, \xi_{n-i-j} \right) \end{aligned}$$

by equation (3.36), then when  $\gamma \geq |\theta|$ , by the above equation,  $g_{\gamma, \theta} \leq 0$ ,  $\delta_{t-i} \geq 0$  and:  $\max \left( v_{n-i}^{(n)}, u_{n-i}^{(n)} \right) = \ln |z_{t-i}| + \left( v_{n-i}^{(n)} - \ln |z_{t-i}| \right)^+$ , then:

$$v_{n-i}^{(n)} \leq \ln |z_{t-i}| + \sum_{j=1}^{\infty} \beta_j \delta_{t-j-i} |z_{t-j-i}|.$$

For the EGARCH(1,1) part in the ARMA-EGARCH process:

$$\begin{aligned}
& |v_n^{(n)} - u_n^{(n)}| \\
& \leq \sum_{j=0}^{\infty} \beta^j |\delta_{t-n-j}| |z_{t-n-j}| \left( \beta^{n-1} + \sum_{p=1}^{n-1} \beta^{n-1-p} \sum_{i_1, \dots, i_p \in I_p^{(n)}} \prod_{j=1}^p \delta_{t-\hat{s}_j} \right. \\
& \quad \left. \exp \left( \sum_{j=1}^p \Delta_{t-\hat{s}_j} \right) \beta_{j+n-\hat{s}_j} \right) + o_p(1) \\
& = \sum_{j=0}^{\infty} \beta^j |\delta_{t-n-j}| |z_{t-n-j}| \left( \beta^{n-1} + \sum_{p=1}^{n-1} \beta^{n-1-p} \right. \\
& \quad \left. \sum_{1 \leq s_1 < \dots < s_p \leq n-1} \exp \left( \sum_{j=1}^p \ln |\delta_{t-\hat{s}_j}| + \sum_{j=1}^p \Delta_{t-\hat{s}_j} \right) \right) + o_p(1).
\end{aligned}$$

Following Proposition 4.1 of Martinet and McAleer (2015), the EGARCH(1,1) part is invertible under Condition (3.38):

$$E \left[ \frac{\delta_t |z_t|}{1 - \beta} \right] + \ln (\beta + E [\delta_t |z_t|]) < 0. \quad (3.38)$$

This means that in the ARMA-EGARCH process:

$$|\ln |\tilde{z}_t| - \ln |z_t|| = o_p(1),$$

and:

$$|\ln |\tilde{h}_t| - \ln |h_t|| = o_p(1).$$

### 3.4.3 Invertibility of the HYEGARCH process

A similar procedure to that of Martinet and McAleer (2015) is used to show the invertibility of the HYEGARCH process. The HYEGARCH process can be rewritten as:

$$\ln |z_t| = \ln |\xi_t| - \frac{\omega}{2} + \sum_{j=1}^{\infty} b_j g_{\theta} (\ln |z_{t-j}|, \xi_{t-j}). \quad (3.39)$$

At a fixed point  $t$  and with an independent shock  $z_t$ , we are able to get an exact value for any positive  $n$  from the observed shocks  $\xi_t$ , when  $(z_t)_{s \leq t-n}$  are known. By extending  $n$  steps backward the 'exact' recursion, the process  $U_k^{(n)}$  is defined as:

$$\begin{cases} U_1^{(n)} = \ln |\xi_{t-n+1}| - \frac{\omega}{2} + \sum_{j=0}^{\infty} b_{j+1} g_{\theta} (\ln |z_{t-n-j}|, \xi_{t-n-j}), \\ U_{k+1}^{(n)} = \ln |\xi_{t-n+k+1}| - \frac{\omega}{2} + \sum_{i=1}^k b_i g_{\theta} (U_{k+1-i}^{(n)}, \xi_{t-n+k+1-i}) \\ \quad + \sum_{j=0}^{\infty} b_{j+1+k} g_{\theta} (\ln |z_{t-n-j}|, \xi_{t-n-j}). \end{cases}$$

As  $n$  tends to infinity, the value of  $\ln |z_t|$  depends more and more on the past observations of  $\xi_t$  and less on the past information on  $(z_t)_{s \leq t-n}$ . Thus if we assume the past values of  $(z_t)_{s \leq t-n}$  to be zero, then the series  $V_k^{(n)}$ , which is a  $\sigma(\xi_t, \xi_{t-1}, \dots)$ -adapted process and is identical to  $U_k^{(n)}$ , and is defined as:

$$\begin{cases} V_1^{(n)} = \ln |\xi_{t-n+1}| - \frac{\omega}{2}, \\ V_{k+1}^{(n)} = \ln |\xi_{t-n+k+1}| - \frac{\omega}{2} + \sum_{i=1}^k b_i g_{\theta} (V_{k+1-i}^{(n)}, \xi_{t-n+k+1-i}). \end{cases}$$

To derive the invertibility conditions of the HYGARCH process, it is essential to find an upper bound for the following inequality. That is if  $\left| V_n^{(n)} - \ln |z_t| \right| = \left| V_n^{(n)} - U_n^{(n)} \right|$  converges to 0 almost surely as  $n$  tends to infinity, then the invertibility of the HYGARCH process holds. Therefore, by Lemma 1.1 of Martinet and McAleer (2015):

$$\begin{aligned} & \left| V_n^{(n)} - U_n^{(n)} \right| \tag{3.40} \\ & \leq \sum_{j=0}^{\infty} b_{j+n} \left| g_{\theta} (\ln |z_{t-n-j}|, \xi_{t-n-j}) \right| + \\ & \quad \sum_{i=1}^{n-1} b_i \delta_{t-i} \left| \exp \left( \max \left( V_{n-i}^{(n)}, U_{n-i}^{(n)} \right) \right) \right| \left| V_{n-i}^{(n)} - U_{n-i}^{(n)} \right| \\ & \leq \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| + \\ & \quad \sum_{i=1}^{n-1} b_i \delta_{t-i} \left| \exp (\Delta_{t-i}) \right| \left| V_{n-i}^{(n)} - U_{n-i}^{(n)} \right| \\ & = a_1(t), \end{aligned}$$

where<sup>8</sup>:

$$g_\theta (\ln |z_{t-n-j}|, \xi_{t-n-j}) = -\delta_{t-n-j} |z_{t-n-j}|,$$

$$\max \left( V_{n-i}^{(n)}, U_{n-i}^{(n)} \right) \leq \Delta_{t-i} = \ln |z_{t-i}| + \sum_{j=1}^{\infty} b_j \delta_{t-j-i} |z_{t-j-i}|,$$

when  $|\theta| \leq 1$ . Follow Lemma 2.2 of Martinet and McAleer (2015), we are able to get that the HYGARCH process also satisfies this Lemma, in this case that is:

**Lemma 3.4.1**

$$|V_n^{(n)} - U_n^{(n)}| \leq a_k, \text{ for any } k \in [1, n).$$

Therefore, similar to the inequality in Proposition 2.1 of Martinet and McAleer (2015), the  $|V_n^{(n)} - U_n^{(n)}|$  can also be rewritten as follows:

$$\begin{aligned} & |V_n^{(n)} - U_n^{(n)}| \tag{3.41} \\ & \leq \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| \\ & \quad + \sum_{j=0}^{\infty} \delta_{t-n-j} |z_{t-n-j}| \left( \sum_{p=1}^{n-1} \sum_{j_1, \dots, j_p \in A_p^{(n)}} \hat{\Pi}_p \hat{D}_p \exp \left( \sum_{i=1}^p \Delta_{t-\hat{S}_i} \right) b_{j+n-\hat{S}_p} \right), \end{aligned}$$

---

<sup>8</sup>For the  $\max \left( V_{n-i}^{(n)}, U_{n-i}^{(n)} \right)$ , follow a similar procedure to Martinet and McAleer (2015), for the upper bound of  $\max \left( V_{n-i}^{(n)}, U_{n-i}^{(n)} \right)$ , in the case of  $|\theta| \leq 1$ , since:

$$\begin{aligned} & V_{n-i}^{(n)} \\ & = \ln |\xi_{t-i}| - \frac{\omega}{2} + \sum_{j=1}^{n-i-1} b_j g_\theta \left( V_{n-i-j}^{(n)}, \xi_{t-i-j} \right) \\ & = \ln |z_{t-i}| + \sum_{j=1}^{\infty} b_j \delta_{t-j-i} |z_{t-j-i}| \\ & \quad + \sum_{j=1}^{n-i-1} b_j \cdot g_\theta \left( V_{n-j-i}^{(n)}, \xi_{t-j-i} \right), \end{aligned}$$

and  $g_\theta \leq 0, \delta_{t-j-i} \geq 0$ , and:

$$\max \left( V_{n-i}^{(n)}, U_{n-i}^{(n)} \right) = \ln |z_{t-i}| + \left( V_{n-i}^{(n)} - \ln |z_{t-i}| \right)^+,$$

thus,  $V_{n-i}^{(n)} \leq \ln |z_{t-i}| + \sum_{j=1}^{\infty} b_j \delta_{t-j-i} |z_{t-j-i}|$ .

where:

$$\begin{aligned} \hat{S}_p &= \sum_{i=1}^p j_i & \hat{\Pi}_p &= \prod_{i=1}^p b_{j_i} \\ A_p^{(n)} &= \{j_1 \geq 1, \dots, j_p \geq 1 : \hat{S}_p \leq n-1\} & \hat{D}_p &= \prod_{i=1}^p \delta_{t-\hat{S}_i} \end{aligned}$$

Thus Proposition 2.1 of Martinet and McAleer (2015) also satisfies the HYE-GARCH process.

However, since the lag coefficients in the HYE-GARCH process decay hyperbolically rather than geometrically, so the convergence of  $|V_n^{(n)} - U_n^{(n)}|$  is different from that in the EGARCH processes. It is more complex to derive the conditions for the invertibility of the HYE-GARCH process. This is a work in progress and will continue to be a focus of my future studies.

### 3.4.4 The invertibility of the derivatives of the ARMA-EGARCH process

To establish the asymptotic normality of the ARMA-EGARCH process, it is also essential to derive that the first derivative of  $v_n^{(n)}$  converges to that of the  $u_n^{(n)}$  in the EGARCH error. Firstly, let us consider the EARCH( $\infty$ ) process with respect to the parameter  $\beta$ . These are:

$$\begin{aligned} \frac{\partial v_n^{(n)}}{\partial \beta} &= -\frac{w}{(1-\beta)^2} + \sum_{i=1}^{n-1} (i-1)\beta^{i-2} \cdot g_{\gamma,\theta} \left( |v_{n-i}|^{(n)}, \xi_{n-i} \right) \\ &\quad + \sum_{i=1}^{n-1} \beta^{i-1} \cdot g_{\gamma,\theta} \left( |v_{n-i}|^{(n)}, \xi_{n-i} \right) \frac{\partial v_{n-i}^{(n)}}{\partial \beta}, \end{aligned}$$

and:

$$\begin{aligned}
\frac{\partial u_n^{(n)}}{\partial \beta} &= -\frac{w}{(1-\beta)^2} + \sum_{i=1}^{n-1} (i-1)\beta^{i-2} \cdot g_{\gamma,\theta} \left( |u_{n-i}|^{(n)}, \xi_{n-i} \right) \\
&\quad + \sum_{i=1}^{n-1} \beta^{i-1} \cdot g_{\gamma,\theta} \left( |u_{n-i}|^{(n)}, \xi_{n-i} \right) \frac{\partial u_{n-i}^{(n)}}{\partial \beta} \\
&\quad + \sum_{j=0}^{\infty} (j+n-1)\beta^{j+n-2} \cdot g_{\gamma,\theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \\
&\quad + \sum_{j=0}^{\infty} \beta^{j+n-1} \cdot g_{\gamma,\theta} \left( \ln |z_{t-n-j}|, \xi_{t-n-j} \right) \frac{\partial u_{t-n-j}^{(n)}}{\partial \beta}.
\end{aligned}$$

Then:

$$\begin{aligned}
&\left| \frac{\partial v_n^{(n)}}{\partial \beta} - \frac{\partial u_n^{(n)}}{\partial \beta} \right| \\
&\leq \sum_{i=1}^{n-1} \beta^{i-1} \delta_{t-i} \exp(v_{n-i}^{(n)}) \left| \frac{\partial v_{n-i}^{(n)}}{\partial \beta} - \frac{\partial \ln |z_{n-i}|}{\partial \beta} \right| \\
&\quad + \sum_{i=1}^{n-1} (i-1)\beta^{i-2} \delta_{t-i} \exp(\Delta_{t-i}) \left| v_{n-i}^{(n)} - u_{n-i}^{(n)} \right| \\
&\quad + \sum_{i=1}^{n-1} \beta^{i-1} \delta_{t-i} \exp(\Delta_{t-i}) \left| v_{n-i}^{(n)} - u_{n-i}^{(n)} \right| \left| \frac{\partial \ln |z_{n-i}|}{\partial \beta} \right| \\
&\quad + \sum_{j=0}^{\infty} (j+n-1)\beta^{j+n-2} \delta_{t-n-j} |z_{t-n-j}| \\
&\quad + \sum_{j=0}^{\infty} \beta^{j+n-1} \delta_{t-n-j} |z_{t-n-j}| \left| \frac{\partial \ln z_{t-n-j}}{\partial \beta} \right|.
\end{aligned}$$

Similar to Lemma 2.2 and Proposition 2.1 of Martinet and McAleer (2015), we

are able obtain an upper bound for  $\left| \partial v_n^{(n)} / \partial \beta - \partial u_n^{(n)} / \partial \beta \right|$ , that is:

$$\begin{aligned}
& \left| \frac{\partial v_n^{(n)}}{\partial \beta} - \frac{\partial u_n^{(n)}}{\partial \beta} \right| \\
& \leq \sum_{j=0}^{\infty} \delta_{t-n-j} |z_{t-n-j}| \\
& \quad \left\{ \left( \sum_{p=1}^{n-1} \sum_{l=p}^{n-1} \sum_{j_1, \dots, j_l \in A_l^{(n)}} (j_p - 1) \beta^{j_p - 2} \prod_{i=1, i \neq p}^l \beta^{j_i - 1} \prod_{i=1}^l \delta_{t-\hat{s}_i} \exp \left( \sum_{i=1}^l \Delta_{t-\hat{s}_i} \right) \beta^{j+n-\hat{s}_l-1} \right) \right. \\
& \quad + \left( \sum_{p=1}^{n-1} \sum_{l=p}^{n-1} \sum_{j_1, \dots, j_l \in A_l^{(n)}} \prod_{i=1}^l \beta^{j_i - 1} \prod_{i=1}^l \delta_{t-\hat{s}_i} \exp \left( \sum_{i=1}^l \Delta_{t-\hat{s}_i} \right) \beta^{j+n-\hat{s}_l-1} \left| \frac{\partial u_{n-\hat{s}_p}}{\partial \beta} \right| \right) \\
& \quad + \left( \beta^{j+n-1} + \sum_{p=1}^{n-1} \sum_{j_1, \dots, j_p \in A_p^{(n)}} \prod_{i=1}^p \beta^{j_i - 1} \prod_{i=1}^p \delta_{t-\hat{s}_i} \exp \left( \sum_{i=1}^p \Delta_{t-\hat{s}_i} \right) \beta^{j+n-\hat{s}_i-1} \right) \left| \frac{\partial \ln z_{t-n-j}}{\partial \beta} \right| \\
& \quad \left. + \left( (j+n-1) \beta^{j+n-2} + \sum_{p=1}^{n-1} \prod_{i=1}^{p-1} \beta^{j_i - 1} \prod_{i=1}^{p-1} \delta_{t-\hat{s}_i} \exp \left( \sum_{i=1}^{p-1} \Delta_{t-\hat{s}_i} \right) (j+n-\hat{s}_i-1) \beta^{j+n-\hat{s}_i-1} \right) \right\}
\end{aligned}$$

Then, similar to Proposition 4.1 of Martinet and McAleer (2015), in the EGARCH(1, 1) case, it is clear that under condition (3.38):

$$E \left[ \frac{\delta_t |z_t|}{1 - \beta} \right] + \ln (\beta + E [\delta_t |z_t|]) < 0,$$

$\left| \partial v_n^{(n)} / \partial \beta - \partial u_n^{(n)} / \partial \beta \right|$  also converges to zero as  $n$  tends to  $\infty$ . And following a similar procedure, we are able to derive the convergence of the second derivative with respect to  $\beta$ , that is:

$$\left| \frac{\partial^2 v_n^{(n)}}{\partial \beta \partial \beta} - \frac{\partial^2 u_n^{(n)}}{\partial \beta \partial \beta} \right| \xrightarrow[n \rightarrow \infty]{a.s.} 0.$$

We are also able to obtain the convergence of first derivation and second derivative with respect to other parameters from EGARCH error in the ARMA-EGARCH process.

### 3.5 Asymptotic theory of the QMLE in ARMA(1,1) with an EGARCH(1,1) error

This section aims to establish the asymptotic properties of the QMLE in the ARMA(1,1)-EGARCH(1,1) process. To establish the asymptotic theory of the QMLE, it is necessary to consider the uniform convergence. Ling and McAleer (2003) introduced the modified uniform convergence theorem, based on Theorem 4.2.1 of Amemiya (1985), which can be applied when deriving the asymptotic properties of the QMLE in the ARMA-EGARCH model. This theorem makes the verification of the limiting properties of the QMLE easier in the ARMA-EGARCH model. The modified uniform convergence theorem is as follows:

**Theorem 3.5.1** (see Theorem 3.1 in Ling and McAleer (2003)): *Let  $f(x_t, \lambda)$  be a measurable function of  $x_t$  in Euclidean space for each  $\lambda \in \Theta$ , a compact subset of  $R^m$  (Euclidean  $m$ -space), and a continuous function of  $\lambda \in \Theta$  for each  $x_t$ . Suppose that  $x_t$  is a sequence of strictly stationary and ergodic time series, such that  $Ef(x_t, \lambda) = 0$  and  $E \sup_{\lambda \in \Theta} |f(x_t, \lambda)| < \infty$ . Then  $\sup_{\lambda \in \Theta} |\frac{1}{n} \sum_{t=1}^n f(x_t, \lambda)| = o_p(1)$ .*

The proof of the uniform convergence theorem can be found in Ling and McAleer (2003). Ling and McAleer (2003) derived the uniform convergence theorem using the ergodic theorem that replaced the Kolmogorov law of large numbers in the proof of Theorem 4.2.1 of Amemiya (1985).



### 3.5.1 Consistency of the QMLE in ARMA(1,1)-EGARCH(1,1) process

This subsection aims to derive the conditions for the consistency of the QMLE. The modified Theorem 4.1.1 of Amemiya (1985) is applied to verify the consistency of the QMLE in the ARMA(1,1)-EGARCH(1,1) process. The observed likelihood process may not be a stationary process. Thus the unobserved log-likelihood function, conditional on infinite past observations, is convenient for showing the convergence of the likelihood function. The following lemmas are using to support the consistency of the QMLEs of the ARMA(1,1)-EGARCH(1,1) process.

**Lemma 3.5.1** *Suppose that  $y_t$  is generated by the ARMA(1,1)-EGARCH(1,1) process under Assumptions 3.3.1–3.3.5 and  $E(y_t)^2 < \infty$ , and define  $\bar{L}(\lambda) = E[l_t(\lambda)]$ . Accordingly,  $\bar{L}(\lambda) < \infty$  for all  $\lambda \in \Theta$  and  $\sup_{\lambda \in \Theta} |L_n(\lambda) - \bar{L}(\lambda)| = o_p(1)$ .*

This lemma proves the existence of  $\bar{L}(\lambda) = E[l_t(\lambda)]$ , which is the corresponding limit of  $L_n(\lambda)$ , and shows that the difference between the unobserved log-likelihood functions  $L_n(\lambda)$  and  $\bar{L}(\lambda)$  converges to zero in probability.

**Lemma 3.5.2** *Suppose that  $y_t$  is generated by the ARMA(1,1)-EGARCH(1,1) process under Assumptions 3.3.1–3.3.5, then  $\bar{L}(\lambda) = E[l_t(\lambda)]$  achieves a unique maximum at  $\lambda_0$ .*

This lemma confirms the identification condition for the consistency. This means that  $E[l_t(\lambda)]$  is uniquely maximised at  $\lambda_0$ . It also verifies the existence of  $\partial \xi_t / \partial \varphi$ . The same procedure as that used in Section 5.4.1 of Straumann (2005) is applied to show the existence of  $\partial \ln h_t / \partial \varsigma$ . This implies that  $(\varphi - \varphi_0)' (\partial \xi_t / \partial \varphi) = 0$

can be obtained if and only if  $\varphi = \varphi_0$ , and that  $(\varsigma - \varsigma_0)' (\partial \ln h_t / \partial \varsigma) = 0$  can be supported if and only if  $\varsigma = \varsigma_0$  almost surely.

**Lemma 3.5.3** *Suppose that  $y_t$  is generated by the ARMA(1,1)-EGARCH(1,1) process under Assumptions 3.3.1–3.3.5, and the invertibility condition (3.38) holds, then  $\sup_{\lambda \in \Theta} |L_n(\lambda) - \tilde{L}_n(\lambda)| = o_p(1)$ .*

This lemma ensures that the difference between the unobserved log-likelihood function  $L_n(\lambda)$  and the likelihood function  $\tilde{L}_n(\lambda)$  of the ARMA(1,1)-EGARCH(1,1) model converges to zero in probability. This means that the gap between the likelihood process that is conditional on infinite past observations and the likelihood function depends on the finite past information can be asymptotically ignored.

Lemmas 3.5.1, 3.5.2 and 3.5.3 ensure that the difference between  $\tilde{L}_n(\lambda)$  and  $\bar{L}(\lambda)$  converges to zero in probability at  $\lambda_0$ . According to these Lemmas, all the conditions for consistency (Amemiya 1985, Theorem 4.1.1) hold and thus  $\hat{\lambda}_n$  is the consistent estimator of  $\lambda_0$ . This result is presented as Theorem 3.5.2.

**Theorem 3.5.2** *Under Assumptions 3.3.1–3.3.5 of the ARMA(1,1)-EGARCH(1,1) process and the invertibility condition (3.38), the QMLE  $\hat{\lambda}_n$  is consistent, which implies that  $\hat{\lambda}_n$  converges to  $\lambda_0$  as  $n$  tends to infinity.*

### **3.5.2 Asymptotic normality of the QMLE in ARMA(1,1) with an EGARCH(1,1) error**

This subsection focuses on establishing the asymptotic normality of the QMLE for the ARMA(1,1)-EGARCH(1,1) model. Theorem 3.5.1 is used, which places a

weaker restriction on the moment conditions. The Taylor expansion of the derivatives of the log-likelihood function is also applied to derive the asymptotic normality of the QMLE. Ling and McAleer (2003) investigated the asymptotic normality of the QMLE for the vector ARMA-GARCH model under a sixth-order moments conditions. Francq and Zakořan (2004) derived the asymptotic normality of the ARMA-GARCH model under a fourth-order moment assumption. This subsection follows a similar procedure to that of Ling and McAleer (2003) to establish the asymptotic normality properties of the ARMA(1, 1)-EGARCH(1, 1) model.

Straumann (2005) derived the conditions for the asymptotic normality of the EGARCH(1, 1) process with  $\beta = 0$  in Theorem 5.7.9. Wintenberger (2013) extended these results to derive the asymptotic normality of the SQMLE by applying the (MX) condition, replacing the uniform moments condition on the compact set  $\Theta$  with the likelihood function and its derivative, and proved that the (MM) assumption is necessary and sufficient for the existence of the asymptotic covariance matrix. The (MX) and (MM) conditions are as follows:

(MX): The EGARCH(1,1) volatility  $(h_t)$  constitutes a geometrically ergodic Markov chain.

$$\text{(MM): } E[z_0^4] < \infty \text{ and } E\left[\left(\beta_0 - \frac{1}{2}((\theta_0 z_0 + \gamma_0 |z_0|))^2\right)\right] < 1.$$

According to Theorem 4.1.3 of Amemiya (1985), some conditions need to be satisfied in order to establish asymptotic normality. Firstly, the QMLE  $\hat{\lambda}_n$  needs to be a consistent estimator of  $\lambda_0$ . This condition was derived in the previous subsection. Secondly, the existence of  $\Omega_0$  needs to be shown and the gradient of  $\tilde{L}_n$ , which is  $n^{-1/2} \sum_{t=1}^n (\partial \tilde{l}_{0t} / \partial \lambda)$ , needs to converge to  $N(0, \Omega_0)$  in distribution. The third condition is that the  $\tilde{L}_n$  function is twice continuously differentiable. This

means the Hessian matrix of  $\tilde{L}_n$ , (i.e.  $n^{-1} \sum_{t=1}^n (\partial \tilde{l}_t^2 / \partial \lambda \partial \lambda')$ ), exists and is continuous in the parameter space  $\Theta$ . It also requires  $n^{-1} \sum_{t=1}^n (\partial \tilde{l}_t^2 / \partial \lambda \partial \lambda')$  to converge to  $\Sigma_0$  for any sequence  $\lambda_n$ , such that  $\lambda_n \rightarrow \lambda_0$  in probability. If all of the conditions hold, then  $\sqrt{n}(\hat{\lambda}_n - \lambda_0)$  converges to  $N(0, \Sigma_0^{-1} \Omega_0 \Sigma_0^{-1})$  in distribution. For a consistent estimator of the covariance matrix, it is also necessary to show that  $\Sigma_0$  and  $\Omega_0$  can be estimated consistently by  $\hat{\Sigma}_n$  and  $\hat{\Omega}_n$ , respectively.

**Lemma 3.5.4** *Suppose that  $y_t$  is generated by the ARMA(1,1)-EGARCH(1,1) model under Assumptions 3.3.1–3.3.5, and the invertibility condition holds, and  $E[y_t]^4 < \infty$ , and  $E[z_{0t}]^4 < \infty$ , and:*

$$E \left| \beta_0 - \frac{1}{2} \theta_0 z_{0t-1} - \frac{1}{2} \gamma_0 |z_{0t-1}| \right|^2 < 1,$$

then  $\Omega_0 = E[(\partial l_{0t} / \partial \lambda)(\partial l_{0t} / \partial \lambda)']$  is finite. Furthermore, if  $\Omega_0 > 0$ , then:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \tilde{l}_{0t}}{\partial \lambda} \rightarrow^d N(0, \Omega_0), \quad (3.42)$$

where  $\lambda = (\varphi', \zeta)'$ ,  $\partial l_{0t} / \partial \lambda = \partial l_t / \partial \lambda|_{\lambda_0}$ ,  $\partial \tilde{l}_{0t} / \partial \lambda = \partial \tilde{l}_t / \partial \lambda|_{\lambda_0}$ .

This lemma shows that the score function converges to  $N(0, \Omega_0)$  in distribution.

**Lemma 3.5.5** *Suppose that  $y_t$  is generated by the ARMA(1,1)-EGARCH(1,1) model under Assumptions 3.3.1–3.3.5, and the invertibility condition holds, and  $E[y_t]^4 < \infty$ , and  $E[z_{0t}]^4 < \infty$ . Then, (i)*

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} - E \left[ \frac{\partial^2 l_{0t}}{\partial \lambda \partial \lambda'} \right] \right\| = o_p(1);$$

(ii)

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \tilde{l}_t}{\partial \lambda \partial \lambda'} \right] \right\| = o_p(1). \quad (3.43)$$

Part (i) of this lemma shows that the difference between  $\frac{1}{n} \sum_{t=1}^n \partial^2 l_t / \partial \lambda \partial \lambda'$  and  $E[\partial^2 l_t / \partial \lambda \partial \lambda']$  converges to zero in probability. Part (ii) of this lemma shows that the difference between the second derivative of the unobserved log-likelihood function and that of the observed function asymptotically equal zero. This lemma implies that  $\widehat{\Sigma}_n = \frac{1}{n} \sum_{t=1}^n \partial^2 \tilde{l}_t / \partial \lambda \partial \lambda'$  is a consistent estimator of  $\Sigma_0 = E[\partial^2 l_{0t} / \partial \lambda \partial \lambda']$ .

**Lemma 3.5.6** *Under Assumptions 3.3.1–3.3.5, suppose that the invertibility condition holds, and  $E[y_t]^4 < \infty$ , and  $E[z_{0t}]^4 < \infty$ , and*

$$E \left| \beta_0 - \frac{1}{2} \theta_0 z_{0t-1} - \frac{1}{2} \gamma_0 |z_{0t-1}| \right|^2 < 1.$$

If  $\sqrt{n}(\lambda_n - \lambda_0) = o_p(1)$ , then, (i)

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} - \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right]_{\lambda_n} = o_p(1);$$

(ii)

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial \tilde{l}_t}{\partial \lambda} \frac{\partial \tilde{l}_t}{\partial \lambda'} \right]_{\lambda_n} = \Omega_0 + o_p(1).$$

This lemma shows that  $\widehat{\Omega}_n = \frac{1}{n} \sum_{t=1}^n \left[ (\partial \tilde{l}_t / \partial \lambda)(\partial \tilde{l}_t / \partial \lambda') \right]_{\lambda_n}$  is a consistent estimator of  $\Omega_0$ .

The conditions for the asymptotic normality (Amemiya, 1985, Theorem 4.1.3) have been verified by Lemmas 3.5.4–3.5.6 under the following assumptions and restrictions. Under the Assumptions 3.3.1–3.3.5, the conditions of  $E[y_t]^4 < \infty$ ,  $E[z_{0t}]^4 < \infty$ , and  $\Omega_0 > 0$  and the invertibility condition, Lemmas 3.5.4–3.5.6 ensure that  $n^{1/2}(\widehat{\lambda}_n - \lambda_0)$  converges to  $N(0, \Sigma_0^{-1} \Omega_0 \Sigma_0^{-1})$  in distribution. The asymptotic variance is  $\Sigma_0^{-1} \Omega_0 \Sigma_0^{-1}$  and it can be estimated consistent by  $\widehat{\Sigma}_n^{-1} \widehat{\Omega}_n \widehat{\Sigma}_n^{-1}$ . Thus, we are able to establish the following theorem.

**Theorem 3.5.3** *Assume that  $y_t$  is generated by the ARMA(1,1)-EGARCH(1,1) model under Assumptions 3.3.1-3.3.5, and the invertibility condition holds, and the conditions of  $E[y_t]^4 < \infty$ , and  $E[z_{0t}]^4 < \infty$ , and*

$$E \left| \beta_0 - \frac{1}{2}\theta_0 z_{0t-1} - \frac{1}{2}\gamma_0 |z_{0t-1}| \right|^2 < 1$$

*hold. Then  $n^{1/2}(\widehat{\lambda}_n - \lambda_0)$  converges to  $N(0, \Sigma_0^{-1}\Omega_0\Sigma_0^{-1})$  in distribution, where  $\lambda = (\varphi', \varsigma)'$ ,  $\Sigma_0 = -E[\partial^2 l_{0t}/\partial\lambda\partial\lambda']$  and  $\Omega_0 = E[(\partial l_{0t}/\partial\lambda)(\partial l_{0t}/\partial\lambda)']$ . Moreover, the  $\widehat{\Sigma}_n$  and  $\widehat{\Omega}_n$  are consistent estimators of  $\Sigma_0$  and  $\Omega_0$ .*

### 3.6 The asymptotic theory of the QMLE in HY/FIEGARCH processes

This section aims to establish the asymptotic theory for the HY/FIEGARCH processes. In the previous chapter, it can be seen that the HYGARCH process embodies the hyperbolic memory in volatility and the FIEGARCH process captures the long memory in volatility. In both cases, the effect of the shocks to volatility decay very slowly. It is hard to establish the asymptotic theory of the FIEGARCH process. This section provides the consistency of the QMLE in the HYGARCH process and investigates the asymptotic normality in the HY/FIEGARCH processes.

### 3.6.1 Asymptotic theory of the QMLE in the HYGARCH process

The slowly decay rate of the lag coefficients makes it harder to derive the asymptotic property of the QMLE for an objective function that is based on finite past information. First, the asymptotic theory of the estimation, given infinite past information, is considered and then an extension to estimations depending on finite past information is derived.

#### Estimations with infinite past information

This part aims to derive the asymptotic property of  $\vartheta_n$ , which is:

$$\vartheta_n = \arg \max_{\vartheta \in \Theta} L_n(\vartheta).$$

Lemmas 3.6.1 to 3.6.2 ensure that  $\vartheta_n$  is a consistent estimator of  $\vartheta_0$ .

**Lemma 3.6.1** *Supposing that  $\xi_t$  is generated by the HYGARCH(0, d, 0) process under Assumptions 3.3.7 and 3.3.9, and  $E[\xi_t^2] < \infty$ , define  $\bar{L}(\vartheta) = E[l_t(\vartheta)]$ . Then,  $\bar{L}(\vartheta) < \infty$  for all  $\vartheta \in \Theta$  and  $\sup_{\lambda \in \Theta} |L_n(\vartheta) - \bar{L}(\vartheta)| = o_p(1)$ .*

Lemma 3.6.1 is similar to Lemma 3.5.1 and shows that the difference between the unobserved log-likelihood function  $L_n(\vartheta)$  and the corresponding limit  $E[l_t(\vartheta)]$  converges to zero in probability in the HYGARCH process. In order to demonstrate the uniform convergence between these two objective functions, the existence of  $E|l_t(\vartheta)|$  first needs to be established. If this is supported, then the uniform convergence theorem can be applied to show the convergence of these two functions.

**Lemma 3.6.2** *If we suppose that  $\xi_t$  is generated by the HYEGARCH(0, d, 0) process under Assumptions 3.3.7 and 3.3.9,  $\bar{L}(\lambda) = E[l_t(\vartheta)]$  achieves a unique maximum at  $\vartheta_0$ .*

This lemma shows that  $\bar{L}(\lambda)$  has a unique maximum value at a true parameter point  $\vartheta_0$ . According to Lemmas 3.6.1 and 3.6.2,  $\vartheta_n$  is a consistent estimator of the true parameter,  $\vartheta_0$ . The remainder of this part studies the asymptotic distribution of  $\vartheta_n$  in the pure HYEGARCH process.

First, the asymptotic distribution of the gradient of the  $L_n(\vartheta)$  function in Lemma 3.6.3 is considered. By the mean value theorem, we have:

$$0 = \frac{1}{n} \sum_{t=1}^n \frac{\partial l_t}{\partial \vartheta} = \frac{1}{n} \sum_{t=1}^n \frac{\partial l_{0t}}{\partial \vartheta} + \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \vartheta \partial \vartheta'} \Big|_{\vartheta^\dagger} (\vartheta_n - \vartheta_0),$$

where  $\vartheta^\dagger$  lies between  $\vartheta_n$  and  $\vartheta_0$ .

**Lemma 3.6.3** *Suppose that  $\xi_t$  is generated by the HYEGARCH(0, d, 0) process under Assumptions 3.3.6, 3.3.7 and 3.3.9, and  $E[\xi_t^2] < \infty$ , and  $E[z_{0t}]^4 < \infty$ , and  $E\left[\sum_{j=1}^{\infty} b_{0j} \left(-\frac{1}{2}\theta_0 z_{0t-j} - \frac{1}{2}|z_{0t-j}|\right)\right]^2 < 1$ . Then  $\Omega_{01} = E[(\partial l_{0t}/\partial \vartheta)(\partial l_{0t}/\partial \vartheta)']$  is finite. Furthermore, if  $\Omega_{01} > 0$ , then:*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial l_{0t}}{\partial \vartheta} \rightarrow^d N(0, \Omega_{01}), \quad (3.44)$$

where  $\vartheta = (\omega, c, d, \theta)'$ ,  $\partial l_{0t}/\partial \vartheta = \partial l_t/\partial \vartheta|_{\vartheta_0}$ .

This lemma shows that the score function  $\frac{1}{\sqrt{n}} \sum_{t=1}^n \partial l_{0t}/\partial \vartheta$  converges to  $N(0, \Omega_{01})$  in distribution. It is also essential to prove that:

$$\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \vartheta \partial \vartheta'} \Big|_{\vartheta^\dagger} - E\left[\frac{\partial^2 l_{0t}}{\partial \vartheta \partial \vartheta'}\right] = o_p(1),$$

as  $n$  tends to  $\infty$ . This is shown in the following lemmas.



**Lemma 3.6.4** Suppose that  $\xi_t$  is generated by the HYGARCH(0,d,0) process under Assumptions 3.3.6, 3.3.7 and 3.3.9, and  $E[z_{0t}]^4 < \infty$ , and:

$$E \left| \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \right|^2 < 1.$$

Then,

$$\sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \vartheta \partial \vartheta'} - E \left[ \frac{\partial^2 l_{0t}}{\partial \vartheta \partial \vartheta'} \right] \right\| = o_p(1)$$

This lemma shows that the Hessian matrix of the  $L_n$  converges to its corresponding limit  $\Sigma_{01}$ .

**Lemma 3.6.5** If we suppose that  $\xi_t$  is generated by the HYGARCH(0,d,0) process under the Assumptions 3.3.6, 3.3.7 and 3.3.9, and  $E[z_{0t}]^4 < \infty$ , and:

$$E \left| \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \right|^2 < 1.$$

If  $\sqrt{n}(\vartheta_n - \vartheta_0) = o_p(1)$  holds, then:

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \vartheta} \frac{\partial l_t}{\partial \vartheta'} - \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right]_{\vartheta_n} = o_p(1),$$

and:

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right]_{\vartheta_n} - E \left[ \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right] = o_p(1).$$

By the Lemma 3.6.1-3.6.5, it is able to derive the following theorem:

**Theorem 3.6.1** Suppose that  $\xi_t$  is generated by the HYGARCH(0,d,0) process under Assumptions 3.3.6, 3.3.7 and 3.3.9,  $E[\xi_t]^2 < \infty$ , and  $E[z_{0t}]^4 < \infty$ , and:

$$E \left| \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \right|^2 < 1,$$

then

$$\sqrt{n}(\vartheta_n - \vartheta_0) \rightarrow^d N(0, \Sigma_{01}^{-1} \Omega_{01} \Sigma_{01}^{-1})$$

as  $n$  tends to infinity, which means that the estimator  $\vartheta_n$  is asymptotically normally distributed.

### Estimations with finite past information

Subsection 3.6.1 has shown the CAN property of  $\vartheta_n$ , however, the likelihood function depends on infinite past information, and this cannot be obtained in empirical applications. Thus, this subsection provides the asymptotic theory of a feasible log-likelihood function that depends on finite past information. However, based on the previous sections, it is hard to derive the invertibility condition for the HYEGARCH process. First, let us define the following:

$$\begin{cases} \ln \tilde{h}_t = \omega + \sum_{j=1}^{n-1} b_j g_1(\tilde{z}_{t-j}) \\ \ln h_t = \omega + \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) \end{cases}.$$

Then,

$$\begin{aligned} & E \left| \ln h_t - \ln \tilde{h}_t \right| \\ &= E \left| \omega + \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) - \omega + \sum_{j=1}^{n-1} b_j g_1(\tilde{z}_{t-j}) \right| \\ &\leq \sum_{j=1}^{n-1} b_j E |g_1(z_{t-j}) - g_1(\tilde{z}_{t-j})| + E \left| \sum_{j=n}^{\infty} b_j g_1(z_{t-j}) \right|. \end{aligned}$$

Following the investigation in Section 3.4.3, the invertibility condition of the HYEGARCH process is very likely to be satisfied. Although the proof of the invertibility conditions is still work in progress, we expect that the HYEGARCH satisfies the invertibility conditions. Here, we assume that  $\sum_{j=1}^{n-1} b_j E |g_1(z_{t-j}) - g_1(\tilde{z}_{t-j})|$

converges to 0 almost surely as  $n$  tends to infinity can be satisfied. Then

$$\begin{aligned} E \left| \ln h_t - \ln \tilde{h}_t \right| &\leq E \left| \sum_{j=n}^{\infty} b_j g_1(z_{t-j}) \right| \\ &\leq E |g_1(z_t)| \sum_{j=n}^{\infty} b_j \\ &= O(n^{-d}). \end{aligned}$$

where  $E |g_1(z_t)| < \infty$ .

**Lemma 3.6.6** *Suppose that  $\xi_t$  is generated by the HYGARCH(0,  $d$ , 0) process under Assumptions 3.3.7 and 3.3.9, and  $E[z_t]^4 < \infty$ , then  $\sup_{\lambda \in \Theta} |L_n(\vartheta) - \tilde{L}_n(\vartheta)| \xrightarrow{L_1} 0$  as  $n \rightarrow \infty$ .*

**Theorem 3.6.2** *Assuming that the Assumptions 3.3.7 and 3.3.9 hold. Then, under conditions of Lemmas 3.6.1, 3.6.2 and 3.6.6,  $\hat{\vartheta}_n$  is a consistent estimator of  $\vartheta_0$ , as  $n \rightarrow \infty$ .*

This theorem shows that the  $\hat{\vartheta}_n$  is a consistent estimator of  $\vartheta_0$ . By the Taylor series expansion, we have:

$$0 = \frac{1}{n} \sum_{t=1}^n \frac{\partial \tilde{l}_t}{\partial \vartheta} = \frac{1}{n} \sum_{t=1}^n \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} + \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \tilde{l}_t}{\partial \vartheta \partial \vartheta'} \bigg|_{\tilde{\vartheta}^\dagger} (\hat{\vartheta}_n - \vartheta_0),$$

where  $\tilde{\vartheta}^\dagger$  lies between  $\hat{\vartheta}_n$  and  $\vartheta_0$ . To establish the asymptotic distribution of the feasible estimator  $\hat{\vartheta}_n$  in the HYGARCH process, it is essential to investigate the asymptotic behaviour of:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \rightarrow^d N(0, \Omega_{01}) \quad (3.45)$$

and:

$$\sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \tilde{l}_t}{\partial \vartheta \partial \vartheta'} \bigg|_{\tilde{\vartheta}^\dagger} - E \left[ \frac{\partial^2 l_{0t}}{\partial \vartheta \partial \vartheta'} \right] \right\| = o_p(1).$$

In Subsection 3.6.1, the convergence of  $\frac{1}{\sqrt{n}} \sum_{t=1}^n \partial l_{0t} / \partial \vartheta$  and  $\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \vartheta \partial \vartheta'} \Big|_{\hat{\vartheta}^\dagger}$  has been derived. Then, to derive the asymptotic distribution of  $\hat{\vartheta}_n$ , it is essential to investigate the difference between the gradient of the  $L_n(\vartheta)$  and that of  $\tilde{L}_n(\vartheta)$ , and the difference between the Hessian matrix of  $L_n(\vartheta)$  and that of  $\tilde{L}_n(\vartheta)$ . The following lemmas can be derived under the existence of fourth moment of the first derivative of  $\ln h_t$  and under Theorem 3.6.2. The Lemma 3.6.7 ensures the convergence of the difference between the gradient.

**Lemma 3.6.7** *Suppose that  $\xi_t$  is generated by the HYGARCH(0,d,0) process under Assumptions 3.3.6, 3.3.7 and 3.3.9, and  $E[z_{0t}]^4 < \infty$ , then:*

$$E \sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right] \right\| = o_p(1). \quad (3.46)$$

**Lemma 3.6.8** *Suppose that  $\xi_t$  is generated by the HYGARCH(0,d,0) process under Assumption 3.3.6, 3.3.7 and 3.3.9, and  $E[z_{0t}]^4 < \infty$ , then:*

$$E \sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial^2 l_t}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 \tilde{l}_t}{\partial \vartheta \partial \vartheta'} \right] \right\| = o_p(1), \quad (3.47)$$

and:

$$\sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \tilde{l}_t}{\partial \vartheta \partial \vartheta'} \Big|_{\hat{\vartheta}^\dagger} - E \left[ \frac{\partial^2 l_{0t}}{\partial \vartheta \partial \vartheta'} \right] \right\| = o_p(1).$$

**Remark 1** *To establish the asymptotic distribution of  $\hat{\vartheta}_n$ , it is essential to show the score function converges to  $N(0, \Omega_{01})$  in distribution, where  $\Omega_{01} = E[(\partial l_{0t} / \partial \vartheta)(\partial l_{0t} / \partial \vartheta)']$ .*

The Lemma 3.6.3 shows that:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial l_{0t}}{\partial \vartheta} \rightarrow^d N(0, \Omega_{01}). \quad (3.48)$$

Then if  $\frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right\|$  converges to zero as  $n$  tends to infinity, the same results can be extended to:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \rightarrow^d N(0, \Omega_{01}). \quad (3.49)$$

Since:

$$\begin{aligned}
& \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right\| \\
&= \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \left( \frac{\xi_t^2}{\tilde{h}_{0t}} - 1 \right) \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\| \\
&\leq \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \left( \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right) \right\| \\
&\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \xi_t^2 \left( \frac{1}{h_{0t}} - \frac{1}{\tilde{h}_{0t}} \right) \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\|,
\end{aligned}$$

then suppose that  $E|\xi_t|^6 < \infty$ , and denotes:

$$\Xi_{12} = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \xi_t^2 \left( \frac{1}{h_{0t}} - \frac{1}{\tilde{h}_{0t}} \right) \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\|,$$

then:

$$\begin{aligned}
E[\Xi_{12}] &= \frac{1}{\sqrt{n}} \sum_{t=1}^n E \left\| \xi_t^2 \left( \frac{1}{h_{0t}} - \frac{1}{\tilde{h}_{0t}} \right) \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\| \tag{3.50} \\
&\leq C \frac{1}{\sqrt{n}} \sum_{t=1}^n \left( E|\xi_t^2|^3 \right)^{\frac{1}{3}} \left( E|\ln h_{0t} - \ln \tilde{h}_{0t}|^2 \right)^{\frac{1}{2}} \left( E \left| \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right|^6 \right)^{\frac{1}{6}} \\
&\leq C_{\Xi_{12}} \frac{1}{\sqrt{n}} \sum_{t=1}^n \left( E|\ln h_{0t} - \ln \tilde{h}_{0t}|^2 \right)^{\frac{1}{2}} \\
&= C'_{\Xi_{12}} n^{-d+\frac{1}{2}}.
\end{aligned}$$

This means that  $E[\Xi_{12}]$  cannot converge to zero. Thus,  $\frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right\|$  does not converge to zero as  $n$  goes to infinity. Therefore, the asymptotic normality of  $\hat{\vartheta}_n$  cannot be shown by that of  $\vartheta_n$  in the HYGARCH process. However, there might be another method that can be applied to derive the asymptotic normality of the  $\hat{\vartheta}_n$  in the HYGARCH process. This is worthy of further studies.

### 3.6.2 Investigation of the asymptotic theory of the QMLE in the FIEGARCH process

If we compared the FIEGARCH and HYGARCH processes, the former can capture long memory in volatility. However, the asymptotic theory of the QMLE in these processes is still an open question. For the asymptotic property of the long memory in volatility, Beran and Schützner (2009) and Schützner (2009) introduced the truncated (and feasible) QMLE of the parameter vector in the long memory case of the LARCH( $\infty$ ) process. This method was also used by Grublytė et al. (2015) to investigate the asymptotic theory of the QMLE in the generalised quadratic ARCH (GQARCH) model, which is also able to capture the long-run dependence of volatility.

The truncated (and feasible) QMLE of the parameter vector was introduced in Definition 6.3 of Schützner (2009). Let  $\epsilon > 0$  and  $0 < \iota < 1$ . Define  $m(n) = [n^\iota] - 1$ . For the sample  $\xi_1, \dots, \xi_n$ , the truncated (and feasible) estimator of the parameter vector  $\vartheta$  is defined by:

$$\vartheta_n^{(\epsilon, \iota)} = \arg \max_{\vartheta \in \Theta} \tilde{L}_{n, \epsilon}(\vartheta),$$

where the truncated objective function is given by:

$$\tilde{L}_{n, h}(\vartheta) = \frac{1}{m(n) + 1} \sum_{t=n-m(n)}^n \frac{\xi_t^2 + \epsilon}{\tilde{h}_t(\vartheta) + \epsilon} + \ln(\tilde{h}_t(\vartheta) + \epsilon),$$

where  $[\cdot]$  denotes the floor function, (i.e.  $[x]$  is the largest integer smaller than  $x$ ).

In this truncated objective function, the additional function  $m(n)$  is used to avoid  $\vartheta_n^{(\epsilon)}$  and  $\hat{\vartheta}_n^{(\epsilon)}$  having different asymptotic distributions because  $\left| h_t(\vartheta_0) - \tilde{h}_t(\vartheta_0) \right|$  and  $\left| \ln h_t(\vartheta_0) - \ln \tilde{h}_t(\vartheta_0) \right|$  converge to zero very slowly. More precisely,  $h_t(\vartheta_0)$  may obtain a poor estimation of  $\tilde{h}_t(\vartheta_0)$  if  $\tilde{h}_t(\vartheta_0)$  depends on only a small amount of the past information, so there will be a large difference between  $\tilde{h}_t(\vartheta_0)$  and  $h_t(\vartheta_0)$ .

Beran and Schützner (2009) proposed that skipping the first  $n - m(n) - 1$  samples is instrumental in solving these problems. Thus, if we use only the most reliable past information to get an approximation of  $\tilde{h}_t(\vartheta_0)$ , the value of  $m(n)$  needs to be chosen to make  $\left| \ln h_t(\vartheta_0) - \ln \tilde{h}_t(\vartheta_0) \right|$  converge to zero when  $-1/2 < d < 0$ .

This truncated (and feasible) estimator can probably also be applied to the FIEGARCH process if the following conditions can be supported. In the FIEGARCH case,  $\ln h_t$  with a lower bound, since:

$$\begin{aligned} \ln h_t &= \omega + \alpha(1 - (1 - L)^d)g_1(z_t) \\ &= \omega + \alpha \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) \\ &= \omega + \alpha \sum_{j=1}^{\infty} b_j \left( \theta \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} + \left| \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} \right| \right), \end{aligned}$$

where all  $\alpha b_j$  are positive values, and under the condition  $|\theta| \leq 1$ , then,  $\ln h_t \geq \omega$ . However, since  $g_1(z_t) = \theta z_t + |z_t|$  does not satisfy the condition of  $L_2$ -mixingale or martingale difference the sum of  $\alpha b_j$  is non-summable, and the Rosenthal's inequality cannot be applied here. This makes it hard to derive the invertibility of the FIEGARCH process. If we define the  $g_1(z_t) = \theta z_t + |z_t| - E|z_t|$ , then the  $g_1(z_t)$  satisfies the condition of the martingale difference then the Rosenthal's inequality can be used to obtain the upper bound of  $|\ln h_t|$ . However, this raises problems in deriving the uniform convergence of  $\sup_{\lambda \in \Theta} |L_n(\vartheta) - \bar{L}(\vartheta)|$ . If these problems can be solved, then the truncated (and feasible) QML estimator of the parameter might be applied to show the asymptotic normality of the FIEGARCH process. This will be investigated deeply in my further studies.

### 3.6.3 Simulations for the HY/FIEGARCH models

To illustrate the asymptotic property of the QMLE for the simulated series of HYGARCH and FIEGARCH models with standard normally distributed  $z_t$ , the true parameters are chosen as follows. In the HYGARCH process,  $\alpha = 1$  is fixed, and setting  $\omega_0 = 1$ , and  $d_0 = 0.1, 0.2, 0.3, 0.4$ ; in the FIEGARCH process,  $\alpha = -1$  is fixed, and setting  $\omega_0 = 1$ , and  $d_0 = -0.1, -0.2, -0.3, -0.4$ . The process  $\xi_t$  is generated via the HYGARCH(0,  $d$ , 0) and FIEGARCH(0,  $d$ , 0) models with 2002 pre-samples. To compare the asymptotic properties of the estimators with finite sample results, a Monte Carlo experiment is applied as follows. In this experiment, the sample sizes  $n = 501, 2001$ , with  $N = 1000$  replications. All of the simulations are using software TSM and results are shown in Tables 3.1 and 3.2. It is obvious that the bias and RMSE are decreasing as the samples increase from 501 to 2001 in both HYGARCH and FIEGARCH processes. The estimators are more consistent as  $n$  tends to infinity.

## 3.7 Conclusion

This study investigated the asymptotic theory of the QMLE in the ARMA models with EGARCH-type errors. The contributions of this study are as follows. Firstly, this chapter investigated the invertibility of the conditional variance and the derivatives of the conditional variance of the QMLE in the ARMA(1,1)-EGARCH(1,1) model by applying the method of Martinet and McAleer (2015). The invertibility of the conditional variance follows the EGARCH(1,1) process has also been investigated by Straumann (2005), Straumann and Mikosch (2006), Wintenberger (2013) and Kyriakopoulou (2015). Straumann (2005), and Straumann and Mikosch



$n$	2001							
	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$
TRUE	1	0.1	1	0.2	1	0.3	1	0.4
BIAS	-0.002	-0.004	-0.004	0.003	0	-0.001	-0.003	-0.002
RMSE	0.039	0.048	0.043	0.052	0.044	0.053	0.044	0.052
$n$	501							
	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$
TRUE	1	0.1	1	0.2	1	0.3	1	0.4
BIAS	-0.006	-0.011	-0.006	0.004	0.005	-0.007	-0.007	-0.004
RMSE	0.074	0.100	0.087	0.109	0.089	0.109	0.086	0.101

Table 3.1: Simulation results for the QMLE in the HYGARCH( $0, d, 0$ ) process. Note: this table reports the simulation results for the QMLE in the HYGARCH( $0, d, 0$ ) processes, with  $d = 0.1, 0.2, 0.3, 0.4$ . These simulations were run with 2002 pre-samples and samples of length  $n = \{501, 2001\}$  with 1000 Monte Carlo replications. Both the bias and RMSE of these estimators, averages of over 1000 replications, are presented in this table. It is can be seen from this table that the bias and RMSE decreased with sample increasing for each parameter. Large samples provide more precise results than small samples.

$n$	2001							
	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$
TRUE	1	-0.1	1	-0.2	1	-0.3	1	-0.4
BIAS	-0.007	0.009	-0.023	0.010	-0.050	0.009	-0.101	0.007
RMSE	0.048	0.039	0.091	0.036	0.190	0.035	0.460	0.031
$n$	501							
	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$	$\omega$	$d$
TRUE	1	-0.1	1	-0.2	1	-0.3	1	-0.4
BIAS	-0.018	0.031	-0.049	0.030	-0.105	0.034	0.039	0.029
RMSE	0.099	0.088	0.178	0.080	0.628	0.080	1.530	0.074

Table 3.2: Simulation results for the QMLE in the FIEGARCH( $0, d, 0$ ) process. Note: In this table, the simulation procedure is similar as that in the Table 3.1, except the DGP considers the FIEGARCH( $0, d, 0$ ) processes, with  $d = -0.1, -0.2, -0.3, -0.4$ . This table also shows that large samples have better performance for the QMLE in the relevant FIEGARCH processes than small samples.

(2006) provided a invertibility condition for the EGARCH(1,1) models by using the SRE approach. However, Wintenberger (2013) argued that the invertibility condition of Straumann and Mikosch (2006) cannot satisfy the conditions for the asymptotic theory of the EGARCH(1,1) models in some cases. He extended Straumann's (2005) results, proposed the notion of continuous invertibility and derived the invertibility of the first and second derivative of the conditional variance. Kyriakopoulou (2015) summarised the invertibility conditions of Straumann and Mikosch (2006) and also considered the continuous invertibility from Wintenberger (2013). However, the studies of Straumann and Mikosch (2006) and Wintenberger (2013) only considered the invertibility for the EGARCH(1,1) model, and it is hard to obtain explicit representations of the (continuous) invertibility conditions, except for the EGARCH(0,1) case. Martinet and McAleer (2015) provided an explicit condition for the invertibility of the conditional variance processes in the some of the EGARCH( $p, q$ ) and some of the EARCH( $\infty$ ) models. This chapter extended the study of Martinet and McAleer (2015) to investigate the invertibility of the derivatives of the conditional variance of the QMLE in the ARMA(1,1) model with an EGARCH(1,1) error. This supports the establishment of the asymptotic normality of the QMLE in the relevant models.

Secondly, this study extended the previous work and finally established the asymptotic theory for the QMLE in the ARMA(1,1)-EGARCH(1,1) model. For empirical studies, it might be too restrictive to assume that the return processes follow pure (exponential) conditional heteroscedasticity models, and more and more researchers consider the return process to be the ARMA model with (exponential) conditional heteroscedasticity errors. However, previous studies mainly considered the asymptotic properties of the QMLE in the pure EGARCH(1,1) model, see Straumann and Mikosch (2006), Wintenberger (2013) and Kyriakopoulou (2015).

In this chapter, to establish the CAN of the QMLE in the ARMA-EGARCH model, I applied the modified asymptotic theory of Amemiya (1985), which was used by Ling and McAleer (2003) to establish the asymptotic theory for the vector ARMA-GARCH process. Under the invertibility condition of the conditional variance in the ARMA(1,1)-EGARCH(1,1) process, and other moment conditions, the consistency of the QMLE in the ARMA-EGARCH process was established. For the asymptotic normality of the QMLE, this study also applied the modified uniform convergence theorem of Ling and McAleer (2003), which is helpful for obtaining a weaker moment condition. Kyriakopoulou (2015) also provided a similar idea by using the ergodic theorem for continuous valued sequences of random function to establish the uniform convergence of the derivatives of the log-likelihood function. Meanwhile, to ensure the stationarity of the derivative of the conditional variance processes, this study found that a moment condition, which is same as the (MM) condition from Wintenberger (2013), also needs to be satisfied. In addition, this chapter also found that the fourth moment of the return processes and the invertibility condition for conditional variance processes are needed to satisfy the asymptotic normality of the QMLE in the ARMA-EGARCH process. Under certain conditions, the asymptotic normality of the QMLE in the ARMA(1,1)-EGARCH(1,1) was established.

Thirdly, this study also established the consistency of the QMLE in the HYE-GARCH process under the assumption of invertibility and provided an investigation into the asymptotic normality of the QMLE in the HYEGARCH model and the asymptotic normality of the QMLE in the FIEGARCH models. The simulation results showed that the QMLE is a consistency estimator for the pure HY/FIEGARHCH models.

Motivated by this study and previous literature, there are several topics are worth further study. Firstly, the method of Martinet and McAleer (2015) can probably also be applied to derive the invertibility of the HYGARCH and FIEGARCH models. This study also tried to apply the method of Martinet and McAleer (2015) to derive the invertibility condition for the HYGARCH process. Although it is hard to derive the explicit invertibility conditions because the lag coefficient of the HYGARCH process has a hyperbolic decay rate, we can see that this method is probably also suitable for obtaining the invertibility condition for the HYGARCH processes.

Moreover, Kyriakopoulou (2015) provided explicit expressions for the moment conditions and bounds for the EGARCH(1,1) models. The method of Kyriakopoulou (2015) can probably also be extended to derive the accurate moment conditions and the explicit expression of the asymptotic covariance matrix for the ARMA(1,1) model with EGARCH(1,1) errors. In addition, it would also be interesting to establish the asymptotic properties of the QMLE in the ARMA( $r, s$ ) models with EGARCH( $p, q$ ) errors following a similar procedure to that in current study.

Furthermore, based on the investigation into the asymptotic theory of the QMLE in the HYGARCH and FIEGARCH models, the method of Beran and Schützner (2009) might be useful for establishing the asymptotic theory for the FIEGARCH model.

## 3.8 Appendix B

### 3.8.1 Proof of Lemma 3.4.1

**Proof of Lemma 3.4.1.** This proof follows the proof of the Lemma 2.2 in Martinet and McAleer (2015). For the HYGARCH process, by inequality (3.40), it can be obtained that:

$$a_1(t) = \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| + \sum_{i=1}^{n-1} b_i \delta_{t-i} |\exp(\Delta_{t-i})| \left| V_{n-i}^{(n)} - U_{n-i}^{(n)} \right|.$$

Also by inequality (3.40), the upper bound of  $\left| V_{n-i}^{(n)} - U_{n-i}^{(n)} \right|$  is:

$$\begin{aligned} & \left| V_{n-i}^{(n)} - U_{n-i}^{(n)} \right| \tag{3.51} \\ & \leq \sum_{j=0}^{\infty} b_{j+n-i} \delta_{t-n-j} |z_{t-n-j}| + \sum_{j_1=1}^{n-i-1} b_{j_1} \delta_{t-i-j_1} |\exp(\Delta_{t-i-j_1})| \left| V_{n-i-j_1}^{(n)} - U_{n-i-j_1}^{(n)} \right|. \end{aligned}$$

It follows on substituting the inequality (3.51) in (3.40) that:

$$\begin{aligned} & \left| V_n^{(n)} - U_n^{(n)} \right| \tag{3.52} \\ & \leq \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| + \sum_{i=1}^{n-1} b_i \delta_{t-i} |\exp(\Delta_{t-i})| \left| \sum_{j=0}^{\infty} b_{j+n-i} \delta_{t-n-j} |z_{t-n-j}| \right| \\ & \quad + \sum_{i=1}^{n-1} b_i \delta_{t-i} |\exp(\Delta_{t-i})| \left| \sum_{j_1=1}^{n-i-1} b_{j_1} \delta_{t-i-j_1} |\exp(\Delta_{t-i-j_1})| \left| V_{n-i-j_1}^{(n)} - U_{n-i-j_1}^{(n)} \right| \right| \\ & = \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| + \sum_{i=1}^{n-1} \sum_{j=0}^{\infty} b_i b_{j+n-i} \delta_{t-i} |\exp(\Delta_{t-i})| \delta_{t-n-j} |z_{t-n-j}| \\ & \quad + \sum_{i=1}^{n-1} \sum_{j_1=1}^{n-i-1} b_i b_{j_1} \delta_{t-i} \delta_{t-i-j_1} |\exp(\Delta_{t-i} + \Delta_{t-i-j_1})| \left| V_{n-i-j_1}^{(n)} - U_{n-i-j_1}^{(n)} \right| \\ & = a_2(t), \end{aligned}$$

where by inequality (3.51), the upper bound of  $\left|V_{n-i-j_1}^{(n)} - U_{n-i-j_1}^{(n)}\right|$  is:

$$\begin{aligned}
& \left|V_{n-i-j_1}^{(n)} - U_{n-i-j_1}^{(n)}\right| \tag{3.53} \\
& \leq \sum_{j=0}^{\infty} b_{j+n-i-j_1} \delta_{t-n-j} |z_{t-n-j}| + \sum_{j_2=1}^{n-i-j_1-1} b_{j_2} \delta_{t-i-j_1-j_2} |\exp(\Delta_{t-i-j_1-j_2})| \\
& \quad \left|V_{n-i-j_1-j_2}^{(n)} - U_{n-i-j_1-j_2}^{(n)}\right|.
\end{aligned}$$

Then, substituting the inequality (3.53) in (3.52), we can get that:

$$\begin{aligned}
& |V_n^{(n)} - U_n^{(n)}| \\
& \leq \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| + \sum_{i=1}^{n-1} \sum_{j=0}^{\infty} b_i b_{j+n-i} \delta_{t-i} |\exp(\Delta_{t-i})| \delta_{t-n-j} |z_{t-n-j}| + \\
& \quad \sum_{i=1}^{n-1} \sum_{j_1=1}^{n-i-1} b_i b_{j_1} \delta_{t-i} \delta_{t-i-j_1} |\exp(\Delta_{t-i} + \Delta_{t-i-j_1})| \sum_{j=0}^{\infty} b_{j+n-i-j_1} \delta_{t-n-j} |z_{t-n-j}| \\
& \quad + \sum_{i=1}^{n-1} \sum_{j_1=1}^{n-i-1} b_i b_{j_1} \delta_{t-i} \delta_{t-i-j_1} |\exp(\Delta_{t-i} + \Delta_{t-i-j_1})| \\
& \quad \sum_{j_2=1}^{n-i-j_1-1} b_{j_2} \delta_{t-i-j_1-j_2} |\exp(\Delta_{t-i-j_1-j_2})| \left|V_{n-i-j_1-j_2}^{(n)} - U_{n-i-j_1-j_2}^{(n)}\right| \\
& = \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| + \sum_{i=1}^{n-1} \sum_{j=0}^{\infty} b_i b_{j+n-i} \delta_{t-i} |\exp(\Delta_{t-i})| \delta_{t-n-j} |z_{t-n-j}| \\
& \quad + \sum_{i=1}^{n-1} \sum_{j_1=1}^{n-i-1} \sum_{j=0}^{\infty} b_i b_{j_1} b_{j+n-i-j_1} \delta_{t-i} \delta_{t-i-j_1} |\exp(\Delta_{t-i} + \Delta_{t-i-j_1})| \delta_{t-n-j} |z_{t-n-j}| \\
& \quad + \sum_{i=1}^{n-1} \sum_{j_1=1}^{n-i-1} \sum_{j_2=1}^{n-i-j_1-1} b_i b_{j_1} b_{j_2} \delta_{t-i} \delta_{t-i-j_1} \delta_{t-i-j_1-j_2} \\
& \quad |\exp(\Delta_{t-i} + \Delta_{t-i-j_1} + \Delta_{t-i-j_1-j_2})| \left|V_{n-i-j_1-j_2}^{(n)} - U_{n-i-j_1-j_2}^{(n)}\right| \\
& = a_3(t).
\end{aligned}$$

Thus, the general  $a_k(t)$  can be obtained as:

$$\begin{aligned}
& a_k(t) \tag{3.54} \\
&= \sum_{j=0}^{\infty} b_{j+n} \delta_{t-n-j} |z_{t-n-j}| + \sum_{j=0}^{\infty} \delta_{t-n-j} |z_{t-n-j}| \\
&\quad \left( \sum_{p=1}^{k-1} \sum_{j_1, \dots, j_p \in A_p^{(n)}} \prod_{i=1}^p b_{j_i} \prod_{i=1}^p \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^p \Delta_{t-\hat{S}_i} \right) b_{j+n-\hat{S}_p} \right) \\
&\quad + \sum_{j_1, \dots, j_k \in A_p^{(n)}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \left| V_{n-\hat{S}_k}^{(n)} - U_{n-\hat{S}_k}^{(n)} \right|.
\end{aligned}$$

Following similar idea to the inequality (3.53), then  $\left| V_{n-\hat{S}_k}^{(n)} - U_{n-\hat{S}_k}^{(n)} \right|$  is:

$$\begin{aligned}
& \left| V_{n-\hat{S}_k}^{(n)} - U_{n-\hat{S}_k}^{(n)} \right| \tag{3.55} \\
&\leq \sum_{j=0}^{\infty} b_{j+n-\hat{S}_k} \delta_{t-n-j} |z_{t-n-j}| \\
&\quad + \sum_{j_{k+1}=1}^{n-\hat{S}_k-1} b_{j_{k+1}} \cdot \delta_{t-\hat{S}_k-j_{k+1}} \left| \exp \left( \Delta_{t-\hat{S}_k-j_{k+1}} \right) \left| V_{n-\hat{S}_k-j_{k+1}}^{(n)} - U_{n-\hat{S}_k-j_{k+1}}^{(n)} \right| \right|.
\end{aligned}$$

Then, substituting the inequality (3.55) in (3.54), we can get that the third term of (3.54) as:

$$\begin{aligned}
& \sum_{j_1, \dots, j_k \in A_p^{(n)}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \left| V_{n-\hat{S}_k}^{(n)} - U_{n-\hat{S}_k}^{(n)} \right| \\
&\leq \sum_{\substack{j_1, \dots, j_k \in A_p^{(n)} \\ \hat{S}_k = n-1}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \left| V_1^{(n)} - U_1^{(n)} \right| \\
&\quad + \sum_{\substack{j_1, \dots, j_k \in A_p^{(n)} \\ \hat{S}_k < n-1}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \\
&\quad \left( \sum_{j_{k+1}=1}^{n-\hat{S}_k-1} b_{j_{k+1}} \cdot \delta_{t-\hat{S}_k-j_{k+1}} \left| \exp \left( \Delta_{t-\hat{S}_k-j_{k+1}} \right) \left| V_{n-\hat{S}_k-j_{k+1}}^{(n)} - U_{n-\hat{S}_k-j_{k+1}}^{(n)} \right| \right| \right) \\
&\quad + \sum_{\substack{j_1, \dots, j_k \in A_p^{(n)} \\ \hat{S}_k < n-1}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \left( \sum_{j=0}^{\infty} b_{j+n-\hat{S}_k} \delta_{t-n-j} |z_{t-n-j}| \right)
\end{aligned}$$

since:

$$\left| V_1^{(n)} - U_1^{(n)} \right| \leq \sum_{j=0}^{\infty} b_{j+1} \delta_{t-n-j} |z_{t-n-j}|,$$

then:

$$\begin{aligned} & \sum_{j_1, \dots, j_k \in A_p^{(n)}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \left| V_{n-\hat{S}_k}^{(n)} - U_{n-\hat{S}_k}^{(n)} \right| \quad (3.56) \\ & \leq \sum_{\substack{j_1, \dots, j_k \in A_p^{(n)} \\ \hat{S}_k < n-1}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \\ & \quad \left( \sum_{j_{k+1}=1}^{n-\hat{S}_k-1} b_{j_{k+1}} \delta_{t-\hat{S}_k-j_{k+1}} \left| \exp \left( \Delta_{t-\hat{S}_k-j_{k+1}} \right) \right| \left| V_{n-\hat{S}_k-j_{k+1}}^{(n)} - U_{n-\hat{S}_k-j_{k+1}}^{(n)} \right| \right) \\ & \quad + \sum_{j_1, \dots, j_k \in A_p^{(n)}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \left( \sum_{j=0}^{\infty} b_{j+n-\hat{S}_k} \delta_{t-n-j} |z_{t-n-j}| \right) \\ & \leq \sum_{\substack{j_1, \dots, j_k \in A_p^{(n)} \\ \hat{S}_k < n-1}} \sum_{j_{k+1}=1}^{n-\hat{S}_k-1} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \cdot b_{j_{k+1}} \delta_{t-\hat{S}_k-j_{k+1}} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} + \Delta_{t-\hat{S}_k-j_{k+1}} \right) \\ & \quad \left| V_{n-\hat{S}_k-j_{k+1}}^{(n)} - U_{n-\hat{S}_k-j_{k+1}}^{(n)} \right| \\ & \quad + \sum_{j_1, \dots, j_k \in A_p^{(n)}} \prod_{i=1}^k b_{j_i} \prod_{i=1}^k \delta_{t-\hat{S}_i} \exp \left( \sum_{i=1}^k \Delta_{t-\hat{S}_i} \right) \left( \sum_{j=0}^{\infty} b_{j+n-\hat{S}_k} \delta_{t-n-j} |z_{t-n-j}| \right) \end{aligned}$$

Then, substituting (3.56) in  $a_k(t)$ , and since that

$$\left\{ j_1, \dots, j_k \in A_p^{(n)}, j_{k+1} \in [1, n - \hat{S}_k - 1] : \hat{S}_k < n - 1 \right\} = A_{k+1}^{(n)}$$

Then, it is straightforward that  $a_k(t) \leq a_{k+1}(t)$ , and this means that:

$$\left| V_n^{(n)} - U_n^{(n)} \right| \leq a_k \leq a_{k+1}.$$

This completes the proof. ■



### 3.8.2 Proof of Lemma 3.5.1

**Proof of Lemma 3.5.1.** The expectation value of the absolute value of the unobserved log-likelihood function is:

$$E|l_t(\lambda)| = E|\ln h_t| + E\left[\frac{\xi_t^2}{h_t}\right].$$

Based on the ARMA(1,1) process:

$$\xi_t = \frac{(1 - \phi_1 L)(y_t - \mu)}{1 + \psi_1 L}. \quad (3.57)$$

Then:

$$|\xi_t| \leq C_{0\xi} + C \sum_{i=0}^{\infty} \rho^{-i} |y_{t-i}|, \quad (3.58)$$

where  $C_{0\xi} = |(1 - \phi_1)\mu/(1 + \psi_1)|$ ,  $C > 0$  are constants, and  $\rho > 1$ , independent to all  $\lambda \in \Theta$ . Suppose  $E(y_t)^2 < \infty$ , then  $E \sup_{\lambda \in \Theta} (\xi_t)^2 < \infty$ . By the Assumption 3.3.5,  $h_t$  has a lower bound uniformly over  $\Theta$ . Then  $E \sup_{\lambda \in \Theta} [\xi_t^2/h_t] < \infty$ . Since the  $h_t$  follows the EGARCH(1,1) process, and by the Assumption 3.3.3  $|\beta| < 1$ , the EGARCH(1,1) process can be reorganised as:

$$\begin{aligned} |\ln h_t| &= \left| \frac{w}{1 - \beta} + \sum_{j=0}^{\infty} \beta^j \left( \theta \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} + \gamma \left| \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} \right| \right) \right| \\ &\leq C + \left| \sum_{j=0}^{\infty} \beta^j \left( \theta \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} + \gamma \left| \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} \right| \right) \right| \\ &\leq C + \sum_{j=0}^{\infty} \left| \beta^j \left( \frac{\theta \xi_{t-j} + \gamma |\xi_{t-j}|}{\sqrt{h_{t-j}}} \right) \right| \\ &= C + C_h \sum_{j=0}^{\infty} |\beta^j (\theta \xi_{t-j} + \gamma |\xi_{t-j}|)| \\ &\leq C + C_{h1} \sum_{j=0}^{\infty} |\rho_1^j| |y_{t-j}|, \end{aligned}$$

where  $C, C_h, C_{h1} > 0$  are constants and  $|\rho_1| < 1$  independent of all  $\lambda \in \Theta$ , then under condition  $E(y_t)^2 < \infty$ ,  $|\ln h_t|$  can be bounded such that,  $E \sup_{\lambda \in \Theta} |\ln h_t| <$

$\infty$ . Thus,  $E|l_t(\lambda)| < \infty$  for all  $\lambda \in \Theta$ . And since  $l_t$  is defined based on all the past information, then  $E \sup_{\lambda \in \Theta} |l_t - El_t| < \infty$ , and  $E[l_t - El_t] = 0$ . Thus, set  $f(\underline{X}_t, \lambda) = l_t - El_t$ , applying the uniform convergence theorem 3.5.1, then  $\sup_{\lambda \in \Theta} |\frac{1}{n} \sum_{t=1}^n l_t - El_t| = o_p(1)$ . Therefore,  $\sup_{\lambda \in \Theta} |L_n(\lambda) - \bar{L}(\lambda)| = o_p(1)$  holds. This completes the proof. ■

### 3.8.3 Proof of Lemma 3.5.2

**Proof of Lemma 3.5.2.** The expectation value of the  $l_t(\lambda)$  function is:

$$\begin{aligned}
E[l_t(\lambda)] &= -E[\ln h_t] - E\left[\frac{\xi_t^2}{h_t}\right] \\
&= -E[\ln h_t] - E\left[\frac{(\xi_t - \xi_{0t} + \xi_{0t})^2}{h_t}\right] \\
&= \left\{-E[\ln h_t] - E\left[\frac{\xi_{0t}^2}{h_t}\right]\right\} - E\left[\frac{(\xi_t - \xi_{0t})^2}{h_t}\right] \\
&= l_1(\lambda) + l_2(\lambda).
\end{aligned} \tag{3.59}$$

To maximise the second term ( $l_2(\lambda)$ ):

$$\max l_2(\lambda) = -E\left[\frac{(\xi_t - \xi_{0t})^2}{h_t}\right] = 0,$$

if and only if  $\xi_t = \xi_{0t}$ , because  $h_t > 0$ , this means:

$$\xi_t - \xi_{0t} = 0.$$

Then, by the mean value expansion:

$$\xi_t - \xi_{0t} = \left.\frac{\partial \xi_t}{\partial \varphi'}\right|_{\varphi_0} (\varphi - \varphi_0) = 0,$$

where  $\varphi = (\mu, \phi_1, \psi_1)'$  and  $\varphi^\dagger$  lies between  $\varphi$  and  $\varphi_0$ , then:

$$\begin{aligned} E(\xi_t - \xi_{0t})^2 &= E \left[ (\varphi - \varphi_0)' \frac{\partial \xi_t}{\partial \varphi} \frac{\partial \xi_t}{\partial \varphi'} \Big|_{\varphi^\dagger} (\varphi - \varphi_0) \right] \\ &= (\varphi - \varphi_0)' E \left[ \frac{\partial \xi_t}{\partial \varphi} \frac{\partial \xi_t}{\partial \varphi'} \Big|_{\varphi^\dagger} \right] (\varphi - \varphi_0) \\ &= 0. \end{aligned} \quad (3.60)$$

The first derivative of  $\xi_t$  with respect to parameter vector  $\varphi$ , in the ARMA(1,1), are given by:

$$\frac{\partial \xi_t}{\partial \varphi} = \begin{pmatrix} (\phi_1 - 1) \\ -(y_{t-1} - \mu) \\ -\xi_{t-1} \end{pmatrix} - \psi_1 \frac{\partial \xi_{t-1}}{\partial \varphi}. \quad (3.61)$$

Multiplying  $(\varphi - \varphi_0)'$  on both sides of Equation (3.61), and let  $(\varphi - \varphi_0)' (\partial \xi_t / \partial \varphi) = 0$ , then by the stationarity of  $(\partial \xi_t / \partial \varphi)$ :

$$(\varphi - \varphi_0)' \begin{pmatrix} (\phi_1 - 1) \\ -(y_{t-1} - \mu) \\ -\xi_{t-1} \end{pmatrix} = 0. \quad (3.62)$$

According to assumption 3.3.2, the Equation (3.62) holds if and only if  $(\varphi - \varphi_0)' = 0$ . Thus:

$$E \left[ \frac{\partial \xi_t}{\partial \varphi} \frac{\partial \xi_t}{\partial \varphi'} \Big|_{\varphi^\dagger} \right] > 0,$$

and:

$$E(\xi_t - \xi_{0t})^2 = (\varphi - \varphi_0)' E \left[ \frac{\partial \xi_t}{\partial \varphi} \frac{\partial \xi_t}{\partial \varphi'} \Big|_{\varphi^\dagger} \right] (\varphi - \varphi_0) = 0$$

holds if and only if  $\varphi - \varphi_0 = 0$ , that is  $\varphi = \varphi_0$ . To maximise the first term, that is  $\max l_1(\lambda)$ :

$$\begin{aligned} l_1(\lambda) &= -E \ln |h_t| - E \left[ \frac{\xi_{0t}^2}{h_t} \right] \\ &= -E \ln \left| h_{0t} \frac{h_t}{h_{0t}} \right| - E \left[ \frac{h_{0t}}{h_t} \right] \\ &= -E \ln(h_{0t}) - \left[ -E \ln \left| \frac{h_{0t}}{h_t} \right| + E \left( \frac{h_{0t}}{h_t} \right) \right]. \end{aligned} \quad (3.63)$$

Then,  $\max l_1(\lambda) = -1 - E \ln(h_{0t})$ , and this can be obtained if and only if  $h_t = h_{0t}$ , this means  $\ln h_t = \ln h_{0t}$ . Thus:

$$\begin{aligned} E[l_t(\lambda)] &\leq \max_{\lambda \in \Theta} l_1(\lambda) + \max_{\lambda \in \Theta} l_2(\lambda) \\ &\leq -1 - E \ln(h_{0t}), \end{aligned} \quad (3.64)$$

and it can be held if and only if:

$$\varphi = \varphi_0,$$

and:

$$\ln h_t = \ln h_{0t}.$$

Then:

$$(\ln h_t - \ln h_{0t})|_{\varphi=\varphi_0} = \frac{\partial \ln h_t}{\partial \zeta'} \Big|_{(\varphi_0, \zeta^\dagger)} (\zeta - \zeta_0) = 0, \quad (3.65)$$

with probability one, where  $\zeta' = (w, \beta, \theta, \gamma)$  and  $\zeta^\dagger$  lies between  $\zeta$  and  $\zeta_0$ . The first order derivative of  $\ln h_t$  with respect to parameter vector  $\zeta$  which is from the EGARCH(1,1) part are:

$$\frac{\partial \ln h_t}{\partial \zeta} = \begin{pmatrix} 1 \\ \ln h_{t-1} \\ z_{t-1} \\ |z_{t-1}| \end{pmatrix} + \left( \beta - \frac{1}{2}\theta z_{t-1} - \frac{1}{2}\gamma |z_{t-1}| \right) \frac{\partial \ln h_{t-1}}{\partial \zeta}, \quad (3.66)$$

where  $\zeta' = (w, \beta, \theta, \gamma)$ . Then, multiplying  $(\zeta - \zeta_0)'$  on both sides of the Equation (3.66), then:

$$\begin{aligned} &(\zeta - \zeta_0)' \frac{\partial \ln h_t}{\partial \zeta} \\ &= (\zeta - \zeta_0)' \begin{pmatrix} 1 \\ \ln h_{t-1} \\ z_{t-1} \\ |z_{t-1}| \end{pmatrix} + \left( \beta - \frac{1}{2}\theta z_{t-1} - \frac{1}{2}\gamma |z_{t-1}| \right) (\zeta - \zeta_0)' \frac{\partial \ln h_{t-1}}{\partial \zeta}, \end{aligned} \quad (3.67)$$

because the stationarity of  $\partial \ln h_t / \partial \varsigma$ , if  $(\varsigma - \varsigma_0)' \partial \ln h_t / \partial \varsigma = 0$ , then:

$$(\varsigma - \varsigma_0)' \begin{pmatrix} 1 \\ \ln h_{t-1} \\ z_{t-1} \\ |z_{t-1}| \end{pmatrix} = 0. \quad (3.68)$$

First consider the  $\beta - \beta_0$ , if  $\beta - \beta_0 \neq 0$ , this implies that  $\ln h_{t-1}$  is linearly dependent with  $z_{t-1}$ . However,  $\ln h_{t-1}$  and  $z_{t-1}$  are independent. Thus  $\beta = \beta_0$ . Then for the rest of the Equation (3.68), denote:

$$f(z_{t-1}) = (w - w_0) + (\theta - \theta_0)z_{t-1} + (\gamma - \gamma_0)(|z_{t-1}|) = 0.$$

For general choice of these parameters,  $f(z_{t-1}) = 0$  almost surely if and only if  $((w - w_0), (\theta - \theta_0), (\gamma - \gamma_0)) = (0, 0, 0)$  holds. That means  $\partial \ln h_t / \partial \varsigma \neq 0$ ,

$$(\varsigma - \varsigma_0)' \frac{\partial \ln h_t}{\partial \varsigma} \Big|_{(\varphi_0, \varsigma^\dagger)} = 0 \quad (3.69)$$

if and only if  $\varsigma = \varsigma_0$ . Therefore,  $\bar{L}(\lambda)$  is uniquely maximised at  $\lambda_0$ . This completes the proof. ■

### 3.8.4 Proof of Lemma 3.5.3

**Proof of Lemma 3.5.3.** The difference between the unobserved likelihood function and the observed likelihood function is as follows:

$$L_n(\lambda) - \tilde{L}_n(\lambda) = \frac{1}{n} \sum_{t=1}^n \left( \ln \frac{\tilde{h}_t(\lambda)}{h_t(\lambda)} \right) + \frac{1}{n} \sum_{t=1}^n \left( \frac{\tilde{\xi}_t^2(\lambda)}{\tilde{h}_t(\lambda)} - \frac{\xi_t^2(\lambda)}{h_t(\lambda)} \right).$$

Then:

$$|l_t(\lambda) - \tilde{l}_t(\lambda)| \leq \left| \ln h_t(\lambda) - \ln \tilde{h}_t(\lambda) \right| + \left| \frac{\xi_t^2(\lambda)}{h_t(\lambda)} - \frac{\tilde{\xi}_t^2(\lambda)}{\tilde{h}_t(\lambda)} \right|. \quad (3.70)$$

By assumption 3.3.2 and the ARMA(1,1) process, the  $\xi_t$  and  $\tilde{\xi}_t$  can be written as:

$$\xi_t = \sum_{i=0}^{\infty} \gamma_i (y_{t-i} - \mu), \quad \tilde{\xi}_t = \sum_{i=0}^{t-1} \gamma_i (y_{t-i} - \mu), \quad (3.71)$$

where  $(1 + \psi_1 L)^{-1}(1 - \phi_1 L) = \sum_{i=0}^{\infty} \gamma_i L^i$ , then:

$$|\xi_t - \tilde{\xi}_t| \leq c \sum_{i=t}^{\infty} \rho_2^{-i} |y_{t-i} - \mu|, \quad (3.72)$$

where  $\rho_2 > 1$  and  $c$  and  $\rho_2$  are constants independent of the parameter  $\lambda$ . Thus:

$$E \sup_{\lambda \in \Theta} (\xi_t - \tilde{\xi}_t) = O(\rho_2^{-t}). \quad (3.73)$$

This means that  $\tilde{\xi}_t$  converges to  $\xi_t$  as  $t$  tends to infinity. For the second term of inequality (3.70):

$$\begin{aligned} \left| \frac{\xi_t^2}{h_t} - \frac{\tilde{\xi}_t^2}{\tilde{h}_t} \right| &\leq \left| \frac{\xi_t^2}{h_t} - \frac{\xi_t^2}{\tilde{h}_t} \right| + \left| \frac{\xi_t^2}{\tilde{h}_t} - \frac{\tilde{\xi}_t^2}{\tilde{h}_t} \right| \\ &= \xi_t^2 \left| \frac{1}{h_t} - \frac{1}{\tilde{h}_t} \right| + \frac{1}{\tilde{h}_t} \left| \xi_t^2 - \tilde{\xi}_t^2 \right|, \end{aligned}$$

where:

$$\begin{aligned} &\frac{1}{\tilde{h}_t} \left| \xi_t^2 - \tilde{\xi}_t^2 \right| \\ &= \frac{1}{\tilde{h}_t} \left| \xi_t^2 + \tilde{\xi}_t^2 - 2\xi_t \tilde{\xi}_t + \xi_t \tilde{\xi}_t + \xi_t \tilde{\xi}_t - \xi_t^2 - \tilde{\xi}_t^2 \right| \\ &\leq \frac{2|\xi_t|}{\tilde{h}_t} \left| \xi_t - \tilde{\xi}_t \right| + \frac{1}{\tilde{h}_t} (\xi_t - \tilde{\xi}_t)^2, \end{aligned}$$

and then by section 5.4.1 in Straumann (2005),  $\ln h_t \geq m$  and  $\ln \tilde{h}_t \geq m - \epsilon$  for all  $t$  where  $m = \inf_{\lambda \in \Theta} w/(1 - \beta)$ , and by the mean value theorem, this yields a constant  $c > 0$  with:

$$\begin{aligned} \sup_{\lambda \in \Theta} \xi_t^2 \left| \frac{1}{h_t} - \frac{1}{\tilde{h}_t} \right| &= \xi_t^2 \left| \exp(-\ln h_t) - \exp(-\ln \tilde{h}_t) \right| \\ &\leq c \xi_t^2 |\ln h_t - \ln \tilde{h}_t|. \end{aligned}$$

Since  $\xi_t$  converges to  $\tilde{\xi}_t$  as  $t$  tends to infinity in the conditional mean equation, then for the conditional variance equation, similar approach to Martinet and McAleer (2015) is used to show the invertibility of EGARCH(1,1) process, see section 3.4, under the conditions:

$$E \left[ \frac{\delta_t |z_t|}{1 - \beta} \right] + \ln (\beta + E [\delta_t |z_t|]) < 0,$$

where  $\delta_t = (\gamma/2) + (\theta/2) \text{sign}(z_t)$  and  $\gamma \geq \theta$ , then:

$$|\ln h_t - \ln \tilde{h}_t| = o_p(1).$$

Thus:

$$\left| \frac{\xi_t^2}{h_t} - \frac{\tilde{\xi}_t^2}{\tilde{h}_t} \right| \leq c \xi_t^2 |\ln h_t - \ln \tilde{h}_t| + \frac{2|\xi_t|}{\tilde{h}_t} |\xi_t - \tilde{\xi}_t| + \frac{1}{\tilde{h}_t} (\xi_t - \tilde{\xi}_t)^2 = o_p(1)$$

and  $\frac{1}{n} \sum_{t=1}^n \xi_t^2 \sup_{\lambda \in \Theta} |\ln h_t - \ln \tilde{h}_t| = o_p(1)$ . Therefore:

$$\sup_{\lambda \in \Theta} |L_n(\lambda) - \tilde{L}_n(\lambda)| = o_p(1). \quad (3.74)$$

This means that the difference between the unobserved likelihood process and the observed likelihood process asymptotically converges to zero. This completes the proof. ■

### 3.8.5 Proof of Theorem 3.5.2

**Proof of Theorem 3.5.2.** By Assumption 3.3.1, the space  $\Theta$  is compact and  $\lambda_0$  is an interior point in  $\Theta$ . Then, by Lemmas 3.5.1 and 3.5.3,  $\tilde{L}_n(\lambda)$  converges to

$\bar{L}(\lambda)$  uniformly in  $\Theta$ , that is:

$$\begin{aligned}
& \sup_{\lambda \in \Theta} |\tilde{L}_n(\lambda) - \bar{L}(\lambda)| & (3.75) \\
&= \sup_{\lambda \in \Theta} |\tilde{L}_n(\lambda) + L_n(\lambda) - L_n(\lambda) - \bar{L}(\lambda)| \\
&\leq \sup_{\lambda \in \Theta} |L_n(\lambda) - \bar{L}(\lambda)| + \sup_{\lambda \in \Theta} |L_n(\lambda) - \tilde{L}_n(\lambda)| \\
&= o_p(1).
\end{aligned}$$

Then, by Lemma 3.5.2, that shows that  $\bar{L}(\lambda)$  has a unique maximum at  $\lambda_0$ . Therefore, the estimator  $\hat{\lambda}_n$  is a consistent estimator of the true parameters. This completes the proof. ■

### 3.8.6 Proof of Lemma 3.5.4

**Proof of Lemma 3.5.4.** The first derivative of the log-likelihood function  $l_t(\lambda)$  with respect to  $\varphi_0 = (\mu_0, \phi_{01}, \psi_{01})'$  is:

$$\frac{\partial l_{0t}(\lambda)}{\partial \varphi} = -\frac{\partial \ln h_{0t}}{\partial \varphi} \left(1 - \frac{\xi_{0t}^2}{h_{0t}}\right) - 2\frac{\xi_{0t}}{h_{0t}} \frac{\partial \xi_{0t}}{\partial \varphi}.$$

Then:

$$E \left\| \frac{\partial l_{0t}}{\partial \varphi} \right\|^2 \leq E \left\| \frac{\partial \ln h_{0t}}{\partial \varphi} \right\|^2 E |1 - z_{0t}^2|^2 + 2E \left\| z_{0t} \frac{1}{\sqrt{h_{0t}}} \frac{\partial \xi_{0t}}{\partial \varphi} \right\|^2 \quad (3.76)$$

by Minkowski's inequality and the independent relationship between  $z_{0t}$  and  $h_{0t}$ .

In the second term of the inequality (3.76), if  $E[y_t^2] < \infty$ , then it ensures that:

$$E \left\| \frac{\partial \xi_{0t}}{\partial \varphi} \right\|^2 < \infty, \quad (3.77)$$

and because  $h_{0t}$  has a lower bound uniformly for all  $\lambda \in \Theta$ , thus:

$$E \sup_{\lambda \in \Theta} \left\| \frac{1}{\sqrt{h_{0t}}} \frac{\partial \xi_{0t}}{\partial \varphi} \right\|^2 < \infty. \quad (3.78)$$



By Lemma 3.5.2,  $E [(\partial\xi_{0t}/\partial\varphi) (\partial\xi_{0t}/\partial\varphi)']$  is positive defined, and  $\sqrt{h_{0t}} > 0$ , thus:

$$E \left[ \frac{1}{\sqrt{h_{0t}}} \frac{\partial\xi_{0t}}{\partial\varphi} \frac{\partial\xi_{0t}}{\partial\varphi'} \right] > 0. \quad (3.79)$$

For the first term of the inequality (3.76), since the  $\ln h_{0t}$  can be rewritten as:

$$\ln h_{0t} = w_0 + \theta_0 \frac{\xi_{0t-1}}{\sqrt{h_{0t-1}}} + \gamma_0 \left| \frac{\xi_{0t-1}}{\sqrt{h_{0t-1}}} \right| + \beta_0 \ln h_{0t-1}.$$

Then, the first derivative of the  $\ln h_t$  with respect to parameters  $\varphi_0$  in the conditional mean equation are:

$$\begin{aligned} \frac{\partial \ln h_{0t}}{\partial\varphi} &= \left( \theta_0 \frac{\partial\xi_{0t-1}}{\partial\varphi} + \gamma_0 \left| \frac{\partial\xi_{0t-1}}{\partial\varphi} \right| \right) \frac{1}{\sqrt{h_{0t-1}}} \\ &\quad + \left( \beta_0 - \frac{1}{2}\theta_0 z_{0t-1} - \frac{1}{2}\gamma_0 |z_{0t-1}| \right) \frac{\partial \ln h_{0t-1}}{\partial\varphi}, \end{aligned}$$

and:

$$\begin{aligned} E \left\| \frac{\partial \ln h_{0t}}{\partial\varphi} \right\|^2 &\leq E \left\| \theta_0 \frac{\partial\xi_{0t-1}}{\partial\varphi} \frac{1}{\sqrt{h_{0t-1}}} \right\|^2 + E \left\| \gamma_0 \left| \frac{\partial\xi_{0t-1}}{\partial\varphi} \right| \frac{1}{\sqrt{h_{0t-1}}} \right\|^2 \\ &\quad + E \left| \beta_0 - \frac{1}{2}\theta_0 z_{0t-1} - \frac{1}{2}\gamma_0 |z_{0t-1}| \right|^2 E \left\| \frac{\partial \ln h_{0t-1}}{\partial\varphi} \right\|^2. \end{aligned} \quad (3.80)$$

According to the stationarity property of  $E \|\partial \ln h_{0t}/\partial\varphi\|^2$ , and if:

$$E \left| \beta_0 - \frac{1}{2}\theta_0 z_{0t-1} - \frac{1}{2}\gamma_0 |z_{0t-1}| \right|^2 < 1, \quad (3.81)$$

then:

$$E \left\| \frac{\partial \ln h_{0t}}{\partial\varphi} \right\|^2 < \infty.$$

And suppose  $E [z_{0t}]^4 < \infty$ , then:

$$E |1 - z_{0t}^2|^2 < \infty.$$

Thus, under the conditions  $E [y_t]^2 < \infty$ ,  $E [z_{0t}]^4 < \infty$ , the inequality (3.81), and

$h_t$  has a lower bound:

$$E \left\| \frac{\partial l_{0t}}{\partial\varphi} \right\|^2 < \infty.$$

The derivatives of the log-likelihood function with respect to  $\varsigma_0$ :

$$\frac{\partial l_{0t}}{\partial \varsigma} = -\frac{\partial \ln h_{0t}}{\partial \varsigma} \left(1 - \frac{\xi_{0t}^2}{h_{0t}}\right),$$

where the parameter vector  $\varsigma'_0 = (w_0, \beta_0, \theta_0, \gamma_0)$ , similarly,

$$\begin{aligned} E \left\| \frac{\partial l_{0t}}{\partial \varsigma} \right\|^2 &= E \left\| -\frac{\partial \ln h_{0t}}{\partial \varsigma} \left(1 - \frac{\xi_{0t}^2}{h_{0t}}\right) \right\|^2 \\ &= E \left\| \frac{\partial \ln h_{0t}}{\partial \varsigma} (1 - z_{0t}^2) \right\|^2 \\ &= E \left\| \frac{\partial \ln h_{0t}}{\partial \varsigma} \right\|^2 E |1 - z_{0t}^2|^2, \end{aligned}$$

where:

$$\frac{\partial \ln h_{0t}}{\partial \varsigma} = \begin{pmatrix} 1 \\ \ln h_{0t-1} \\ z_{0t-1} \\ |z_{0t-1}| \end{pmatrix} + \left( \beta_0 - \frac{1}{2}\theta_0 z_{0t-1} - \frac{1}{2}\gamma_0 |z_{0t-1}| \right) \frac{\partial \ln h_{0t-1}}{\partial \varsigma}.$$

Then:

$$\begin{aligned} &E \left\| \frac{\partial \ln h_{0t}}{\partial \varsigma} \right\|^2 \\ &= E \left\| \begin{pmatrix} 1 \\ \ln h_{0t-1} \\ z_{0t-1} \\ |z_{0t-1}| \end{pmatrix} + \left( \beta_0 - \frac{1}{2}\theta_0 z_{0t-1} - \frac{1}{2}\gamma_0 |z_{0t-1}| \right) \frac{\partial \ln h_{0t-1}}{\partial \varsigma} \right\|^2 \\ &\leq E \left\| \begin{pmatrix} 1 \\ \ln h_{0t-1} \\ z_{0t-1} \\ |z_{0t-1}| \end{pmatrix} \right\|^2 + E \left\| \left( \beta_0 - \frac{1}{2}\theta_0 z_{0t-1} - \frac{1}{2}\gamma_0 |z_{0t-1}| \right) \frac{\partial \ln h_{0t-1}}{\partial \varsigma} \right\|^2. \end{aligned}$$

By the stationarity of  $\partial \ln h_{0t}/\partial \varsigma$ , and if condition (3.81) holds, then:

$$E \left\| \frac{\partial \ln h_{0t}}{\partial \varsigma} \right\|^2 < \infty.$$

By Lemma 3.5.2,  $\partial \ln h_{0t} / \partial \varsigma \neq 0$ , then:

$$E \left\| \frac{\partial \ln h_{0t}}{\partial \varsigma} \right\|^2 > 0.$$

In addition, since:

$$\begin{aligned} & E \left[ \frac{\partial l_{0t}}{\partial \varphi} \frac{\partial l_{0t}}{\partial \varsigma'} \right] \\ = & E \left[ \left( -\frac{\partial \ln h_{0t}}{\partial \varphi} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) - 2 \frac{\xi_{0t}}{h_{0t}} \frac{\partial \xi_{0t}}{\partial \varphi} \right) \left( -\frac{\partial \ln h_{0t}}{\partial \varsigma'} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) \right) \right] \\ = & E \left[ \frac{\partial \ln h_{0t}}{\partial \varphi} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) \frac{\partial \ln h_{0t}}{\partial \varsigma'} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) + 2 \frac{\xi_{0t}}{h_{0t}} \frac{\partial \xi_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) \right] \\ = & E \left[ \frac{\partial \ln h_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right)^2 \right] + 2E \left[ \frac{\xi_{0t}}{h_{0t}} \frac{\partial \xi_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) \right] \\ = & E \left[ \frac{\partial \ln h_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} (1 - z_{0t}^2)^2 \right] + 2E \left[ \frac{\partial \xi_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \frac{1}{\sqrt{h_{0t}}} z_{0t} (1 - z_{0t}^2) \right], \end{aligned}$$

where:

$$\begin{aligned} & 2E \left[ \frac{\partial \xi_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \frac{1}{\sqrt{h_{0t}}} z_{0t} (1 - z_{0t}^2) \right] \\ = & 2E \left[ \frac{\partial \xi_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \frac{1}{\sqrt{h_{0t}}} \right] 2E [z_{0t} (1 - z_{0t}^2)] \\ = & 0, \end{aligned}$$

when  $z_{0t}$  is symmetric distribution. This also holds because the term  $\frac{\partial \xi_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \frac{1}{\sqrt{h_{0t}}}$  is independent with  $z_{0t}$ . Thus:

$$E \left[ \frac{\partial l_{0t}}{\partial \varphi} \frac{\partial l_{0t}}{\partial \varsigma'} \right] = E \left[ \frac{\partial \ln h_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} (1 - z_{0t}^2)^2 \right],$$

and suppose  $E[z_{0t}]^4 < \infty$ , then:

$$\begin{aligned} & E \left[ \frac{\partial \ln h_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} (1 - z_{0t}^2)^2 \right] \\ = & E \left[ \frac{\partial \ln h_{0t}}{\partial \varphi} \frac{\partial \ln h_{0t}}{\partial \varsigma'} \right] E [(1 - z_{0t}^2)^2] \\ < & \infty. \end{aligned}$$

Therefore,  $0 < \Omega_0 < \infty$ . Let:

$$S_n = \sum_{t=1}^n c' \frac{\partial l_{0t}}{\partial \lambda} \quad (3.82)$$

where  $c$  is a constant vector with the same dimension as  $\lambda$ . Then  $S_n$  is a martingale array with respect to  $F_t$ . And since:

$$E \left[ \frac{S_n}{n} \right] = c' E \left[ \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right] c > 0,$$

then by the central limit theorem of Stout (1974),  $\sqrt{n}S_n$  converges to  $N(0, c'\Omega_0 c)$  in distribution. Finally, by the Cramer-Wold device,  $n^{-1/2} \sum_{t=1}^n \partial l_{0t} / \partial \lambda$  converges to  $N(0, \Omega_0)$  in distribution. Similarly to the Lemma 3.5.3, it can be obtained that:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| \frac{\partial l_{0t}}{\partial \lambda} - \frac{\partial \tilde{l}_{0t}}{\partial \lambda} \right\| = o_p(1). \quad (3.83)$$

Since:

$$\begin{aligned} & \left\| \frac{\partial l_{0t}}{\partial \lambda} - \frac{\partial \tilde{l}_{0t}}{\partial \lambda} \right\| \\ \leq & \left\| \frac{\partial \ln \tilde{h}_{0t}}{\partial \lambda} \right\| \left| \frac{\xi_{0t}^2}{h_{0t}} - \frac{\tilde{\xi}_{0t}^2}{\tilde{h}_{0t}} \right| + \left\| \frac{\partial \ln \tilde{h}_{0t}}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} \right\| \left| 1 - \frac{\xi_{0t}^2}{h_{0t}} \right| \\ & + 2 \left| \frac{\tilde{\xi}_{0t}}{h_{0t}} - \frac{\xi_{0t}}{h_{0t}} \right| \left\| \frac{\partial \tilde{\xi}_{0t}}{\partial \varphi} \right\| + 2 \left| \frac{\xi_{0t}}{h_{0t}} \right| \left\| \frac{\partial \tilde{\xi}_{0t}}{\partial \lambda} - \frac{\partial \xi_{0t}}{\partial \lambda} \right\|, \end{aligned}$$

by Lemma 3.5.3,

$$\sup_{\lambda \in \Theta} \frac{1}{n} \sum_{t=1}^n \left| \frac{\xi_{0t}^2}{h_{0t}} - \frac{\tilde{\xi}_{0t}^2}{\tilde{h}_{0t}} \right| = o_p(1), \quad (3.84)$$

and since:

$$\left| \frac{\tilde{\xi}_{0t}}{\tilde{h}_{0t}} - \frac{\xi_{0t}}{h_{0t}} \right| \leq |\xi_{0t}| \left| \frac{1}{\tilde{h}_{0t}} - \frac{1}{h_{0t}} \right| + \frac{1}{\tilde{h}_t} |\xi_{0t} - \tilde{\xi}_{0t}|,$$

and:

$$\sup_{\lambda \in \Theta} \frac{1}{n} \sum_{t=1}^n \left| \frac{1}{\tilde{h}_{0t}} - \frac{1}{h_{0t}} \right| = o_p(1), \quad (3.85)$$

then:

$$\sup_{\lambda \in \Theta} \frac{1}{n} \sum_{t=1}^n \left| \frac{\tilde{\xi}_{0t}}{\tilde{h}_{0t}} - \frac{\xi_{0t}}{h_{0t}} \right| = o_p(1). \quad (3.86)$$

And because:

$$\left\| \frac{\partial \tilde{\xi}_{0t}}{\partial \lambda} - \frac{\partial \xi_{0t}}{\partial \lambda} \right\| = \left\| \frac{\partial \tilde{\xi}_{0t}}{\partial \varphi} - \frac{\partial \xi_{0t}}{\partial \varphi} \right\|,$$

and:

$$\left\| \frac{\partial \tilde{\xi}_{0t}}{\partial \varphi} - \frac{\partial \xi_{0t}}{\partial \varphi} \right\| = \left\| \begin{array}{c} 0 \\ \sum_{i=t}^{\infty} (-\psi_1)^i (Y_{t-i} - \mu) \\ \sum_{i=t}^{\infty} (-\psi_1)^i (\xi_{t-1} - \tilde{\xi}_{t-1}) \end{array} \right\|_{\varphi_0}.$$

Thus:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \left\| \frac{\partial \tilde{\xi}_{0t}}{\partial \lambda} - \frac{\partial \xi_{0t}}{\partial \lambda} \right\| \right] \right\| = o_p(1). \quad (3.87)$$

For term:

$$\frac{1}{n} \sum_{t=1}^n \left\| \frac{\partial \ln \tilde{h}_{0t}}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} \right\|,$$

where  $\lambda' = \text{vec}(\mu, \phi_1, \psi_1, w, \beta, \theta, \gamma)$ , and by subsection 3.4.4:

$$\left| \frac{\partial \ln \tilde{h}_{0t}}{\partial \beta} - \frac{\partial \ln h_{0t}}{\partial \beta} \right| = o_p(1),$$

and similarly it is able to obtain the convergences for other parameters, then:

$$\frac{1}{n} \sum_{t=1}^n \left\| \frac{\partial \ln \tilde{h}_{0t}}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} \right\| = o_p(1).$$

Thus,  $n^{-\frac{1}{2}} \sum_{t=1}^n \left\| \partial \tilde{l}_{0t} / \partial \lambda \right\|$  converges to  $N(0, \Omega_0)$  in distribution. This completes the proof. ■

### 3.8.7 Proof of Lemma 3.5.5

**Proof of Lemma 3.5.5.** To show (i), according to the expression of the ARMA(1,1)-EGARCH(1,1) model and the log-likelihood functions, the second order derivative of the log-likelihood with respect to the parameter vector  $\lambda$ :

$$\begin{aligned} \frac{\partial^2 l_t(\lambda)}{\partial \lambda \partial \lambda'} &= 2 \frac{\xi_t}{h_t} \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} - \frac{\xi_t^2}{h_t} \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} + \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \\ &\quad - 2 \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \xi_t}{\partial \lambda} \frac{1}{h_t} - 2 \frac{\xi_t}{h_t} \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} + 2 \frac{\xi_t}{h_t} \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'}. \end{aligned}$$

Taking the expectation on both sides of the above equation:

$$\begin{aligned}
E \left[ \frac{\partial^2 l_t(\lambda)}{\partial \lambda \partial \lambda'} \right] &= E \left[ 2 \frac{\xi_t}{h_t} \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} - \frac{\xi_t^2}{h_t} \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} + \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right. \\
&\quad \left. - 2 \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \xi_t}{\partial \lambda} \frac{1}{h_t} - 2 \frac{\xi_t}{h_t} \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} + 2 \frac{\xi_t}{h_t} \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} \right] \\
&= -E \left[ \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right] - 2E \left[ \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \xi_t}{\partial \lambda} (h_t)^{-1} \right],
\end{aligned}$$

since  $\xi_t = \xi_{0t}$  and  $\ln h_t = \ln h_{0t}$  when  $\lambda = \lambda_0$ , and  $z_t = z_{0t} \sim iid(0, 1)$ , then:

$$\begin{aligned}
2E \left[ \frac{\xi_t}{h_t} \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right] &= 2E \left[ z_t (h_t)^{-\frac{1}{2}} \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right] = 0, \\
E \left[ \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right] &= E \left[ (z_t^2 - 1) \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right] = 0, \\
2E \left[ \frac{\xi_t}{h_t} \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} \right] &= 2E \left[ z_t (h_t)^{-\frac{1}{2}} \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} \right] = 0, \\
2E \left[ \frac{\xi_t}{h_t} \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} \right] &= 2E \left[ z_t (h_t)^{-\frac{1}{2}} \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} \right] = 0,
\end{aligned}$$

and:

$$E \left[ \frac{\xi_t^2}{h_t} \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right] = E \left[ z_t^2 \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right] = E \left[ \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right].$$

So, to show  $E[\partial^2 l_t(\lambda)/\partial \lambda \partial \lambda'] < \infty$ , it is essential to prove that:

$$E \left[ \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right] < \infty, \tag{3.88}$$

and:

$$E \left[ \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \xi_t}{\partial \lambda} \frac{1}{h_t} \right] < \infty. \tag{3.89}$$

By Lemma 3.5.4, the conditions (3.88) and (3.89) can be held. Therefore:

$$E \sup_{\lambda \in \Theta} \left\| \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} \right\| < \infty.$$

Thus, according to the modified uniform convergence theorem:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} - E \left[ \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} \right] \right\| = o_p(1), \tag{3.90}$$

holds. And also:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} - E \left[ \frac{\partial^2 l_{0t}}{\partial \lambda \partial \lambda'} \right] \right\| = o_p(1). \quad (3.91)$$

Proof of (ii) Similarly to the Lemma 3.5.3:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \tilde{l}_t}{\partial \lambda \partial \lambda'} \right] \right\| = o_p(1). \quad (3.92)$$

Since:

$$\begin{aligned} & \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \tilde{l}_t}{\partial \lambda \partial \lambda'} \\ = & \left( \frac{\xi_t^2}{h_t} - 1 \right) \left( \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \ln \tilde{h}_t}{\partial \lambda \partial \lambda'} \right) + \left( \frac{\xi_t^2}{h_t} - \frac{\tilde{\xi}_t^2}{\tilde{h}_t} \right) \left( \frac{\partial^2 \ln \tilde{h}_t}{\partial \lambda \partial \lambda'} - \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} \frac{\partial \ln \tilde{h}_t}{\partial \lambda} \right) \\ & + \left( 2 \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} - 2 \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} + 2 \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} \right) \left( \frac{\xi_t}{h_t} - \frac{\tilde{\xi}_t}{\tilde{h}_t} \right) \\ & + 2 \left( \frac{\partial^2 \tilde{\xi}_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} \right) \frac{\tilde{\xi}_t}{\tilde{h}_t} + 2 \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \xi_t}{\partial \lambda} \left( \frac{1}{\tilde{h}_t} - \frac{1}{h_t} \right) \\ & + L_{11} + L_{12} + L_{13} + L_{14}, \end{aligned}$$

where:

$$\begin{aligned} L_{11} &= 2 \left( \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} - \frac{\partial \tilde{\xi}_t}{\partial \lambda'} \frac{\partial \ln \tilde{h}_t}{\partial \lambda} \right) \frac{\tilde{\xi}_t}{\tilde{h}_t} \\ &= 2 \frac{\partial \xi_t}{\partial \lambda'} \left( \frac{\partial \ln h_t}{\partial \lambda} - \frac{\partial \ln \tilde{h}_t}{\partial \lambda} \right) \frac{\tilde{\xi}_t}{\tilde{h}_t} + 2 \left( \frac{\partial \xi_t}{\partial \lambda'} - \frac{\partial \tilde{\xi}_t}{\partial \lambda'} \right) \frac{\partial \ln \tilde{h}_t}{\partial \lambda} \frac{\tilde{\xi}_t}{\tilde{h}_t}, \\ L_{12} &= 2 \left( \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \tilde{\xi}_t}{\partial \lambda} \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} \right) \frac{\tilde{\xi}_t}{\tilde{h}_t} \\ &= 2 \frac{\partial \xi_t}{\partial \lambda} \left( \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} \right) \frac{\tilde{\xi}_t}{\tilde{h}_t} + 2 \left( \frac{\partial \xi_t}{\partial \lambda} - \frac{\partial \tilde{\xi}_t}{\partial \lambda} \right) \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} \frac{\tilde{\xi}_t}{\tilde{h}_t}, \\ L_{13} &= \frac{\xi_t^2}{h_t} \left( \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} \frac{\partial \ln \tilde{h}_t}{\partial \lambda} - \frac{\partial \ln h_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \right) \\ &= \frac{\xi_t^2}{h_t} \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} \left( \frac{\partial \ln \tilde{h}_t}{\partial \lambda} - \frac{\partial \ln h_t}{\partial \lambda} \right) + \frac{\xi_t^2}{h_t} \left( \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} - \frac{\partial \ln h_t}{\partial \lambda'} \right) \frac{\partial \ln h_t}{\partial \lambda}, \end{aligned}$$

$$\begin{aligned}
L_{14} &= 2 \left( \frac{\partial \tilde{\xi}_t}{\partial \lambda'} \frac{\partial \tilde{\xi}_t}{\partial \lambda} - \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \xi_t}{\partial \lambda} \right) \frac{1}{\tilde{h}_t} \\
&= 2 \frac{\partial \tilde{\xi}_t}{\partial \lambda'} \left( \frac{\partial \tilde{\xi}_t}{\partial \lambda} - \frac{\partial \xi_t}{\partial \lambda} \right) \frac{1}{\tilde{h}_t} + 2 \left( \frac{\partial \tilde{\xi}_t}{\partial \lambda'} - \frac{\partial \xi_t}{\partial \lambda'} \right) \frac{\partial \xi_t}{\partial \lambda} \frac{1}{\tilde{h}_t}.
\end{aligned}$$

Similar to the subsection 3.4.4, it is able to obtain:

$$\left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \ln \tilde{h}_t}{\partial \lambda \partial \lambda'} \right\| = o_p(1),$$

and similar to the proof of the Lemma 3.5.4:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \left\| \frac{\partial^2 \tilde{\xi}_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} \right\| \right] \right\| = o_p(1), \quad (3.93)$$

then:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \left( 1 - \frac{\xi_t^2}{h_t} \right) \left( \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \ln \tilde{h}_t}{\partial \lambda \partial \lambda'} \right) \right] \right\| = o_p(1), \quad (3.94)$$

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \left( \frac{\partial^2 \tilde{\xi}_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \xi_t}{\partial \lambda \partial \lambda'} \right) \frac{\tilde{\xi}_t}{\tilde{h}_t} \right] \right\| = o_p(1), \quad (3.95)$$

and by Lemma 3.5.3 and 3.5.4:

$$\sup_{\lambda \in \Theta} \frac{1}{n} \sum_{t=1}^n \left| \frac{\xi_t^2}{h_t} - \frac{\tilde{\xi}_t^2}{\tilde{h}_t} \right| = o_p(1), \quad (3.96)$$

and:

$$\sup_{\lambda \in \Theta} \frac{1}{n} \sum_{t=1}^n \left| \frac{\xi_t}{h_t} - \frac{\tilde{\xi}_t}{\tilde{h}_t} \right| = o_p(1). \quad (3.97)$$

In addition:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial \xi_t}{\partial \lambda'} - \frac{\partial \tilde{\xi}_t}{\partial \lambda'} \right] \right\| = o_p(1), \quad (3.98)$$

and:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial \ln \tilde{h}_t}{\partial \lambda'} - \frac{\partial \ln h_t}{\partial \lambda'} \right] \right\| = o_p(1). \quad (3.99)$$

Thereby:

$$\sup_{\lambda \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial^2 l_t}{\partial \lambda \partial \lambda'} - \frac{\partial^2 \tilde{l}_t}{\partial \lambda \partial \lambda'} \right] \right\| = o_p(1). \quad (3.100)$$

This completes the proof. ■



### 3.8.8 Proof of Lemma 3.5.6

**Proof of Lemma 3.5.6.** To prove (i), since:

$$\begin{aligned}
& \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} - \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right] \\
= & \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \\
& - 2 \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\xi_t}{h_t} \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \xi_t}{\partial \lambda'} \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \xi_t}{\partial \lambda'} + 2 \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \frac{\xi_{0t}}{h_{0t}} \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \xi_{0t}}{\partial \lambda'} \\
& - 2 \frac{\xi_t}{h_t} \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} + \frac{\xi_{0t}}{h_{0t}} \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \\
& + 4 \frac{\xi_t}{h_t} \frac{\xi_t}{h_t} \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \xi_t}{\partial \lambda'} - 4 \frac{\xi_{0t}}{h_{0t}} \frac{\xi_{0t}}{h_{0t}} \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \xi_{0t}}{\partial \lambda'}.
\end{aligned}$$

Then, it is necessary to hold the following conditions (ignore the constant) to ensure the convergence of the difference between  $(\partial l_t / \partial \lambda)^2$  and  $(\partial l_{0t} / \partial \lambda)^2$ , firstly considers the convergence of the following equation:

$$\begin{aligned}
& \left\| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \tag{3.101} \\
\leq & \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\
& + \left| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2 \\
= & \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - 2 \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} + \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right. \\
& \left. - 2 \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} + 2 \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| + \left| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2 \\
\leq & \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} \right\|^2 + 2 \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} \right\| \left\| \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\
& + \left| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2.
\end{aligned}$$

In the third term of the inequality (3.101):

$$\begin{aligned}
& \left| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \\
& \leq \left| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 - 2 \left( \frac{\xi_t^2}{h_t} - 1 \right) \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) + \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \\
& \quad + \left| 2 \left( \frac{\xi_t^2}{h_t} - 1 \right) \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) - 2 \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \\
& = \left| \left( \frac{\xi_t^2}{h_t} - 1 \right) - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \right|^2 + 2 \left| \left( \frac{\xi_t^2}{h_t} - 1 \right) - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \right| \left| \frac{\xi_{0t}^2}{h_{0t}} - 1 \right| \\
& = \left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right|^2 + 2 \left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right| \left| \frac{\xi_{0t}^2}{h_{0t}} - 1 \right|,
\end{aligned}$$

where:

$$\begin{aligned}
\left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right| &= \left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right| \\
&= \left| \frac{\xi_t^2}{h_t} - 2 \frac{\xi_t \xi_{0t}}{\sqrt{h_t h_{0t}}} + \frac{\xi_{0t}^2}{h_{0t}} + 2 \frac{\xi_t \xi_{0t}}{\sqrt{h_t h_{0t}}} - 2 \frac{\xi_{0t}^2}{h_{0t}} \right| \\
&\leq \left| \frac{\xi_t}{\sqrt{h_t}} - \frac{\xi_{0t}}{\sqrt{h_{0t}}} \right|^2 + 2 \left| \frac{\xi_t}{\sqrt{h_t}} - \frac{\xi_{0t}}{\sqrt{h_{0t}}} \right| \left| \frac{\xi_{0t}}{\sqrt{h_{0t}}} \right|,
\end{aligned}$$

and:

$$\begin{aligned}
& \left| \frac{\xi_t}{\sqrt{h_t}} - \frac{\xi_{0t}}{\sqrt{h_{0t}}} \right| \\
& \leq |\xi_t - \xi_{0t}| \frac{1}{\sqrt{h_t}} + \left| \sqrt{h_t} - \sqrt{h_{0t}} \right| \frac{|\xi_{0t}|}{\sqrt{h_t h_{0t}}} \\
& \leq \|\sqrt{n}(\lambda_n - \lambda_0)\| \frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \left( \left\| \frac{\partial \xi_t}{\partial \lambda} \right\|_{\lambda_{1n}^\dagger} \right) \frac{1}{\sqrt{h_t}} \\
& \quad + \frac{|\xi_{0t}|}{\sqrt{h_{0t}}} \frac{1}{\sqrt{n}} \|\sqrt{n}(\lambda_n - \lambda_0)\| \max_{1 \leq t \leq n} \left( \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|_{\lambda_{2n}^\dagger} \right) \\
& = o_p(1) + o_p(1) |z_{0t}|^2,
\end{aligned}$$

where  $\lambda_{1n}^\dagger$  and  $\lambda_{2n}^\dagger$  lie between  $\lambda_n$  and  $\lambda_0$ . Thus:

$$\left| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \leq o_p(1) + o_p(1) |z_{0t}|^4.$$

And by the Taylor expansion:

$$\frac{\partial \ln h_t}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} = \left. \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right|_{\lambda_n^\dagger} (\lambda_n - \lambda_0), \quad (3.102)$$

where  $\lambda_n^\dagger$  lies between  $\lambda_n$  and  $\lambda_0$ . For the conditional variance equation, by Lemma 3.5.4:

$$E \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2 < \infty.$$

Then, according to Chung (1968):

$$\frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \sup_{\lambda \in \Theta} \left| \frac{\partial \ln h_t}{\partial \lambda} \right| = o_p(1), \quad (3.103)$$

and:

$$\frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \sup_{\lambda \in \Theta} \left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right\| = o_p(1) \quad (3.104)$$

as well. Then:

$$\begin{aligned} & \left\| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\ \leq & \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \left\| \frac{\partial \ln h_t}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} \right\|^2 + 2 \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \frac{\partial \ln h_t}{\partial \lambda} - \frac{\partial \ln h_{0t}}{\partial \lambda} \right\| \left\| \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\ & + (o_p(1) + o_p(1)|z_{0t}|^4) \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2 \\ = & \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \left\| \left. \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right|_{\tilde{\lambda}_n} (\lambda_n - \lambda_0) \right\|^2 + 2 \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \left\| \left. \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right|_{\tilde{\lambda}_n} (\lambda_n - \lambda_0) \right\| \left\| \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\ & + (o_p(1) + o_p(1)|z_{0t}|^4) \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2 \\ \leq & \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \frac{1}{\sqrt{n}} \|\sqrt{n}(\lambda_n - \lambda_0)\|^2 \left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right\|^2 + \left| \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \right| \frac{1}{\sqrt{n}} \|\sqrt{n}(\lambda_n - \lambda_0)\| \\ & \left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right\| \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\| + (o_p(1) + o_p(1)|z_{0t}|^4) \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2 \\ = & o_p(1) \left[ \left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right\|^2 + \left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right\| \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\| + \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2 \right] [1 + |z_{0t}|^4] \\ = & o_p(1)H [1 + |z_{0t}|^4], \end{aligned}$$

where  $H = \left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right\| \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\| + \left\| \frac{\partial^2 \ln h_t}{\partial \lambda \partial \lambda'} \right\|^2 + \left\| \frac{\partial \ln h_t}{\partial \lambda} \right\|^2$ . Then, under the conditions  $E[y_t]^4 < \infty$ ,  $E[z_t]^4 < \infty$ , thus,  $E[H(1 + |z_{0t}|^4)] < \infty$ . According to the ergodic theorem,

$$\frac{1}{n} \sum_{t=1}^n H(1 + |z_{0t}|^4) = o_p(1). \quad (3.105)$$

That is:

$$\frac{1}{n} \sum_{t=1}^n \left\| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 \frac{\partial \ln h_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right)^2 \frac{\partial \ln h_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| = o_p(1). \quad (3.106)$$

Then, for the term:

$$\begin{aligned} & \left\| \frac{\xi_t^2}{h_t^2} \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \xi_t}{\partial \lambda'} - \frac{\xi_{0t}^2}{h_{0t}^2} \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \xi_{0t}}{\partial \lambda'} \right\| \\ & \leq \left\| \frac{\xi_t^2}{h_t^2} - \frac{\xi_{0t}^2}{h_{0t}^2} \right\| \left\| \frac{\partial \xi_t}{\partial \varphi} \frac{\partial \xi_t}{\partial \varphi'} \right\| + \left\| \frac{\xi_{0t}^2}{h_{0t}^2} \right\| \left\| \frac{\partial \xi_t}{\partial \varphi} \frac{\partial \xi_t}{\partial \varphi'} - \frac{\partial \xi_{0t}}{\partial \varphi} \frac{\partial \xi_{0t}}{\partial \varphi'} \right\| \\ & \leq \left\| \frac{z_t^2}{h_t} - \frac{z_{0t}^2}{h_{0t}} \right\| \left\| \frac{\partial \xi_t}{\partial \varphi} \right\|^2 + \left\| \frac{z_{0t}^2}{h_{0t}} \right\| \left\| \frac{\partial \xi_t}{\partial \varphi} - \frac{\partial \xi_{0t}}{\partial \varphi} \right\|^2 \\ & \quad + 2 \left\| \frac{z_{0t}^2}{h_{0t}} \right\| \left\| \frac{\partial \xi_t}{\partial \varphi} - \frac{\partial \xi_{0t}}{\partial \varphi} \right\| \left\| \frac{\partial \xi_{0t}}{\partial \varphi'} \right\|, \end{aligned}$$

since

$$\left\| \frac{\partial \xi_t}{\partial \varphi} \right\| \leq c_0 + c_1 \sum_{i=1}^{\infty} \rho_1^i |y_{t-i}| = R_{1t},$$

$E[y_t^2] < \infty$ . Then, following Chung (1968):

$$\frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \sup_{\lambda \in \Theta} \left| \frac{\partial \xi_t}{\partial \varphi} \right| = o_p(1). \quad (3.107)$$

And according to the Taylor expansion:

$$\frac{\partial \xi_t}{\partial \varphi} - \frac{\partial \xi_{0t}}{\partial \varphi} = \frac{\partial^2 \xi_{0t}}{\partial \varphi \partial \varphi'} \Big|_{\varphi_n^\dagger} (\varphi_n - \varphi_0), \quad (3.108)$$

where  $\varphi_n^\dagger$  lies between  $\varphi_n$  and  $\varphi_0$ . Since:

$$\begin{aligned} & \left| \frac{\xi_t^2}{h_t^2} - \frac{\xi_{0t}^2}{h_{0t}^2} \right| \\ & \leq \left| \frac{\xi_t^2}{h_t^2} - \frac{\xi_{0t}^2}{h_t h_{0t}} \right| + \left| \frac{\xi_{0t}^2}{h_t h_{0t}} - \frac{\xi_{0t}^2}{h_{0t}^2} \right| \\ & = \frac{1}{h_t} \left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right| + \left| \frac{1}{h_t} - \frac{1}{h_{0t}} \right| |z_{0t}^2|, \end{aligned}$$

where:

$$\left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right| = o_p(1) + o_p(1)|z_{0t}|^2,$$

and:

$$\max \left| \frac{1}{h_t} - \frac{1}{h_{0t}} \right| = \|\sqrt{n}(\lambda_n - \lambda_0)\| \frac{1}{\sqrt{n}} \max \left\| \frac{1}{h_t} \frac{\partial \ln h_t}{\partial \lambda} \right\| = o_p(1).$$

Then:

$$\begin{aligned} & \left| \frac{1}{h_t} \left( 1 - \frac{\xi_t^2}{h_t} \right) - \frac{1}{h_{0t}} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) \right| \\ = & \left| \frac{1}{h_t} \left( 1 - \frac{\xi_t^2}{h_t} \right) - \frac{1}{h_t} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) + \frac{1}{h_t} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) - \frac{1}{h_{0t}} \left( 1 - \frac{\xi_{0t}^2}{h_{0t}} \right) \right| \\ \leq & \frac{1}{h_t} \left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right| + \left| \frac{1}{h_t} - \frac{1}{h_{0t}} \right| \left| 1 - \frac{\xi_{0t}^2}{h_{0t}} \right| \\ = & o_p(1) + o_p(1)|z_{0t}|^2. \end{aligned}$$

Thus:

$$\frac{1}{n} \sum_{t=1}^n 4 \left\| \frac{\xi_t^2}{h_t^2} \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \xi_t}{\partial \lambda'} - \frac{\xi_{0t}^2}{h_{0t}^2} \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \xi_{0t}}{\partial \lambda'} \right\| = o_p(1). \quad (3.109)$$

Then, for the term:

$$\begin{aligned} & \left\| \frac{\xi_t}{h_t} \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\xi_{0t}}{h_{0t}} \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \quad (3.110) \\ \leq & \left| \frac{\xi_t}{h_t} \left( \frac{\xi_t^2}{h_t} - 1 \right) - \frac{\xi_{0t}}{h_{0t}} \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \right| \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} \right\| \\ & + \left\| \frac{\xi_{0t}}{h_{0t}} \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \right\| \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\|, \end{aligned}$$

and for the first term of the right hand side of inequality (3.110) :

$$\begin{aligned} & \left| \frac{\xi_t}{h_t} \left( \frac{\xi_t^2}{h_t} - 1 \right) - \frac{\xi_{0t}}{h_{0t}} \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \right| \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} \right\| \\ \leq & \left( \left| \frac{\xi_t}{h_t} - \frac{\xi_{0t}}{h_{0t}} \right| \left| \frac{\xi_t^2}{h_t} - 1 \right| + \left| \frac{\xi_{0t}}{h_{0t}} \right| \left| \frac{\xi_t^2}{h_t} - \frac{\xi_{0t}^2}{h_{0t}} \right| \right) \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} \right\| \\ = & o_p(1), \end{aligned}$$

and:

$$\begin{aligned}
& \left| \frac{\xi_{0t}}{h_{0t}} \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \right| \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\
&= \left| \frac{z_{0t}}{\sqrt{h_{0t}}} (z_{0t}^2 - 1) \right| \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\
&= o_p(1),
\end{aligned}$$

because of:

$$\begin{aligned}
& \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\
&= \left\| \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} + \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} - \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\
&\leq \left\| \frac{\partial \xi_t}{\partial \lambda} \right\| \left\| \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| + \left\| \frac{\partial \xi_t}{\partial \lambda} - \frac{\partial \xi_{0t}}{\partial \lambda} \right\| \left\| \frac{\partial \ln h_{0t}}{\partial \lambda'} \right\| \\
&= o_p(1).
\end{aligned}$$

Then:

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\xi_t}{h_t} \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial \xi_t}{\partial \lambda} \frac{\partial \ln h_t}{\partial \lambda'} - \frac{\xi_{0t}}{h_{0t}} \left( \frac{\xi_{0t}^2}{h_{0t}} - 1 \right) \frac{\partial \xi_{0t}}{\partial \lambda} \frac{\partial \ln h_{0t}}{\partial \lambda'} \right]_{\lambda_n} = o_p(1). \quad (3.111)$$

Finally:

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} - \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right]_{\lambda_n} = o_p(1) \quad (3.112)$$

(i) holds. Assume  $(\partial l_{0t}/\partial \lambda)(\partial l_{0t}/\partial \lambda')$  is strictly stationary and ergodic and:

$$E\|(\partial l_{0t}/\partial \lambda)(\partial l_{0t}/\partial \lambda')\| < \infty.$$

By the ergodic theorem, it is able to obtain that:

$$\begin{aligned}
\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} \right]_{\lambda_n} &= E \left[ \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right] + o_p(1) \\
&= \Omega_0 + o_p(1).
\end{aligned}$$

And similarly to Lemma 3.5.3, it is able to obtain:

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} - \frac{\partial \tilde{l}_t}{\partial \lambda} \frac{\partial \tilde{l}_t}{\partial \lambda'} \right]_{\lambda_n} = o_p(1). \quad (3.113)$$

Then:

$$\begin{aligned}
& \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial \tilde{l}_t}{\partial \lambda} \frac{\partial \tilde{l}_t}{\partial \lambda'} \right]_{\lambda_n} - E \left[ \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right] \tag{3.114} \\
&= \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial \tilde{l}_t}{\partial \lambda} \frac{\partial \tilde{l}_t}{\partial \lambda'} - \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right]_{\lambda_n} \\
&= \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} - \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} + \frac{\partial \tilde{l}_t}{\partial \lambda} \frac{\partial \tilde{l}_t}{\partial \lambda'} - \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} \right]_{\lambda_n} \\
&\leq \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} - \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right]_{\lambda_n} + \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \lambda} \frac{\partial l_t}{\partial \lambda'} - \frac{\partial \tilde{l}_t}{\partial \lambda} \frac{\partial \tilde{l}_t}{\partial \lambda'} \right]_{\lambda_n} \\
&= o_p(1).
\end{aligned}$$

Thus (ii) holds, that is:

$$\widehat{\Omega}_n - E \left[ \frac{\partial l_{0t}}{\partial \lambda} \frac{\partial l_{0t}}{\partial \lambda'} \right] = o_p(1). \tag{3.115}$$

This means that  $\widehat{\Omega}_n = \frac{1}{n} \sum_{t=1}^n \left[ (\partial \tilde{l}_t / \partial \lambda) (\partial \tilde{l}_t / \partial \lambda') \right]_{\lambda_n}$  is a consistent estimator of  $\Omega_0 = E [(\partial l_{0t} / \partial \lambda) (\partial l_{0t} / \partial \lambda')]$ . This completes the proof. ■

### 3.8.9 Proof of Lemma 3.6.1

**Proof of Lemma 3.6.1.** Since:

$$E|l_t(\vartheta)| = E|\ln h_t| + E \left[ \frac{\xi_t^2}{h_t} \right], \tag{3.116}$$

and by the expression of the HYGARCH(0, d, 0):

$$\begin{aligned}
E|\ln h_t| &= E \left| \omega + \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) \right| \\
&\leq |\omega| + E \left[ \sum_{j=1}^{\infty} \max_{\vartheta \in \Theta} |b_j g_1(z_{t-j})| \right] \\
&\leq |\omega| + E |g_1(z_1)| E \left[ \sum_{j=1}^{\infty} \max_{\vartheta \in \Theta} (c j^{-d-1}) \right] \\
&< \infty,
\end{aligned}$$

and suppose  $|\theta| \leq 1$ , then:

$$\begin{aligned}
\ln h_t &= \omega + \sum_{j=1}^{\infty} b_j g_1(z_{t-j}) \\
&= \omega + \sum_{j=1}^{\infty} b_j \left( \theta \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} + \left| \frac{\xi_{t-j}}{\sqrt{h_{t-j}}} \right| \right) \\
&\geq \omega.
\end{aligned}$$

Thus,  $\ln h_t \geq \omega$ , and  $h_t \geq \exp(\omega)$ , this ensures  $h_t$  has lower bound over  $\Theta$ . Thereby, under  $E[\xi_t^2] < \infty$ , and:

$$\begin{aligned}
\sup_{\vartheta \in \Theta} E|l_t(\vartheta)| &= \sup_{\vartheta \in \Theta} \left( E|\ln h_t| + E \left[ \frac{\xi_t^2}{h_t} \right] \right) \\
&\leq \sup_{\vartheta \in \Theta} (E|\ln h_t|) + CE[\xi_t^2] \\
&< \infty,
\end{aligned}$$

$L_n(\vartheta)$  converges to the function  $\bar{L}(\vartheta)$  by ergodicity of  $\xi_t^2$  and  $\ln h_t$  for each individual  $\vartheta \in \Theta$ . This completes the proof. ■

### 3.8.10 Proof of Lemma 3.6.2

**Proof of Lemma 3.6.2.** Since:

$$\begin{aligned}
&\bar{L}(\vartheta) - \bar{L}(\vartheta_0) \\
&= E[l_t(\vartheta)] - E[l_{0t}(\vartheta)] \\
&= -E \ln |h_t(\vartheta)| - E \left[ \frac{\xi_t^2}{h_t(\vartheta)} \right] + E \ln |h_{0t}(\vartheta)| + E \left[ \frac{\xi_t^2}{h_{0t}(\vartheta)} \right] \\
&= E \ln \left| \frac{h_{0t}(\vartheta)}{h_t(\vartheta)} \right| - E \left[ \frac{\xi_t^2}{h_t(\vartheta)} \right] + E \left[ \frac{\xi_t^2}{h_{0t}(\vartheta)} \right] \\
&= E \ln \left| \frac{h_{0t}(\vartheta)}{h_t(\vartheta)} \right| - E \left[ \frac{h_{0t}(\vartheta)}{h_t(\vartheta)} \right] + E \left[ \frac{h_{0t}(\vartheta)}{h_{0t}(\vartheta)} \right] \\
&= E \left[ \ln \left| \frac{h_{0t}(\vartheta)}{h_t(\vartheta)} \right| - \frac{h_{0t}(\vartheta)}{h_t(\vartheta)} + 1 \right],
\end{aligned}$$



and because of  $\ln x < x - 1$  for  $1 \neq x > 0$ , thus:

$$\bar{L}(\vartheta) < \bar{L}(\vartheta_0),$$

and:

$$\bar{L}(\vartheta) = \bar{L}(\vartheta_0),$$

for all  $\vartheta$  if and only if  $h_{0t}(\vartheta)/h_t(\vartheta) = 1$ , that is  $h_t(\vartheta_0) = h_t(\vartheta)$  almost surely, and also  $\ln h_t(\vartheta_0) = \ln h_t(\vartheta)$ ,  $g(z_t)(\vartheta_0) = g(z_t)(\vartheta)$  which means that:

$$\omega_0 + \sum_{j=1}^{\infty} c_0 j^{-d_0-1} g(z_{0t-j}) = \omega + \sum_{j=1}^{\infty} c j^{-d-1} g(z_{t-j}). \quad (3.117)$$

Taking the variance of both sides of the above equation and reorganised it, then:

$$\sum_{j=1}^{\infty} E [g(z_{0t-j})^2] (c_0 j^{-d_0-1} - c j^{-d-1})^2 = 0,$$

and  $0 < E [g(z_{0t-j})^2] < \infty$ , then:

$$\sum_{j=1}^{\infty} (c_0 j^{-d_0-1} - c j^{-d-1})^2 = 0, \quad (3.118)$$

this implies that  $c_0 j^{-d_0-1} = c j^{-d-1}$  for all  $j \geq 1$  and  $c_0, c$  are positive constants. Then, the Equation (3.118) holds almost surely if and only if  $c = c_0$  and  $d = d_0$ . Then, substituting  $c = c_0$  and  $d = d_0$  into the Equation (3.117), it is clear that  $\omega_0 = \omega$  almost surely. This completes the proof. ■

### 3.8.11 Proof of Lemma 3.6.3

**Proof of Lemma 3.6.3.** Since:

$$\frac{\partial l_{0t}}{\partial \vartheta} = \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial \ln(h_{0t})}{\partial \vartheta},$$

then:

$$E \left[ \frac{\partial l_{0t}}{\partial \vartheta} \middle| \mathcal{F}_{t-1} \right] = E \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial \ln(h_{0t})}{\partial \vartheta} \middle| \mathcal{F}_{t-1} \right] = 0,$$

and:

$$E \left[ \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right] = E \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \frac{\partial \ln(h_{0t})}{\partial \vartheta} \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \right] < \infty,$$

if  $E[z_{0t}]^4 < \infty$  and  $E \left[ \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \right]^2 < 1$  by Lemma 3.6.4.

Then:

$$\begin{aligned} & C' E \left[ \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right] C \\ &= E \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 C' \frac{\partial \ln(h_{0t})}{\partial \vartheta} \frac{\partial \ln(h_{0t})}{\partial \vartheta'} C \right] \\ &= E \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \left( C' \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right)^2 \right] \\ &> 0, \end{aligned}$$

where  $C$  is with the same dimension as  $\vartheta$ , and:

$$\begin{aligned} \frac{\partial \ln(h_{0t})}{\partial \vartheta} &= \begin{pmatrix} 1 \\ \sum_{j=1}^{\infty} c_0 j^{-d_0-1} z_{0t-j} \\ \sum_{j=1}^{\infty} j^{-d_0-1} g_1 \left( \frac{\xi_{t-j}}{\sqrt{h_{0t-j}}} \right) \\ - \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j \cdot g_1 \left( \frac{\xi_{t-j}}{\sqrt{h_{0t-j}}} \right) \end{pmatrix} \\ &+ \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \frac{\partial \ln h_{0t-j}}{\partial \vartheta}. \end{aligned} \quad (3.119)$$

If assume that:

$$\left( C \frac{\partial \ln(h_t)}{\partial \vartheta'} \right)^2 = 0,$$

almost surely, and by the stationarity of  $\partial \ln(h_{0t}) / \partial \vartheta$ , then:

$$C \begin{pmatrix} 1 \\ \sum_{j=1}^{\infty} c_0 j^{-d_0-1} z_{0t-j} \\ \sum_{j=1}^{\infty} j^{-d_0-1} g \left( \frac{\xi_{t-j}}{\sqrt{h_{0t-j}}} \right) \\ - \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j \cdot g \left( \frac{\xi_{t-j}}{\sqrt{h_{0t-j}}} \right) \end{pmatrix} = 0.$$

That is:

$$c_1 + c_2 \sum_{j=1}^{\infty} c_0 j^{-d_0-1} z_{0t-j} + c_3 \sum_{j=1}^{\infty} j^{-d_0-1} g(z_{0t-j}) - c_4 \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j g(z_{0t-j}) = 0,$$

which means that:

$$\begin{aligned} & c_1 + \left( c_2 \sum_{j=1}^{\infty} c_0 j^{-d_0-1} + \theta c_3 \sum_{j=1}^{\infty} j^{-d_0-1} - \theta c_4 \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j \right) z_{0t-j} \\ & + \left( c_3 \sum_{j=2}^{\infty} j^{-d_0-1} - c_4 \sum_{j=2}^{\infty} c_0 j^{-d_0-1} \ln j \right) |z_{0t-j}| \\ & = -c_3 |z_{0t-1}|. \end{aligned}$$

Then, similar to Straumann (2005) Section 5.4.1, since  $z_t$  with mean 0 and variance 1 and its distribution is not concentrated in two points, it is able to obtain that the left hand side of the above equation is independent to the right hand of the equation, thus  $c_3 = 0$  almost surely; and:

$$\begin{aligned} & c_1 + \left( c_2 \sum_{j=1}^{\infty} c_0 j^{-d_0-1} - \theta c_4 \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j \right) z_{0t-j} \\ & + \left( -c_4 \sum_{j=2}^{\infty} c_0 j^{-d_0-1} \ln j \right) |z_{0t-j}| \\ & = 0, \end{aligned}$$

almost surely if and only if:

$$c_1 = 0, \quad c_4 = 0, \quad \text{and then } c_2 = 0.$$

Thus,  $C = 0$ . Therefore, similar to the Theorem 6.3 of Schützner (2009), by the Theorem 23.1 in Billingsley (1968)<sup>9</sup> and the Cramer-Wold device<sup>10</sup>, we can get:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial l_{0t}}{\partial \vartheta} \rightarrow^d N(0, \Omega_{01}). \quad (3.120)$$

<sup>9</sup>The theorem 23.1 in Billingsley (1968): Let  $\{\varepsilon_1, \varepsilon_2, \dots\}$  be a stationary, ergodic process for which

$$E\{\varepsilon_n | \varepsilon_1, \varepsilon_2, \dots, \varepsilon_{n-1}\} = 0$$

with probability 1 and for which  $E[\varepsilon_n^2] = \sigma^2$  is positive and finite. If  $X_n(t, \omega) = S_{[n,t]}(\omega)/\sigma\sqrt{n}$ , then  $X_n \rightarrow^d W$ .

<sup>10</sup>Cramer-Wold device: Suppose  $\{x_n\}_{n=1}^{\infty}$  is a sequence of random  $K$ -vectors that satisfies  $c'Y_n \rightarrow^d c'Y$  as  $n \rightarrow \infty$  for all  $c \in R^k$ . Then  $Y_n \rightarrow^d Y$ .

This completes the proof. ■

### 3.8.12 Proof of Lemma 3.6.4

**Proof of Lemma 3.6.4.** Since:

$$\frac{\partial^2 l_t(\vartheta)}{\partial \vartheta \partial \vartheta'} = \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial^2 \ln(h_t)}{\partial \vartheta \partial \vartheta'} - \frac{\xi_t^2}{h_t} \frac{\partial \ln(h_t)}{\partial \vartheta'} \frac{\partial \ln(h_t)}{\partial \vartheta}$$

Taking expectation value on both sides of the above equation on true parameters:

$$\begin{aligned} E \left[ \frac{\partial^2 l_{0t}(\vartheta)}{\partial \vartheta \partial \vartheta'} \right] &= E \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial^2 \ln(h_{0t})}{\partial \vartheta \partial \vartheta'} - \frac{\xi_t^2}{h_{0t}} \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right] \\ &= E \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial^2 \ln(h_{0t})}{\partial \vartheta \partial \vartheta'} \right] - E \left[ \frac{\xi_t^2}{h_{0t}} \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right] \\ &= -E \left[ \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right], \end{aligned}$$

where:

$$E \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial^2 \ln(h_{0t})}{\partial \vartheta \partial \vartheta'} \right] = 0.$$

Thus, in order to prove the existence of  $E[\partial^2 l_{0t}(\vartheta) / (\partial \vartheta \partial \vartheta')]$ , it is essential to show the existence of  $E[(\partial \ln(h_{0t}) / \partial \vartheta') (\partial \ln(h_{0t}) / \partial \vartheta)]$ . Where:

$$\begin{aligned} \frac{\partial \ln(h_{0t})}{\partial \vartheta} &= \begin{pmatrix} 1 \\ \sum_{j=1}^{\infty} c_0 j^{-d_0-1} z_{0t-j} \\ \sum_{j=1}^{\infty} j^{-d_0-1} g_1(z_{0t-j}) \\ - \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j \cdot g_1(z_{0t-j}) \end{pmatrix} \\ &\quad + \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \frac{\partial \ln h_{0t-j}}{\partial \vartheta}. \end{aligned}$$

Then,

$$E \left\| \frac{\partial \ln h_{0t}}{\partial \vartheta} \right\|^2 \leq E \left\| \begin{pmatrix} 1 \\ \sum_{j=1}^{\infty} c_0 j^{-d_0-1} z_{0t-j} \\ \sum_{j=1}^{\infty} j^{-d_0-1} g_1(z_{0t-j}) \\ - \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j \cdot g_1(z_{0t-j}) \end{pmatrix} \right\|^2 \\ + E \left\| \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \frac{\partial \ln h_{0t-j}}{\partial \vartheta} \right\|^2$$

Thus, by the stationarity of  $\partial \ln h_{0t} / \partial \vartheta$ , and if  $E [g_1(z_{0t-j})^2] < \infty$ , and:

$$E \left| \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta_0 z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \right|^2 < 1,$$

then  $E \|\partial \ln h_{0t} / \partial \vartheta\|^2 < \infty$ , and  $E [\partial^2 l_{0t}(\vartheta) / \partial \vartheta \partial \vartheta'] < \infty$ . Thus, by the uniform convergence theorem and the  $\vartheta_n$  is a consistent estimator of  $\vartheta_0$ , the following result

$$\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \vartheta \partial \vartheta'} \Big|_{\vartheta^\dagger} - E \left[ \frac{\partial^2 l_{0t}}{\partial \vartheta \partial \vartheta'} \right] = o_p(1)$$

holds. This completes the proof. ■

### 3.8.13 Proof of Lemma 3.6.5

**Proof of Lemma 3.6.5.** Similar procedure to the proof of Lemma 3.5.6, since:

$$\frac{\partial l_t}{\partial \vartheta} \frac{\partial l_t}{\partial \vartheta'} = \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 \frac{\partial \ln(h_t)}{\partial \vartheta} \frac{\partial \ln(h_t)}{\partial \vartheta'},$$

then:

$$\begin{aligned}
& \left\| \frac{\partial l_t}{\partial \vartheta} \frac{\partial l_t}{\partial \vartheta'} - \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right\| \\
& \leq \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \left\| \frac{\partial \ln(h_t)}{\partial \vartheta} \frac{\partial \ln(h_t)}{\partial \vartheta'} - \frac{\partial \ln(h_{0t})}{\partial \vartheta} \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \right\| \\
& \quad + \left\| \left( \frac{\xi_t^2}{h_t} - 1 \right)^2 - \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \right\| \left\| \frac{\partial \ln(h_t)}{\partial \vartheta} \frac{\partial \ln(h_t)}{\partial \vartheta'} \right\| \\
& \leq \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \left\| \frac{\partial \ln(h_t)}{\partial \vartheta} - \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right\|^2 + 2 \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \left\| \frac{\partial \ln(h_t)}{\partial \vartheta} - \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right\| \\
& \quad \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \right\| + \|o_p(1) + o_p(1)\| |z_{0t}|^4 \left\| \frac{\partial \ln h_t}{\partial \vartheta} \right\|^2 \\
& \leq \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \frac{1}{\sqrt{n}} \|\sqrt{n}(\vartheta_n - \vartheta_0)\|^2 \left\| \frac{\partial^2 \ln h_t}{\partial \vartheta \partial \vartheta'} \right\|^2 + 2 \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 \frac{1}{\sqrt{n}} \|\sqrt{n}(\vartheta_n - \vartheta_0)\| \\
& \quad \left\| \frac{\partial^2 \ln h_t}{\partial \vartheta \partial \vartheta'} \right\| \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \right\| + \|o_p(1) + o_p(1)\| |z_{0t}|^4 \left\| \frac{\partial \ln h_t}{\partial \vartheta} \right\|^2 \\
& = \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 o_p(1) + 2 \left( \frac{\xi_t^2}{h_{0t}} - 1 \right)^2 o_p(1) \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \right\| + \|o_p(1) + o_p(1)\| |z_{0t}|^4 \left\| \frac{\partial \ln h_t}{\partial \vartheta} \right\|^2.
\end{aligned}$$

Thus:

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_t}{\partial \vartheta} \frac{\partial l_t}{\partial \vartheta'} - \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right]_{\vartheta_n} = o_p(1),$$

and since:

$$\frac{\partial l_{0t}}{\partial \vartheta} = \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial \ln(h_{0t})}{\partial \vartheta},$$

then:

$$E \left\| \frac{\partial l_{0t}}{\partial \vartheta} \right\|^2 = E \left| \frac{\xi_t^2}{h_{0t}} - 1 \right|^2 E \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right\|^2.$$

Suppose  $E[z_{0t}]^4 < \infty$ , and since  $E \|\partial \ln h_{0t} / \partial \vartheta\|^2 < \infty$ , thus by the uniform ergodic

theorem:

$$\frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right]_{\vartheta_n} - E \left[ \frac{\partial l_{0t}}{\partial \vartheta} \frac{\partial l_{0t}}{\partial \vartheta'} \right] = o_p(1).$$

This completes the proof. ■

### 3.8.14 Proof of Lemma 3.6.6

**Proof of Lemma 3.6.6.** The difference between  $L_n(\vartheta)$  and  $\tilde{L}_n(\vartheta)$  in the HYE-GARCH is as follows:

$$\begin{aligned}
& \sup_{\vartheta \in \Theta} \left| L_n(\vartheta) - \tilde{L}_n(\vartheta) \right| \\
& \leq \frac{1}{n} \sum_{t=1}^n |\xi_t^2| \sup_{\vartheta \in \Theta} \left| \frac{1}{h_t(\vartheta)} - \frac{1}{\tilde{h}_t(\vartheta)} \right| \\
& \quad + \frac{1}{n} \sum_{t=1}^n \sup_{\vartheta \in \Theta} \left| \ln(h_t(\vartheta)) - \ln(\tilde{h}_t(\vartheta)) \right| \\
& \leq C \left( \frac{1}{n} \sum_{t=1}^n |\xi_t^2| \sup_{\vartheta \in \Theta} \left| \ln h_t(\vartheta) - \ln \tilde{h}_t(\vartheta) \right| \right) + \\
& \quad \frac{1}{n} \sum_{t=1}^n \sup_{\vartheta \in \Theta} \left| \ln(h_t(\vartheta)) - \ln(\tilde{h}_t(\vartheta)) \right|.
\end{aligned}$$

The first inequality follows direct from the triangle inequality, and the second inequality is obtained by mean value theorem. Then:

$$\begin{aligned}
& E \left[ \sup_{\vartheta \in \Theta} \left| L_n(\vartheta) - \tilde{L}_n(\vartheta) \right| \right] \\
& \leq C \frac{1}{n} \sum_{t=1}^n \left( E \left[ |\xi_t^2| \sup_{\vartheta \in \Theta} \left| \ln h_t(\vartheta) - \ln \tilde{h}_t(\vartheta) \right| \right] \right) + \\
& \quad C \frac{1}{n} \sum_{t=1}^n E \left[ \sup_{\vartheta \in \Theta} \left| \ln h_t(\vartheta) - \ln \tilde{h}_t(\vartheta) \right| \right] \\
& \leq C \frac{1}{n} \sum_{t=1}^n \left( \left[ E |\xi_t^2|^2 \right]^{\frac{1}{2}} \left[ \sup_{\vartheta \in \Theta} E \left| \ln h_t(\vartheta) - \ln \tilde{h}_t(\vartheta) \right|^2 \right]^{\frac{1}{2}} \right) + \\
& \quad C \frac{1}{n} \sum_{t=1}^n E \left[ \sup_{\vartheta \in \Theta} \left| \ln h_t(\vartheta) - \ln \tilde{h}_t(\vartheta) \right| \right]
\end{aligned}$$

Since:

$$\left| \ln h_t(\vartheta) - \ln \tilde{h}_t(\vartheta) \right| = O(n^{-d}),$$

then:

$$\frac{1}{n} \sum_{t=1}^n \left[ \sup_{\vartheta \in \Theta} E \left| \ln h_t(\vartheta) - \ln \tilde{h}_t(\vartheta) \right|^2 \right]^{\frac{1}{2}} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

And  $E[\xi_t]^4 < \infty$  if  $E[z_t]^4 < \infty$  since  $E[h_t]^2$  can be finite by Section 2.4, finally:

$$E \left[ \sup_{\vartheta \in \Theta} \left| L_n(\vartheta) - \tilde{L}_n(\vartheta) \right| \right] \rightarrow 0 \text{ as } n \rightarrow \infty.$$

This completes the proof. ■

### 3.8.15 Proof of Theorem 3.6.2

**Proof of Theorem 3.6.2.** Similar to Theorem 3.5.2, the Lemmas 3.6.1 and 3.6.2 show uniform convergence of the  $L_n(\vartheta)$  to the function  $\bar{L}(\vartheta)$  in probability. Moreover,  $\bar{L}(\vartheta)$  obtain a unique maximum at  $\vartheta_0$ . And Lemma 3.6.6 shows the difference between  $L_n(\vartheta)$  and  $\tilde{L}_n(\vartheta)$  converges to zero in probability. Then since:

$$\begin{aligned} & \sup_{\vartheta \in \Theta} |\tilde{L}_n(\vartheta) - \bar{L}(\vartheta)| && (3.121) \\ &= \sup_{\vartheta \in \Theta} |\tilde{L}_n(\vartheta) + L_n(\vartheta) - L_n(\vartheta) - \bar{L}(\vartheta)| \\ &\leq \sup_{\vartheta \in \Theta} |L_n(\vartheta) - \bar{L}(\vartheta)| + \sup_{\vartheta \in \Theta} |L_n(\vartheta) - \tilde{L}_n(\vartheta)| \\ &= o_p(1), \end{aligned}$$

thus, the  $\hat{\vartheta}_n$  is a consistent estimator of  $\vartheta_0$ , as  $n \rightarrow \infty$ . This completes the proof.

■



### 3.8.16 Proof of Lemma 3.6.7

**Proof of Lemma 3.6.7.** Since:

$$\begin{aligned}
& \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right] \\
&= \frac{1}{n} \sum_{t=1}^n \left[ \left( \frac{\xi_t^2}{h_{0t}} - 1 \right) \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \left( \frac{\xi_t^2}{\tilde{h}_{0t}} - 1 \right) \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right] \\
&= \frac{1}{n} \sum_{t=1}^n \left[ \left( \frac{\xi_t^2}{\tilde{h}_{0t}} - 1 \right) \left( \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right) \right] \\
&\quad + \frac{1}{n} \sum_{t=1}^n \left[ \xi_t^2 \left( \frac{1}{h_{0t}} - \frac{1}{\tilde{h}_{0t}} \right) \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right],
\end{aligned}$$

then:

$$\begin{aligned}
& \sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right] \right\| \\
&\leq \frac{1}{n} \sum_{t=1}^n \left| \frac{\xi_t^2}{\tilde{h}_{0t}} - 1 \right| \sup_{\vartheta \in \Theta} \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\| \\
&\quad + \frac{C}{n} \sum_{t=1}^n \xi_t^2 \sup_{\vartheta \in \Theta} \left| \ln h_{0t} - \ln \tilde{h}_{0t} \right| \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right\|,
\end{aligned}$$

and:

$$\begin{aligned}
& E \left[ \sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right] \right\| \right] \\
&\leq \frac{1}{n} \sum_{t=1}^n \left[ E \left| \frac{\xi_t^2}{\tilde{h}_{0t}} - 1 \right|^2 \right]^{\frac{1}{2}} \left[ E \sup_{\vartheta \in \Theta} \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\|^2 \right]^{\frac{1}{2}} \\
&\quad + \frac{C}{n} \sum_{t=1}^n \left[ E |\xi_t^2|^2 \right]^{\frac{1}{2}} E \left[ \sup_{\vartheta \in \Theta} \left| \ln h_{0t} - \ln \tilde{h}_{0t} \right|^4 \right]^{\frac{1}{4}} E \left[ \sup_{\vartheta \in \Theta} \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right\|^4 \right]^{\frac{1}{4}}.
\end{aligned}$$

by the invertibility property of HYGARCH process and Lemma 3.6.6, we can

get:

$$E \left[ \sup_{\vartheta \in \Theta} \left| \ln h_{0t} - \ln \tilde{h}_{0t} \right|^4 \right]^{\frac{1}{4}} = o_p(1),$$

and since:

$$\frac{\partial \ln(h_{0t})}{\partial \vartheta} = \begin{pmatrix} 1 \\ \sum_{j=1}^{\infty} c_0 j^{-d_0-1} z_{0t-j} \\ \sum_{j=1}^{\infty} j^{-d_0-1} g_1 \left( \frac{\xi_{t-j}}{\sqrt{h_{0t-j}}} \right) \\ - \sum_{j=1}^{\infty} c_0 j^{-d_0-1} \ln j \cdot g_1 \left( \frac{\xi_{t-j}}{\sqrt{h_{0t-j}}} \right) \end{pmatrix} + \sum_{j=1}^{\infty} b_{0j} \left( -\frac{1}{2} \theta z_{0t-j} - \frac{1}{2} |z_{0t-j}| \right) \frac{\partial \ln h_{0t-j}}{\partial \vartheta}, \quad (3.122)$$

and:

$$\frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} = \begin{pmatrix} 1 \\ \sum_{j=1}^{n-1} c_0 j^{-d_0-1} \tilde{z}_{0t-j} \\ \sum_{j=1}^{n-1} j^{-d_0-1} g_1 \left( \frac{\xi_{t-j}}{\sqrt{\tilde{h}_{0t-j}}} \right) \\ - \sum_{j=1}^{n-1} c_0 j^{-d_0-1} \ln j \cdot g_1 \left( \frac{\xi_{t-j}}{\sqrt{\tilde{h}_{0t-j}}} \right) \end{pmatrix} + \sum_{j=1}^{n-1} b_{0j} \left( -\frac{1}{2} \theta \tilde{z}_{0t-j} - \frac{1}{2} |\tilde{z}_{0t-j}| \right) \frac{\partial \ln \tilde{h}_{0t-j}}{\partial \vartheta}, \quad (3.123)$$

and  $\tilde{z}_{0t}$  converges to  $z_{0t}$  and  $\tilde{h}_{0t}$  converges to  $h_{0t}$  as  $n$  tends to infinity. Thus:

$$\left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\| = o_p(1),$$

and then:

$$\left[ E \sup_{\vartheta \in \Theta} \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta} - \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} \right\|^2 \right]^{\frac{1}{2}} = o_p(1),$$

and suppose  $E[z_{0t}]^4 < \infty$  and  $E \|\partial \ln(h_{0t}) / \partial \vartheta\|^4 < \infty$ , then:

$$E \left[ \sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial l_{0t}}{\partial \vartheta} - \frac{\partial \tilde{l}_{0t}}{\partial \vartheta} \right] \right\| \right] = o_p(1),$$

as  $n$  tends to infinity. This completes the proof. ■

### 3.8.17 Proof of Lemma 3.6.8

**Proof of Lemma 3.6.8.** Since:

$$\begin{aligned} \frac{\partial^2 l_t(\vartheta)}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 \tilde{l}_t(\vartheta)}{\partial \vartheta \partial \vartheta'} &= \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial^2 \ln(h_t)}{\partial \vartheta \partial \vartheta'} - \left( \frac{\xi_t^2}{\tilde{h}_t} - 1 \right) \frac{\partial^2 \ln(\tilde{h}_t)}{\partial \vartheta \partial \vartheta'} \\ &\quad + \frac{\xi_t^2}{\tilde{h}_t} \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta'} \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta} - \frac{\xi_t^2}{h_t} \frac{\partial \ln(h_t)}{\partial \vartheta'} \frac{\partial \ln(h_t)}{\partial \vartheta}, \end{aligned}$$

set:

$$\begin{aligned} L_{21} &= \left( \frac{\xi_t^2}{h_t} - 1 \right) \frac{\partial^2 \ln(h_t)}{\partial \vartheta \partial \vartheta'} - \left( \frac{\xi_t^2}{\tilde{h}_t} - 1 \right) \frac{\partial^2 \ln(\tilde{h}_t)}{\partial \vartheta \partial \vartheta'} \\ &= \left( \frac{\xi_t^2}{\tilde{h}_t} - 1 \right) \left( \frac{\partial^2 \ln(h_t)}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 \ln(\tilde{h}_t)}{\partial \vartheta \partial \vartheta'} \right) + \left( \frac{\xi_t^2}{h_t} - \frac{\xi_t^2}{\tilde{h}_t} \right) \frac{\partial^2 \ln(h_t)}{\partial \vartheta \partial \vartheta'}, \end{aligned}$$

and:

$$\begin{aligned} L_{22} &= \frac{\xi_t^2}{\tilde{h}_t} \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta'} \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta} - \frac{\xi_t^2}{h_t} \frac{\partial \ln(h_t)}{\partial \vartheta'} \frac{\partial \ln(h_t)}{\partial \vartheta} \\ &= \frac{\xi_t^2}{\tilde{h}_t} \left( \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta} - \frac{\partial \ln(h_t)}{\partial \vartheta} \right) \left( \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta'} + \frac{\partial \ln(h_t)}{\partial \vartheta'} \right) \\ &\quad + \left( \frac{\xi_t^2}{\tilde{h}_t} - \frac{\xi_t^2}{h_t} \right) \frac{\partial \ln(h_t)}{\partial \vartheta'} \frac{\partial \ln(h_t)}{\partial \vartheta}. \end{aligned}$$

Then:

$$\begin{aligned} \sup_{\vartheta \in \Theta} \frac{1}{n} \sum_{t=1}^n \|L_{21}\| &\leq \frac{1}{n} \sum_{t=1}^n \left| \frac{\xi_t^2}{\tilde{h}_t} - 1 \right| \sup_{\vartheta \in \Theta} \left\| \frac{\partial^2 \ln(h_t)}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 \ln(\tilde{h}_t)}{\partial \vartheta \partial \vartheta'} \right\| \\ &\quad + \frac{1}{n} \sum_{t=1}^n \sup_{\vartheta \in \Theta} \left| \frac{\xi_t^2}{h_t} - \frac{\xi_t^2}{\tilde{h}_t} \right| \left\| \frac{\partial^2 \ln(h_t)}{\partial \vartheta \partial \vartheta'} \right\|, \end{aligned}$$

and:

$$\begin{aligned}
& \sup_{\vartheta \in \Theta} \frac{1}{n} \sum_{t=1}^n \|L_{22}\| \\
\leq & \frac{1}{n} \sum_{t=1}^n \frac{\xi_t^2}{\tilde{h}_t} \sup_{\vartheta \in \Theta} \left\| \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta} - \frac{\partial \ln(h_t)}{\partial \vartheta} \right\| \left\| \frac{\partial \ln(\tilde{h}_t)}{\partial \vartheta'} + \frac{\partial \ln(h_t)}{\partial \vartheta'} \right\| \\
& + \frac{1}{n} \sum_{t=1}^n \sup_{\vartheta \in \Theta} \left| \frac{\xi_t^2}{\tilde{h}_t} - \frac{\xi_t^2}{h_t} \right| \left\| \frac{\partial \ln(h_t)}{\partial \vartheta'} \right\|^2.
\end{aligned}$$

Thus:

$$\begin{aligned}
& E \sup_{\vartheta \in \Theta} \frac{1}{n} \sum_{t=1}^n \|L_{22}\|_{\vartheta_0} \\
\leq & \frac{1}{n} \sum_{t=1}^n \left[ E \left( \frac{\xi_t^2}{\tilde{h}_{0t}} \right)^2 \right]^{\frac{1}{2}} \left[ E \sup_{\vartheta \in \Theta} \left\| \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} - \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right\|^4 \right]^{\frac{1}{4}} \\
& \left[ E \left\| \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta'} + \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \right\|^4 \right]^{\frac{1}{4}} \\
& + \frac{1}{n} \sum_{t=1}^n \left[ E \sup_{\vartheta \in \Theta} \left| \frac{\xi_t^2}{\tilde{h}_{0t}} - \frac{\xi_t^2}{h_{0t}} \right|^2 \right]^{\frac{1}{2}} \left[ E \left\| \frac{\partial \ln(h_{0t})}{\partial \vartheta'} \right\|^4 \right]^{\frac{1}{2}}.
\end{aligned}$$

As from Lemma 3.6.7:

$$\sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n E \left[ \frac{\partial \ln(\tilde{h}_{0t})}{\partial \vartheta} - \frac{\partial \ln(h_{0t})}{\partial \vartheta} \right] \right\| = o_p(1), \quad (3.124)$$

and suppose  $E[z_{0t}]^4 < \infty$  and  $E \|\partial \ln(h_t) / \partial \vartheta'\|^4 < \infty$ , and from Lemma 3.6.6:

$$\sup_{\lambda \in \Theta} \frac{1}{n} \sum_{t=1}^n E \left| \frac{\xi_t^2}{h_{0t}} - \frac{\xi_t^2}{\tilde{h}_{0t}} \right| = o_p(1), \quad (3.125)$$

then:

$$E \sup_{\vartheta \in \Theta} \frac{1}{n} \sum_{t=1}^n \|L_{22}\|_{\vartheta_0} \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

And for:

$$\begin{aligned}
& E \sup_{\vartheta \in \Theta} \frac{1}{n} \sum_{t=1}^n \|L_{21}\|_{\vartheta_0} \\
& \leq \frac{1}{n} \sum_{t=1}^n \left[ E \left| \frac{\xi_t^2}{\tilde{h}_{0t}} - 1 \right|^2 \right]^{\frac{1}{2}} \left[ E \sup_{\vartheta \in \Theta} \left\| \frac{\partial^2 \ln(h_{0t})}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 \ln(\tilde{h}_{0t})}{\partial \vartheta \partial \vartheta'} \right\|^2 \right]^{\frac{1}{2}},
\end{aligned}$$

suppose  $E[z_{0t}^4] < \infty$  and similar to Lemma 3.6.7, we can obtain:

$$\left[ E \sup_{\vartheta \in \Theta} \left\| \frac{\partial^2 \ln(h_{0t})}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 \ln(\tilde{h}_{0t})}{\partial \vartheta \partial \vartheta'} \right\|^2 \right]^{\frac{1}{2}} = o_p(1), \quad (3.126)$$

Thus:

$$E \sup_{\vartheta \in \Theta} \frac{1}{n} \sum_{t=1}^n \|L_{21}\|_{\vartheta_0} \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

By the Theorem 3.6.2 and Lemma 3.6.4:

$$\sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \vartheta \partial \vartheta'} - E \left[ \frac{\partial^2 l_{0t}}{\partial \vartheta \partial \vartheta'} \right] \right\| = o_p(1).$$

Then:

$$\sup_{\vartheta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^n \frac{\partial^2 \tilde{l}_t}{\partial \vartheta \partial \vartheta'} \Big|_{\tilde{\vartheta}^\dagger} - E \left[ \frac{\partial^2 l_{0t}}{\partial \vartheta \partial \vartheta'} \right] \right\| = o_p(1).$$

This completes the proof. ■

**STRICT STATIONARITY, PERSISTENCE AND VOLATILITY  
FORECASTING IN ARCH( $\infty$ ) PROCESSES**

### 4.1 Introduction

As discussed in the previous chapters, the ARCH, GARCH and related classes of volatility models are extensively employed to exploit the fact of local persistence in the volatility of returns processes, so as to predict volatility a number of steps into the future. Notwithstanding the large volume of research that has been devoted to understanding these models since their inception, there remains a degree of mystery surrounding their dynamic properties, and hence the degree to which they assist the effective forecasting of further volatility. Analogies drawn from the theory of linear processes in levels have sometimes been invoked inappropriately in attempts to explain their behaviour, as has been detailed in Ding and Granger (1996), Davidson (2004), Giraitis et al. (2009) among other commentaries.

This chapter considers the ARCH( $\infty$ ) model, which was introduced by Robinson (1991), of an uncorrelated returns sequence  $\{\xi_t\}$  in which, for  $-\infty < t < +\infty$ ,

$$\xi_t = \sqrt{h_t} z_t, \quad (4.1)$$

and:

$$h_t = \omega + \sum_{j=1}^{\infty} \theta_j \xi_{t-j}^2, \quad (4.2)$$

where  $z_t \sim i.i.d.(0, 1)$ ,  $\omega > 0$ ,  $\theta_j \geq 0$  for all  $j$  and  $S = \sum_{j=1}^{\infty} \theta_j < \infty$ . The stationarity and persistence of the ARCH( $\infty$ ) have been following developed by Kokoszka and Leipus (2000), Giraitis et al. (2000a), Giraitis and Surgailis (2002), Giraitis et al. (2007), among others. Kokoszka and Leipus (2000) proposed a

strictly and weakly stationary solution to the ARCH( $\infty$ ). Giraitis et al. (2000a) also studied the existence of a stationary solution to the ARCH( $\infty$ ) and emphasised that conditions on the lag coefficients and sequence  $\{z_t\}$  are significant for the existence of stationarity. They improved the results of Kokoszka and Leipus (2000) and showed the uniqueness of the stationary solution. Giraitis and Surgailis (2002) investigated the fourth moment conditions of the ARCH( $\infty$ ) model.

However, the strict stationarity in covariance nonstationary ARCH( $\infty$ ) has not been fully explored. The question of strict stationarity in covariance nonstationary processes was first examined by Nelson (1990) who derived the necessary and sufficient condition for strict stationarity in the GARCH(1,1) model. Bougerol and Picard (1992) extended the results of Nelson to GARCH( $p, q$ ) models, and pointed out the importance of the negativity of the top Lyapunov exponent of a certain sequence of random matrices. Kazakevičius and Leipus (2002) showed that a necessary condition for a stationary solution in the ARCH( $\infty$ ) class models and Douc et al. (2008) provided a sufficient condition. Motivated by the above results, this chapter focuses on the strict stationarity in covariance nonstationary ARCH( $\infty$ ) processes.

This chapter also considers the wider question of the persistence of stationary volatility processes, and proposes a new persistence measure. This notion of persistence, which is independent of the existence of moments, is made precise in Section 4.4, where we define it in terms of the (in)frequency of crossings of the median in successive steps. Thus, a process which crosses the median at most a finite number of times in a realisation of length  $T$ , as  $T \rightarrow \infty$ , is necessarily nonstationary, either converging or diverging. At the other extreme, a serially independent process crosses the median with probability  $1/2$  at each step, by construction.

Conditions for strict stationarity of a process in effect define the boundary beyond which persistence becomes divergence, and there is no reversion tendency defining a stationary distribution.

Moreover, once we know how long episodes of high volatility tend to persist, it is worth investigating how far into the future variations in volatility may feasibly be forecast. For the volatility forecasting, the squared of the return processes is normally replaced by its (assumed) conditional expectation. This might not be very appropriate, thus an alternative forecasting method is proposed in the Section 4.6.

To investigate these questions, it is vital to focus on the three salient features of models of this type: the value of the sum of the lag coefficients  $S$ ; the decay rate of the lag coefficients  $\theta_j$ ; and the distribution of the underlying process  $z_t$ . Having regard to the first of these features, it is well known that  $S < 1$  is a necessary condition for covariance stationarity (see e.g. Kokoszka and Leipus, 2000; Giraitis et al., 2000a). Unless this condition applies it is inappropriate to speak of  $h_t$  as the conditional variance although it is always well-defined as a volatility indicator. With regard to the second feature, it is also well known that the Bollerslev (1986) GARCH class of models imposes exponential decay rates on the coefficients, and the HYGARCH class due to Davidson (2004) which includes the FIGARCH model of Baillie et al. (1996), embodies hyperbolic decay rates. Considering the third feature, the underlying processes are often specified to be Gaussian, even though it is a well-known stylised fact that the residuals from estimated GARCH models in financial data can exhibit excess kurtosis.

The structure of the rest of chapter is as follows. Section 4.2 briefly reviews the covariance stationary, strictly stationary and persistence of the relevant ARCH-



type volatility models. In Section 4.3, a decomposition of the ARCH( $\infty$ ) equation is introduced which simplifies the problem of seeing how persistence and stationarity depends on the various model features. We use this representation to derive a new sufficient condition for strict stationarity. In the GARCH(1, 1) case where the stationarity boundary in the parameter space is known, we show numerically that our condition is not too far from necessity, in contrast to a strong condition such as (4.16). The properties of these models are shown to be the result of rather complex interactions between the shock distribution and the linear structure. Section 4.5 reports a comprehensive set of simulations, covering the persistence, covariance stationary, strictly stationary and nonstationary cases of the ARCH-type models. Section 4.6 evaluates the performance of the volatility forecasting, considers the implications of our analysis for the optimal forecasting of volatility, and investigates alternatives to the minimum mean squared error criterion, which is conventional but not necessarily optimal in the context of highly skewed volatility processes. Section 4.7 contains concluding remarks, and proofs of the propositions stated in Section 4.3 are gathered in the Appendix C. All simulations results are obtained by using software OxMetrics and TSM.

## 4.2 Literature review

### 4.2.1 The stationarity of the ARCH-type models

Many scholars have investigated the stationarity of ARCH-type processes. These results have been reviewed recently by, for example, Giraitis et al. (2007), Giraitis et al. (2009), and Beran et al. (2013). This section briefly reviews the literature on covariance stationarity and strict stationarity for ARCH-type models.

#### Second and fourth-moment conditions of ARCH-type models

The stationarity properties of the ARCH( $\infty$ ) model under second-moment and fourth-moment conditions have been investigated by many researchers, including Kokoszka and Leipus (2000), Giraitis et al. (2000a), and Giraitis and Surgailis (2002). With regard to the covariance stationarity of the ARCH( $\infty$ ) process, Kokoszka and Leipus (2000) showed that if  $(Ez_1^4)^{1/2} \sum_{j=1}^{\infty} \theta_j < 1$ , then  $\xi_t^2$  is a strict and weak stationary solution to the non-negative ARCH( $\infty$ ) processes. Giraitis et al. (2000a) also investigated the existence of a stationary solution to non-negative ARCH( $\infty$ ) models and suggested that the conditions on these coefficients and the distribution of  $\{z_t^2\}$  play a vital role in guaranteeing the stationarity of these models. They assumed that the first and second moments of  $z_t^2$  are existent and

define the non-negative ARCH( $\infty$ ) process as:

$$\begin{aligned}
\xi_t^2 &= \left( \omega + \sum_{j=1}^{\infty} \theta_j \xi_{t-j}^2 \right) z_t^2 & (4.3) \\
&= z_t^2 \omega \left( 1 + \sum_{k=1}^{\infty} \sum_{j_1, j_2, \dots, j_k=1}^{\infty} \theta_{j_1, \dots, j_k} z_{t-j_1}^2 \cdots z_{t-j_1 \cdots -j_k}^2 \right) \\
&= \omega \sum_{k=0}^{\infty} \sum_{j_1, j_2, \dots, j_k=1}^{\infty} \theta_{j_1, \dots, j_k} z_t^2 z_{t-j_1}^2 \cdots z_{t-j_1 \cdots -j_k}^2.
\end{aligned}$$

where  $\omega \geq 0$ ,  $\theta_j \geq 0$ ,  $j = 1, 2, \dots$ . This can be rewritten as:

$$\xi_t^2 = \omega \sum_{k=0}^{\infty} M_k(t), \quad (4.4)$$

with:

$$\begin{aligned}
M_k(t) &= \sum_{j_1, j_2, \dots, j_k=1}^{\infty} \theta_{j_1, \dots, j_k} z_t^2 z_{t-j_1}^2 \cdots z_{t-j_1 \cdots -j_k}^2 & (4.5) \\
&= \sum_{j_k < j_{k-1} < \dots < j_1}^{\infty} \theta_{j_1, \dots, j_k} z_t^2 z_{t-j_1}^2 \cdots z_{t-j_k}^2,
\end{aligned}$$

where  $M_0(t) = z_t^2$  and  $M_k(t)$  for all  $k \geq 1$ . Giraitis et al. (2000a) refined the results of Kokoszka and Leipus (2000) in their Theorem 2.1 and provided a sufficient conditions for the existence of a stationary solution to non-negative ARCH( $\infty$ ) class models using a Volterra series type expansion of the ARCH process. The conditions are:

$$E z_1^2 \sum_{j=1}^{\infty} \theta_j < 1, \quad (4.6)$$

and

$$(E z_1^4)^{1/2} \sum_{j=1}^{\infty} \theta_j < 1, \quad (4.7)$$

with  $E z_1^2 < \infty$  and  $E z_1^4 < \infty$ , respectively. In their Theorem 2.1 Condition (4.7) is a sufficient condition for the existence of a fourth-order stationary solution to the non-negative ARCH( $\infty$ ) processes. It can be seen that Condition (4.7) implies (4.6) because  $E z_1^2 \leq (E z_1^4)^{1/2}$ . This shows that Giraitis et al. (2000a) obtained the

same result as Kokoszka and Leipus (2000) for the covariance stationary solution but improved the strict stationary solution (see Section 4.2.1). Based on Condition (4.7), Giraitis et al. (2007) proposed the following sufficient condition:

$$E(|z_1|^{2p})^{\frac{1}{p}} \sum_{j=1}^{\infty} \theta_j < 1, \quad (4.8)$$

for the existence of the  $E(\xi_t^2)^p$  with  $p \geq 1$ . This condition is the same as (4.7) when  $p = 2$ . However, compared with the results of Nelson (1990), the inequality of (4.7) is not a necessary condition for the existence of a fourth-order stationary solution to the ARCH( $\infty$ ) process.

For the existence of a fourth-order stationary solution to the ARCH( $\infty$ ) process, Giraitis and Surgailis (2002) derived a sufficient and necessary condition by applying an orthogonal Volterra representation of  $\xi_t^2$ . The fourth-moment conditions require that (4.6) holds and:

$$\left( \sum_{j=1}^{\infty} \left( \left( \frac{E(z_1^4)}{E(z_1^2)} \right) g_j \right)^2 \right)^{1/2} < 1, \quad (4.9)$$

where:

$$\sum_{j=1}^{\infty} g_j z^j = \left( 1 - E z_1^2 \sum_{i=1}^{\infty} \theta_i z^i \right)^{-1},$$

which means that:

$$g_j = \sum_{k=1}^j (E z_1^2)^k \sum_{0 < i_1 < \dots < i_{k-1}} \theta_{i_1} \theta_{i_2 - i_1} \dots \theta_{i_{k-2} - i_{k-1}} \quad (j \geq 1), \quad g_0 = 1,$$

need to hold. In comparison, Giraitis et al. (2007) provided a sufficient but not necessary condition for the fourth-order solution by applying Minkowski's norm inequality and pointed out that Kazakevičius et al. (2004) also provided an equivalent condition to the results of Giraitis and Surgailis (2002) by applying a different method.

## Strict stationarity of the covariance nonstationarity in GARCH-type models

Nelson (1990) first examined the strict stationarity in covariance nonstationary processes. He considered the GARCH(1,1) model below:

$$h_t = \gamma + \alpha \xi_{t-1}^2 + \beta h_{t-1}, \quad (4.10)$$

which has the same form as Condition (4.2) with  $\omega = \gamma/(1 - \beta)$  and  $\theta_j = \alpha\beta^{j-1}$  so that  $S = \alpha/(1 - \beta)$ , and showed the necessary and sufficient condition for strict stationarity to be:

$$E \ln(\alpha z_1^2 + \beta) < 0 \quad (4.11)$$

by applying the theory of products of random matrices. Giraitis et al. (2007) reviewed the relevant literature in this area. They emphasised that this condition is weaker than the conditions of Kokoszka and Leipus (2000), Giraitis et al. (2000) and He and Teräsvirta (1999) for the GARCH(1,1) case, and that it implies the existence of the finite second and fourth moments of the return series. Subsequently, Bougerol and Picard (1992) extended Nelson's (1991) results, and provided a necessary and sufficient condition for the existence of a strictly stationary solution to GARCH( $p, q$ ) models, which may have infinite variance. Theorem 1.3 of Bougerol and Picard (1992) shows that the negativity of the associated Lyapunov exponent is a necessary and sufficient condition for strict stationarity in GARCH( $p, q$ ) models; the strictly stationary property of IGARCH models is also included in their Corollary 2.2.

## Strict stationarity in ARCH( $\infty$ ) processes

Kazakevičius and Leipus (2002) generalised the results of Nelson (1990), and

Bougerol and Picard (1992) to ARCH( $\infty$ ) models. They investigated the existence of a strictly stationary solution to ARCH( $\infty$ ) models without moment conditions and established the conditions for the existence of an ARCH( $\infty$ ) process, similar to the results of Bougerol and Picard (1992). In their Corollary 2.3, Bougerol and Picard (1992) proposed that  $\beta(1) < 1$  is a necessary condition for the existence of a GARCH( $p, q$ ) process under a finite second-moment condition of  $z_t$ . Kazakevičius and Leipus (2002) weakened this condition and claimed that the strictly stationary solution to GARCH( $p, q$ ) also satisfies the associated ARCH( $\infty$ ) process in Theorem 2.2 without the moment conditions on  $z_t$ .

In their Theorem 2.3, Kazakevičius and Leipus (2002) also showed that a necessary condition for a stationary solution in the ARCH( $\infty$ ) class is:

$$\ln S < -E \ln(z_1^2), \quad (4.12)$$

where  $S < \infty$  and  $E \ln(z_1^2)$  is well defined. However, this condition is not sufficient. For instance, in the GARCH(1,1) case,  $S = \alpha_1/(1 - \beta_1)$ ; if  $\beta_1 > 0$ , this condition is weaker than that of Nelson (1990). Kazakevičius and Leipus (2002) also applied the convergence radius of the random power series, which is almost surely equal to some non-random constant  $R$ , to show the existence of the ARCH( $\infty$ ) process rather than the top Lyapunov exponent when there is no moment condition placed on  $z_t^2$ . Their Theorem 2.4 proved that the condition  $R > 1$  is sufficient and condition  $R \geq 1$  is necessary for the existence of an ARCH( $\infty$ ) process. They also provided some sufficient conditions for the existence of ARCH( $\infty$ ) processes in their Theorem 2.5.

## Strict stationarity of integrated ARCH ( $\infty$ ) processes without moment conditions

Kazakevičius and Leipus (2003) extended the previous results to the integrated ARCH (IARCH( $\infty$ )) process, which is under the condition:

$$Ez_1^2 \sum_{j=1}^{\infty} \theta_j = 1. \quad (4.13)$$

This is the same as  $\sum_{j=1}^{\infty} \theta_j = 1$  here, since  $Ez_1^2 = 1$ . They applied a Kesten and Spitzer's (1984) method to solve the problem of the existence of IARCH processes, for which they provided conditions in their Theorem 3.2, which is a top Lyapunov exponent  $\gamma < 0$  under the following conditions:

$$E|\ln(z_1^2)| < \infty, \quad (4.14)$$

and:

$$\sum_j \theta_j q^j < \infty \text{ for some } q > 1. \quad (4.15)$$

Condition (4.15) implies that the lag coefficients  $\theta_j$  exhibit exponential decay. They also emphasised that if conditions (4.13) and (4.14) hold, where as (4.15) is not satisfied, then the top Lyapunov exponent  $\gamma = 0$  for the ARCH( $\infty$ ) process in their theorem 3.3. It is obvious that Theorem 3.3 cannot solve the existence problem of the FIGARCH models. Kazakevičius and Leipus (2003) referred to Giraitis et al.'s (2000) analogous condition (4.6) for the ARCH( $\infty$ ) model under the moment conditions on  $\xi_t$ . The strictly stationary solution to the ARCH( $\infty$ ) process becomes more complicated without moment conditions. However, Kazakevičius and Leipus's (2003) results show the stationary conditions for ARCH( $\infty$ ) processes with geometrically decaying lag coefficients, which rules out the power-law decay of the coefficients  $\{\theta_j\}$ . (see also the reviews in Giraitis et al., 2007, Giraitis et al., 2009, Douc et al., 2008).

In their Theorem 1, Douc et al. (2008) provided a new sufficient condition for the existence of a strictly stationary solution to the ARCH( $\infty$ ) model also including the FIGARCH model:

$$E|z_1|^{2p} \sum_{j=1}^{\infty} \theta_j^p < 1 \text{ with } p \in (0, 1]. \quad (4.16)$$

They also investigated stationary IARCH( $\infty$ ) processes in their Corollary 2. It was shown that if  $E[z_1^2] = 1$  and the (4.16) holds for some  $p \in (0, 1]$  if and only if  $p^* < 1$  exists such that  $\sum_{j=1}^{\infty} \theta_j^{p^*} < \infty$  and:

$$\sum_{j=1}^{\infty} \theta_j \ln(\theta_j) + E[z_1^2 \ln(z_1^2)] \in (0, \infty]. \quad (4.17)$$

A solution to the ARCH( $\infty$ ) equation then exists under the conditions  $E[|\xi_t|^q] < \infty$  for all  $q \in [0, 2)$  and  $E[\xi_t^2] = \infty$ .

In contrast to Corollary 2 of Douc et al. (2008) with the conditions in Kazakevičius and Leipus (2003), Douc et al.'s (2008) Corollary 2 derives the stationarity property of IARCH process. And it can be applied to show the causal strictly stationary of some FIGARCH( $p, d, q$ ) models under certain conditions. They proved this result in their Corollary 3, which demonstrated a unique causal stationary solution for FIGARCH( $0, d, 0$ ).

### 4.2.2 Persistence in the ARCH( $\infty$ ) process

Some of the relevant studies have been reviewed in Chapter 2. Here, we briefly review the rest of the relevant literature. Nelson and Cao (1992) proved that the lag coefficients in the ARCH( $\infty$ ) process, which is transformed from the GARCH( $p, q$ ), decay at an exponential rate. The covariance function of the GARCH( $p, q$ ) process presents has exponential decay (see e.g. Bollerslev, 1988; He and Teräsvirta, 1999a).



Kokoszka and Leipus (2000) showed that in the short memory ARCH( $\infty$ ) process, if the lag coefficients  $\theta_j$  decay exponentially, then the covariance function of the squared return series also decays exponentially quickly. Giraitis et al. (2000a), in their Proposition 3.2, showed that the hyperbolic decay of lag coefficients implies the hyperbolic decay of the covariances of the squared of returns. In other words, if  $\theta_j \leq cj^{-r}$ , with  $r > 1$ ,  $c > 0$ , then  $Cov(\xi_k^2, \xi_0^2) \leq ck^{-r}$  for  $k \geq 1$ , where  $c > 0$ ; and if  $c_1j^{-r} \leq \theta_j \leq c_2j^{-r}$  for a large  $j$ , then  $Cov(\xi_k^2, \xi_0^2) \asymp k^{-r}$  for a sufficiently large  $k$ . They emphasised that, although the weakly stationary condition, summability of the lag coefficients, implies summability of the covariances, the autocorrelation function decays very slowly when the chosen  $r$  (fixed) tends to 1. This type memory property of the squared returns may be defined as moderate memory (see e.g. Giraitis et al., 2007). Based on previous literature, Giraitis et al. (2000a) claimed that the rate of decay of the covariance function of the stationary ARCH( $\infty$ ) process is implied by the asymptotic behaviour of its lag coefficients. Giraitis and Surgailis (2002) also obtained similar results by applying an orthogonal Volterra representation of the squared return process. Davidson (2004) demonstrated that GARCH and IGARCH processes display geometric memory, and the stationary HYGARCH and FIGARCH processes embody hyperbolic memory by applying the NED approach. This was reviewed in more detail in Chapter 2.

Most of the literature above has been summarised by Giraitis et al. (2007) and Giraitis et al. (2009), where the authors reviewed the memory or persistence property of the stationary ARCH( $\infty$ ) process, and showed that the squares ( $\xi_t^2$ ) of the covariance stationary ARCH( $\infty$ ) process has a distributional short memory under finite fourth moment condition. However, nonsummable autocorrelations may exist in the absolute values  $|\xi_t|$  or some (fractional) powers  $|\xi_t|^l$ ,  $l > 0$ , which can have a long memory property. More investigation is required into this.

### 4.3 Stationarity and persistence in ARCH( $\infty$ )-class processes

In this section, we propose to investigate the stationarity and persistence in the ARCH( $\infty$ )-class processes. An analogous expression of ARCH( $\infty$ ) processes can be written as:

$$h_t = \omega + \sum_{j=1}^{\infty} \psi_{jt} h_{t-j} = \omega + \psi_t(L)h_t, \quad (4.18)$$

where:

$$\psi_{jt} = \theta_j z_{t-j}^2. \quad (4.19)$$

In words, this expression can be described as an infinite-order linear difference equation with independently distributed random coefficients.

To focus our attention on the persistence properties of (4.18), it is helpful to apply a variant of the so-called Beveridge and Nelson (1981) decomposition (henceforth, BN), which was introduced as a tool of econometric analysis by Phillips and Solo (1992). The BN decomposition is the easily verified identity for polynomials  $\lambda(x) = \sum_{j=0}^{\infty} \lambda_j x^j$  having the form:

$$\lambda(z) = \lambda(1) + \lambda^*(z)(1 - z),$$

where  $\lambda_j^* = -\sum_{k=j+1}^{\infty} \lambda_k$ . In the present application we consider, for each  $t$ , the stochastic polynomial in the lag operator:

$$\psi_t(L) = \sum_{j=0}^{\infty} \psi_{jt} L^j,$$

where the coefficients  $\psi_{jt}$  are given by (4.19) with  $\psi_{0t} = \theta_0 = 0$ . Then, the BN decomposition form of this expression is:

$$\psi_t(L) = \Psi_t + \psi_t^*(L)(1 - L),$$

where:

$$\Psi_t = \psi_t(1) = \sum_{j=1}^{\infty} \psi_{jt}, \quad (4.20)$$

and note that:

$$E(\Psi_t) = S. \quad (4.21)$$

The coefficients of  $\psi_t^*(L)$  are  $\psi_{0t}^* = 0$ , for  $k \geq 1$ ,

$$\psi_{kt}^* = - \sum_{l=k+1}^{\infty} \theta_l z_{t-l}^2 \leq 0. \quad (4.22)$$

Then the corresponding form of (4.18) is:

$$\begin{aligned} h_t &= \omega + \sum_{j=1}^{\infty} \psi_{jt} h_{t-j} \\ &= \omega + \left( \sum_{j=1}^{\infty} \psi_{jt} \right) h_{t-1} - \psi_{2t} (h_{t-1} - h_{t-2}) - \psi_{3t} (h_{t-1} - h_{t-3}) - \dots \\ &= \omega + \left( \sum_{j=1}^{\infty} \psi_{jt} \right) h_{t-1} - \theta_2 z_{t-2}^2 (h_{t-1} - h_{t-2}) - \theta_3 z_{t-3}^2 (h_{t-1} - h_{t-3}) - \dots \\ &= \omega + \left( \sum_{j=1}^{\infty} \psi_{jt} \right) h_{t-1} + \sum_{k=1}^{\infty} \psi_{kt}^* \Delta h_{t-k}. \end{aligned}$$

It can be written as:

$$h_t = \omega + \Psi_t h_{t-1} + R_t, \quad (4.23)$$

where:

$$R_t = \sum_{k=1}^{\infty} \psi_{kt}^* \Delta h_{t-k}. \quad (4.24)$$

It also can be written as:

$$\begin{aligned} R_t &= -\theta_2 z_{t-2}^2 (h_{t-1} - h_{t-2}) - \theta_3 z_{t-3}^2 (h_{t-1} - h_{t-3}) - \dots \\ &= - \sum_{k=2}^{\infty} \theta_k z_{t-k}^2 (h_{t-1} - h_{t-k}). \end{aligned}$$

Note that if  $\{h_t\}$  is a stationary process, the terms  $\Delta h_t$  are negatively autocorrelated and their contribution to the dynamics is therefore high-frequency, in general.

That the longer-run persistence and stationarity properties of the process depend critically on the distribution of the sequence  $\{\Psi_t\}$  is shown by the Proposition 4.3.1.

**Proposition 4.3.1** *If the stochastic process  $\{h_t^*\}_{t=-\infty}^{\infty}$  where*

$$h_t^* = \omega + \Psi_t h_{t-1}^* \quad (4.25)$$

*satisfies a sufficient condition for  $P(h_t^* < \infty) = 1$ , then  $P(h_t < \infty) = 1$  also holds for (4.18).*

With this consideration in mind we give the following result, establishing a sufficient condition for stationarity of  $\{h_t\}$ . For convenience of notation, let the symbol  $\zeta$  denote the constant  $E(\ln \Psi_t)$ , not depending on  $t$  since  $\{z_t\}$  is i.i.d.

**Proposition 4.3.2** *If*

$$\zeta < 0 \quad (4.26)$$

*then  $\{h_t^*\}_{t=-\infty}^{\infty}$  defined by (4.25) is strictly stationary and ergodic.*

Sufficiency of the covariance stationarity condition  $S = E(\Psi_t) < 1$  follows from Proposition 4.3.2 by the Jensen inequality.

Consider this result in the case of the GARCH(1, 1) process:

$$\begin{aligned} h_t &= \gamma + \alpha \xi_{t-1}^2 + \beta h_{t-1} \\ &= \frac{\gamma}{1 - \beta} + \frac{\alpha}{1 - \beta L} \xi_{t-1}^2 \\ &= \omega + \sum_{j=1}^{\infty} \alpha \beta^{j-1} \xi_{t-j}^2. \end{aligned}$$

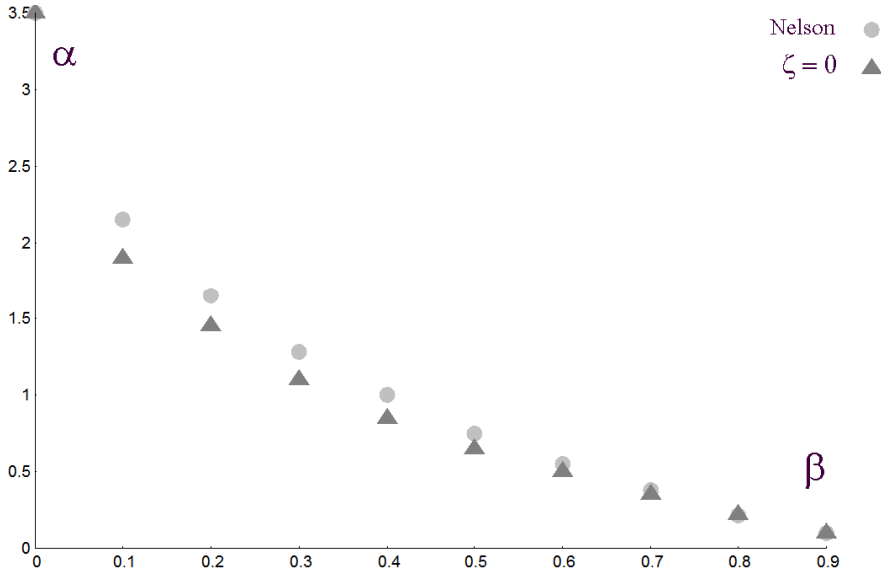


Figure 4.1: Gaussian GARCH(1, 1) model:  $(\alpha, \beta)$  pairs where  $\zeta = 0$  and the stationarity boundary points of Nelson (1990). Note: this figure provides some numerical experiments with Gaussian shocks showing  $\alpha$ -values at which  $\zeta \approx 0$  for  $\beta = 0, 0.1, 0.2, \dots, 0.9$ . The mean is estimated in each case as the average of 20, 000 values of  $\ln(\Psi_t)$ . The actual stationarity boundary points from Nelson (1990) are also shown for comparison.

This is a special case because, uniquely among ARCH( $\infty$ ) processes, it can be expressed exactly in the form of (4.25). In other words, we may write the model as:

$$h_t = \gamma + \Psi_t^\dagger h_{t-1} \quad (4.27)$$

where  $\Psi_t^\dagger = \alpha z_{t-1}^2 + \beta$  and  $\gamma = \omega(1 - \beta)$ . Proposition 4.3.2 can be applied directly to (4.27) to obtain condition (4.11), which Nelson (1990) shows to be necessary as well as sufficient. However, writing the model in its ARCH( $\infty$ ) representation with:

$$\begin{aligned} \zeta &= E \ln(\Psi_t) \\ &= E \ln(\alpha z_{t-1}^2 + \alpha \beta z_{t-2}^2 + \alpha \beta^2 z_{t-3}^2 + \dots) \\ &= E [\ln(\alpha z_{t-1}^2 + \beta \Psi_{t-1})]. \end{aligned} \quad (4.28)$$

In the case  $\beta = 0$ , so that  $S = \alpha$ , the condition (4.11) and (4.26) match. They also

match the necessary condition (4.12) which for the GARCH(1,1) case becomes:

$$E \ln(\alpha z_1^2) < \ln(1 - \beta).$$

Also, letting  $\beta \rightarrow 1$  while letting  $\alpha$  tend to zero at such a rate as to fix the sum of the coefficients at  $S = \alpha/(1 - \beta)$ , note that condition  $\zeta < 0$  in case of (4.28) implies the covariance stationarity condition  $S < 1$ . This follows because  $\Psi_t \rightarrow S$  almost surely as  $\alpha \rightarrow 0$  by the strong law of large numbers, noting that it is a weighted average of i.i.d. random variables with means of unity and weights with finite sum  $S$ .

For the intermediate cases with  $0 < \beta < 1$ , conditions (4.11) and (4.26) do not match but can be compared, giving an opportunity to verify the sharpness of the latter condition. Some numerical experiments with Gaussian shocks are illustrated in Figure 4.1, showing  $\alpha$ -values at which  $\zeta \approx 0$  for  $\beta = 0, 0.1, 0.2, \dots, 0.9$ . The mean is estimated in each case as the average of 20,000 values of  $\ln(\Psi_t)$  where  $\Psi_t$  is calculated from a generated i.i.d. Gaussian sequence  $\{z_t\}$  and the recursion  $\Psi_t = \alpha z_{t-1}^2 + \beta \Psi_{t-1}$ . The actual stationarity boundary points from (4.11) are shown for comparison, as plotted in Figure 4.1 of Nelson (1990)<sup>1</sup>. By comparison, note that the sufficient condition of Douc et al. (2008) is substantially stronger than the bound of Proposition 4.3.2. For the cases illustrated in Figure 4.1, the boundary value of  $S = \alpha/(1 - \beta)$  ranges from 1 at  $\beta = 0.9$  up to 2.1 at  $\beta = 0.1$ . In the Gaussian case, a lower bound on  $E|z_1|^{2p}$  is  $\sqrt{2/\pi} = 0.798$  at  $p = 0.5$ , whereas  $S$  is a lower bound on the second factor of condition (4.16). For most of these cases, there is no value  $p \in (0, 1]$  close to meeting the stated condition.

The way in which these conditions depend on the distribution of  $z_1^2$  can be appreciated by considering Figures 4.2-4.4, which show simulated paths ( $T = 5000$ ,

---

<sup>1</sup>Note that the axes in our figure are interchanged relative to Nelson's figure.

with 10,000 presample steps) for three cases of the IGARCH(1,1) model, with  $\omega = 1$  and  $\beta = 0.9$  in each case. These are among the models studied in Section 4.5 of this chapter. The sole difference between the three cases comes from the shock distributions, which are, respectively, the Student  $t$  with 3 degrees of freedom, the Gaussian, and the uniform, in each case normalised to zero mean and unit variance. Estimates of  $-E(\ln z_1^2)$  (computed as averages of samples of size 20,000) are, respectively, 2.02 for the Student(3), 1.25 for the Gaussian, and 0.87 for the uniform case. These may be compared with  $\ln(S) = 0$  in the light of the necessary stationarity condition (4.12). The plots show how these characteristics map into differences in persistence, pointing up the somewhat counter-intuitive effect of fat tails on persistence.

Turning now to the general ARCH( $\infty$ ) case, note first that from  $E(\Psi_t) = S$  and  $\omega > 0$  it follows that the existence of  $E(h_t^*)$  requires  $S < 1$ , mirroring the full model  $h_t = \omega + \Psi_t h_{t-1} + R_t$ ; in the same case, observe that  $E(R_t) = 0$ . Except in the case where  $S < 1$ , stationarity depends on the distribution of  $\Psi_t$  and particularly on the degree of positive skewness which, as a moving average of squared shocks,  $\Psi_t$  must exhibit in some degree. If the mass of the distribution of  $\Psi_t$  falls below one, the mass of the distribution of  $\ln \Psi_t$  is in the negative part of the line. While  $E(\ln \Psi_t) < \ln S$  by the Jensen inequality, the logarithm of a positive and positively skewed random variable has a more nearly symmetric distribution than the variable itself. Hence,  $E(\ln \Psi_t)$  lies correspondingly closer to  $\text{Median}(\ln \Psi_t) = \ln(\text{Median}(\Psi_t))$ , which in turn lies further below  $\ln S$ , as the skewness is greater. In terms of the dynamics of the process, to the extent that  $\Psi_t$  is symmetrically distributed about its mean  $S$ , and  $S \geq 1$ , the probability that a step is convergent, in the sense of Proposition 4.3.1, is relatively small. The stochastic difference equation defined by (4.25) must, with the complementary probability,

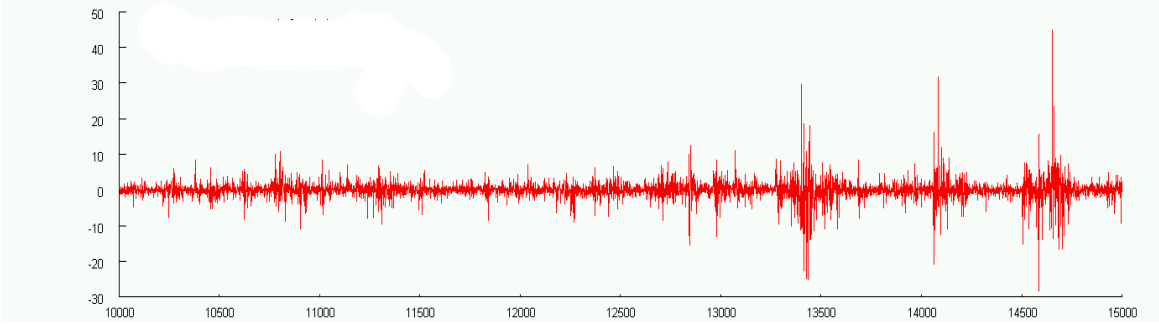


Figure 4.2: Simulation of IGARCH(1,1) with Student(3) shocks. Note: This figure provides the simulation results for samples of size 5000, with 10,000 pre-sample steps, for the IGARCH(1,1) model with  $\omega = 1$ ,  $\beta = 0.9$  and Student(3) shocks. It shows the persistence properties of the IGARCH(1,1) process with Student(3) shocks. More interpretations of this figure are shown in Section 4.3.

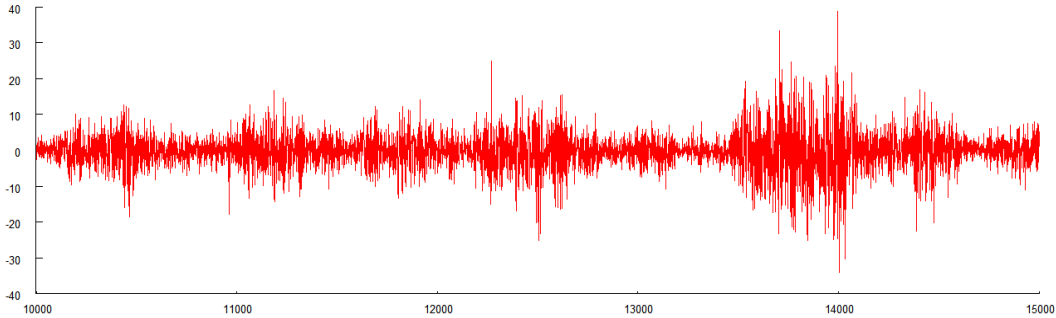


Figure 4.3: Simulation of IGARCH(1,1) with Gaussian shocks. This figure follows same simulation procedure to figure 4.2, except it considers the IGARCH(1,1) process with Gaussian shocks. This figure shows the persistence properties of the IGARCH(1,1) process with Gaussian shocks. More interpretations of this figure are shown in Section 4.3.

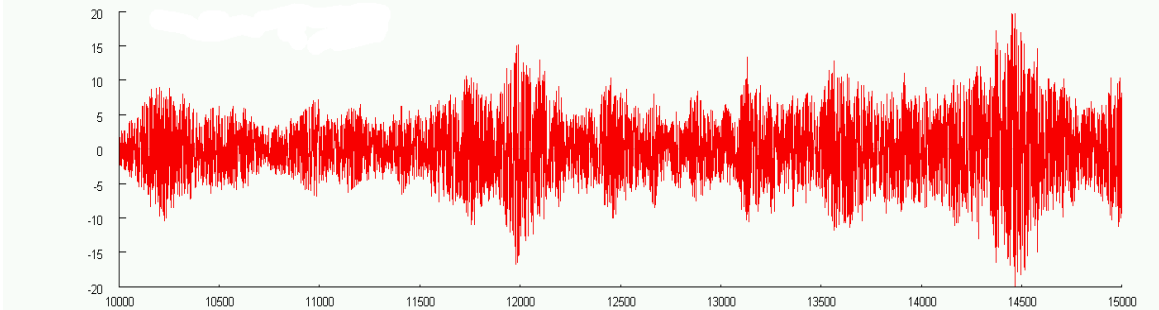


Figure 4.4: Simulation of IGARCH(1,1) with uniformly distributed shocks. Note: This figure follows same simulation procedure to figure 4.2, except it considers the IGARCH(1,1) process with uniformly distributed shocks. This figure shows the persistence properties of the IGARCH(1,1) process with uniformly distributed shocks. More interpretations about this figure are shown in Section 4.3.



behave like either a unit root process with positive drift or an explosive process. However, skewness will increase the proportion of the realisations falling below the mean, yielding stationary behaviour on more frequent occasions, compensated by less frequent but larger excursions above the mean.

In this context we can appreciate the rather complex role played by the rate of decay of the nonnegative sequence  $\{\theta_j\}_{j=1}^{\infty}$ , given its fixed sum  $S = E(\Psi_1)$ . First, note that the skewness of  $\Psi_1$  derives from and is bounded by the skewness in the distribution of the increments  $\{z_s^2, s \leq 0\}$ . Hence, the necessary condition (4.12) can be understood as the minimal condition for non-divergence when  $S \geq 1$ . This condition would also be sufficient in the case  $\theta_j = 0$  for  $j > 1$  and  $S = \theta_1 = 1$  (the IARCH(1) model), in which case the distribution of  $\Psi_1$  and  $z_1^2$  match. However, when  $\Psi_1$  is a moving average of the  $\{z_s^2\}$  process, the distribution of  $\Psi_1$  depends critically on the distribution of the lag coefficients. Since the lag weights have a finite sum  $S$ , the effects of a longer or shorter average lag are to introduce different degrees of averaging of the squared shocks. The somewhat complex nature of this relation depends on the existence of a trade-off between two countervailing effects. Assuming the  $z_1$  possesses a fourth moment, the central limit theorem implies that  $\Psi_1$  is attracted to the normal distribution, with skewness increasingly attenuated, as lag decay gets slower. At the same time, the law of large numbers implies that the variance of  $\Psi_1$  is smaller. The first of these effects is tending to increase the persistence of the  $\{h_t^*\}$  process, while the second is tending to lower the influence of  $h_t^*$  on the volatility of  $\xi_t^* = \sqrt{h_t^*} z_t$ , simply because the noise contribution from  $z_t$  becomes more dominant as the variation in  $h_t^*$  are attenuated. It is therefore difficult to predict the effect of changing the lag decay in any given case.

To summarize: if the contribution of the term  $R_t$  in (4.23) to the persistence

properties can be largely discounted, as we argue, the persistence and stationarity of the ARCH( $\infty$ ) process can be related, through the distribution of  $\Psi_1$ , to the three key factors:  $S$ , the rate of decay of the lag coefficients, and the marginal distribution of  $z_1$ . Greater/smaller kurtosis of  $z_1$  implies greater/smaller positive skewness in the distribution of  $z_1^2$ , and hence gives rise to less/more persistence in  $\{h_t\}$ , other things equal. A longer average lag can, counterintuitively, imply a lesser degree of persistence in the observed process, virtually the opposite of the role of lag decay in models of levels, where the sum of the lag coefficients is not constrained in the same way, and shocks are viewed implicitly as having a symmetric distribution. Finally, it is most important to note that the distinction between exponential and hyperbolic decay rates has quite different implications here than in models of levels. There is no counterpart to so-called long memory in levels, otherwise called fractional integration. The dynamics are nonlinear and there is no simple parallel with linear time series models. The closest analogy is with a single autoregressive root which in the covariance nonstationary cases is local to unity.

#### **4.4 Measuring the persistence of stationary time series**

This section aims to provide a framework for comparing persistence in general time series processes, and also considers some alternative approaches including GPH estimators and GPH estimators for the normalised ranks series. The persistence, or equivalently memory, of a strictly stationary process can be thought of heuristically in terms of the degree to which the history of the process contains information to predict its future path, more accurately than by simple knowledge of the marginal distribution. In the context of univariate forecasting, forecastability must entail

that changes in the level of the process are relatively sluggish. It is customary to measure this type of property with reference to the autocovariance sequence, but this is not a valid approach in the absence of second moments.

#### 4.4.1 $J_T$ statistics

In this study, we resort instead to the idea the key indicator of persistence is the (in)frequency of reversion towards a point of central tendency. We may formalize this notion by defining the persistence of an arbitrary sequence  $\{X_t\}_{t=1}^T$  specifically in terms of the number of occasions on which the series crosses its median point. The direct measure of this property, which is well defined and comparable in any sample sequence whatever, is the relative median-crossing frequency, although it is more convenient to consider the complementary relative frequency of non-crossings. We therefore define:

$$J_T = \frac{1}{T} \sum_{t=2}^T I((X_t - M_T)(X_{t-1} - M_T) > 0), \quad (4.29)$$

where  $T$  is sample length,  $I(\cdot)$  denotes the indicator of its argument and  $M_T$  is the sample median.  $J_T$  measures the persistence of a sample as a point in the unit interval. When the sequence is serially independent,  $J_T \rightarrow 1/2$  as  $T \rightarrow \infty$ , almost surely, by construction. In other words, under independence half of the pairs of successive drawings must fall on different sides of the median on average. The extreme cases are  $J_T \rightarrow 0$  (anti-persistence) and  $J_T \rightarrow 1$  (persistence). In the latter case, at most a finite number of median crossings as  $T \rightarrow \infty$  implies that the sequence either converges, or diverges to infinity. In neither case can it be strictly stationary. The condition  $\limsup J_T < 1$  is evidently necessary for strict stationarity.

$J_T$  in (4.29) applied to a given sequence measures what we may designate persistence in levels. Persistence in volatility is measured by the statistic analogous to  $J_T$  for the squared or (equivalently) absolute values of the series. From the standpoint of returns it is second order persistence, so defined, that is our interest in the present analysis. The  $J_T$  statistic can be computed for arbitrary transformations of the variables, and a necessary and sufficient condition for strict stationarity would appear to be that the sequences  $\{J_T, T \geq 2\}$  are bounded below 1 for all such variants. However, the two leading cases mentioned appear the important ones in the usual time series context.

$J_T$  is an ordinal measure that is well defined regardless of the existence of moments and is also invariant under monotone transformations. Thus, the cases  $X_t = \xi_t^2$  and  $X_t = |\xi_t|$  must yield the same value of  $J_T$ . More interestingly, it is invariant under the operation of forming the normalised ranks of the series,  $\{x_t\}_{t=1}^T$ . Letting  $\hat{F}_T$  denote the empirical distribution function:

$$\hat{F}_T(z) = \frac{1}{T} \sum_{s=1}^T I(X_s \leq z),$$

$x_t = \hat{F}_T(X_t)$  denotes the relative position of  $X_t$  in the sorted sequence  $X_{(1)}, \dots, X_{(T)}$ . The sample median of the normalised ranks tends to 1/2 by construction, and when the sample is large enough,  $J_T$  must have the same value for  $\{x_t\}_{t=1}^T$  as it does for the original series  $\{X_t\}_{t=1}^T$ . The ranks are also invariant under monotone transformations of the series, so yielding the same values for  $X_t = \xi_t^2$  and  $X_t = |\xi_t|$  in particular.

#### 4.4.2 GPH estimator and that for the normalised ranks series

Conventional approaches to measuring persistence, for levels or squares/absolute values as the case may be, are based on the autocovariance sequence. There is particular interest in the property of absolute summability of this sequence, often called weak dependence, with strong dependence defining the non-summable case.<sup>2</sup> Popular persistence measures based on the autocovariance sequence are the so-called GPH log-periodogram regression estimators (for different bandwidths) of the fractional persistence parameter  $d$ , originally due to Geweke and Porter-Hudak (1983). In principle, GPH estimators provide a test of the null hypothesis of weak dependence, although they are well-known to be subject to finite sample bias except under the null of white noise.

Our present interest is due to the fact that the long memory paradigm has proved popular in volatility modelling, and GPH estimation can be validly performed on the normalised ranks of a series regardless of the covariance stationarity property. The particular problem faced in the context of nonstationary volatility is the existence of excessively influential outlying observations, which may invalidate the usual assumptions for valid inference. Rank autocorrelations are free of these influences and may focus more specifically on measuring persistence as characterized here. We should emphasise, though, that our present concerns are not primarily hypothesis testing, but rather to compare and rank different models according to their persistence characteristics.

---

<sup>2</sup>The well-known difficulty of discriminating between these cases in a finite sample has recently been studied in detail by one of the present authors, see Davidson (2009).

### 4.4.3 Persistence measures

To calibrate the performance of these alternative measures, Monte Carlo simulations are used, we generated some pure fractional series, otherwise known as  $I(d)$  processes, for a range of values of  $d$ . The sample size for this experiment is  $T = 10,000$ , with 5000 pre-sample observations. However, the driving shocks were generated to have an  $\alpha$ -stable distribution with  $\alpha = 1.8$  and  $\beta = 1$ , where  $\beta$  is the skewness parameter. The series so constructed do not have second moments and superficially resemble volatility series (after centring) while having a conventional and well-understood linear dependence structure.

Three statistics were computed for these series:  $J_T$  is defined as (4.29), the GPH estimator ( $\hat{d}$ ) with bandwidth  $\sqrt{T}$  for the original series, and also the same GPH estimator for the series of normalised ranks ( $\hat{d}^R$ ). The simulations were repeated 100 times, and the means and standard deviations (in parentheses) of the replications are shown in Table 4.1.

The  $J_T$  statistics discriminate rather clearly between the independent case at one end of the dependence spectrum and the strictly nonstationary unit root at the other. The GPH estimator for the raw data in fact behaves like consistent estimates of  $d$ , while the rank correlation-based estimator appears biased upwards. This is a slightly counter-intuitive result that may or may not be specific to the example considered. However, in our application we are seeking only to rank models, in contexts where a parameter  $d$  with the usual linear property is not typically well defined. (In particular, it does not correspond to the ‘ $d$ ’ appearing in FIGARCH and HYGARCH models.) We carry this alternative along, chiefly, in a spirit of curiosity about the performance of a seemingly natural measure in the context of an exploration of ‘long memory in volatility’.

$d$	$J_T$	$\hat{d}$	$\hat{d}^R$
0	0.498 (0.004)	-0.033 (0.061)	-0.002 (0.065)
0.3	0.663 (0.009)	0.281 (0.061)	0.330 (0.063)
0.5	0.835 (0.024)	0.496 (0.069)	0.544 (0.068)
0.7	0.948 (0.016)	0.718 (0.078)	0.741 (0.075)
0.9	0.985 (0.006)	0.921 (0.013)	0.986 (0.006)
1	0.992 (0.004)	0.985 (0.056)	0.976 (0.065)

Table 4.1: Persistence measures in a fractional linear time series Note: This table provides the simulations are used to generate some pure fractional series, with samples size of 10,000, with 5000 pre-sample observations. The driving shocks were generated to have an  $\alpha$ -stable distribution. The persistent measure  $J_t$ -statistic, the GPH estimator and also the same GPH estimator for the series of normalised ranks are used to measure the persistence of the fractional linear time series and the results are reported in this table. The simulations were repeated 100 times and the means and standard deviations in parentheses were reported in this table.

## 4.5 Some simulation experiments

This section evaluates and compares the properties discussed in section 4.3 by using Monte Carlo experiments.

### 4.5.1 Date generation process

In the first stage, the GARCH(1,1) type models are used to generate the return series. The data generation processes are as following  $\xi_t = \sqrt{h_t}z_t$  where  $z_t \sim$  i.i.d.(0,1) and:

$$h_t = \omega + \left(1 - \frac{1 - \delta L}{1 - \beta L}\right) \xi_t^2, \quad (4.30)$$

where  $\delta > 0$  and  $0 \leq \beta < \min(1, \delta)$ . In (4.30), which matches the GARCH(1,1) model on setting  $\delta = \alpha + \beta$ , and  $S = (\delta - \beta)/(1 - \beta)$ . The pure FI/HYGARCH

is also used to generate the return series and is given by:

$$h_t = \omega + \alpha(1 - (1 - L)^d)\xi_t^2, \quad (4.31)$$

where  $\alpha > 0$  and  $0 < d \leq 1$ . (See e.g. Davidson (2004) for the context of these examples.) whereas in (4.31),  $S = \alpha$ . Setting  $\delta = 1$  and  $\alpha = 1$ , respectively, yields the covariance nonstationary IGARCH and FIGARCH models, whereas setting these parameters strictly less than one implies covariance stationarity.

In the second stage, according to the conditions for the existence of the stationarity, the simulations set a range of values for each of the parameter pairs  $(\delta, \beta)$  in GARCH(1,1) processes. Covariance stationary cases are specified having  $\delta = 0.8$ . For simulating the nonstationary cases, with  $\delta = 1$ , and  $\delta = 1.2$ . For each of these cases, three values of  $\beta$  are chosen, more specifically,  $\beta$  are 0.1, 0.4, 0.7 when  $\delta = 0.8$ , and 0.1, 0.5 and 0.9 when both  $\delta = 1$ , and  $\delta = 1.2$ . It is worth to note that the degree of the persistence depends on the values of  $\beta$ . For the FI/HYGARCH process, the simulations set a range of values for each of the parameter pairs  $(\alpha, d)$ . Covariance stationary cases are specified having  $\alpha = 0.8$ . We also simulate nonstationary cases,  $\alpha = 1$ ,  $\alpha = 1.2$ . For each of these cases, three values of  $d$ , they are 0.9, 0.5 and 0.1, are chosen, being careful to note that the degree of volatility persistence varies inversely with  $d$  (which is of course to be understood as a differencing parameter, not an integration parameter).

In the third stage, for each parameter pair selected, the generation for the underlying processes  $z_t$ . There are three different generation processes: in decreasing order of kurtosis, these are the normalised Student  $t(3)$ ,  $z_{St(3)} = t(3)/\sqrt{3}$ ; the standard Gaussian,  $z_G$ ; and the normalised uniform distribution,  $z_U = \sqrt{12}(U[0, 1] - 1/2)$ .



## 4.5.2 Measurement approach and simulation results

This has been discussed in Section 4.3, and in terms of the results of simulation for persistence measurements, the  $J_T$  statistics and ranked GPH estimators are considered.

Tables 4.2-4.5 show the results for the samples of size  $T = 10,000$ , with 5000 pre-sample observations to account for any start-up effects. The reported values are the averages of  $N = 100$  Monte Carlo replications of the generation process, with the replication standard deviations are shown in parentheses as a guide to the stability of these persistence indicators.

The columns of the tables show the following: first, the parameters and distributions settings for the models; second the sample mean, sample median, and sample logarithmic mean of the random sequences  $\{\Psi_t\}_{t=1}^T$  is defined in (4.20); third, the values of  $J_T$  for various series defined in Section 4.3: the squared returns, the conditional volatilities  $h_t$ , and  $R_t = h_t - \omega - \Psi_t h_{t-1}$ . The final columns of the tables show, for an alternative view of the persistence, the GPH estimators based on the normalised rank correlations of the squared returns.

The salient points of interest in these experimental results seem to us to be the following. First, the relationships between the proximity of the mean of  $\Psi_t$  (measuring  $S$ ) to the corresponding median<sup>3</sup>, and also the proximity of the logarithmic mean to zero, and the measured persistence of the squared returns.

Second, we note that the measured persistence of  $R_t$  is in general much lower than that of  $h_t$ , confirming the fact that  $\Psi_t$  is the key determinant of persistence.

---

<sup>3</sup>The medians are much better determined than the skewness coefficients, which were also computed, but not reported since they convey a very similar picture to the mean-median gaps.

Third, we draw attention to the relative persistence of the squared returns and of the volatility series. In the former case, for given  $\delta$  (or  $\alpha$ ), and given shock distribution, the median-crossing frequencies (measured by  $1 - J_T$ ) actually rise as the lag decay rates decrease, either through  $\beta$  increasing, or  $d$  decreasing. In other words, longer average lags imply less persistence. The reason for this phenomenon has been discussed in Section 4.3, and the interesting observation is that this effect is large enough to counteract the increased persistence in volatility,  $h_t$ , which is also observed.

Finally, we draw attention to the cases with  $\delta = 1.2$  and  $\alpha = 1.2$ , where instances of the logarithmic mean exceeding zero are recorded. In the GARCH case, there is clearly a close correspondence between this occurrence and the evidence that stationarity is violated, in the sense that the median is crossed fewer than ten times in 10,000 steps. The necessary condition (4.12) can also be checked out. Compare the estimated values of  $-E(\ln z_t^2)$  for the three distributions, as reported in Section 4.3. When  $S = 3$  so that  $\ln(S) = 1.09$ , which is the GARCH case corresponding to  $\delta = 1.2$  and  $\beta = 0.9$ , only the uniform distribution case actually violates the necessary condition, but all the distribution alternatives appear nonstationary. All the HYGARCH examples appear stationary, although the uniform case with  $d = 0.5$  appears the closest to divergent.

The estimates of the fractional integration parameter in the last column of the tables are of interest in reflecting the persistence measured by  $J_T$  quite closely, increasing across the range with  $\beta$ , but are non-monotone with respect to  $d$ . Observe that, for the normal and uniform cases in Tables 4.4 and 4.5, the values obtained for  $d = 0.5$  are generally greater than those for either  $d = 0.9$  or  $d = 0.1$ . When the volatility is covariance nonstationary these measures can be quite large, and

$\delta$	Model		$\Psi_t$			$\xi_t^2$	$J_T$		$d^R$ $\xi_t^2$
	$\beta$	Dist'n	Mean	Median	MeanLog		$h_t$	$R_t$	
0.8	0.1	<i>St(3)</i>	0.772 (0.119)	0.204 (0.004)	-1.630 (0.021)	0.571 (0.005)	0.634 (0.005)	0.468 (0.005)	-0.004 (0.070)
		<i>N</i>	0.777 (0.011)	0.411 (0.007)	-1.015 (0.018)	0.613 (0.005)	0.662 (0.005)	0.460 (0.004)	-0.002 (0.076)
		<i>U</i>	0.778 (0.007)	0.605 (0.011)	-0.746 (0.014)	0.639 (0.006)	0.677 (0.006)	0.469 (0.005)	0.006 (0.071)
	0.4	<i>St(3)</i>	0.663 (0.065)	0.265 (0.005)	-1.286 (0.023)	0.553 (0.005)	0.748 (0.004)	0.634 (0.004)	-0.009 (0.078)
		<i>N</i>	0.667 (0.009)	0.480 (0.009)	-0.788 (0.015)	0.576 (0.006)	0.751 (0.006)	0.615 (0.004)	0.014 (0.062)
		<i>U</i>	0.666 (0.006)	0.594 (0.007)	-0.618 (0.011)	0.585 (0.006)	0.739 (0.006)	0.574 (0.005)	0.013 (0.061)
	0.7	<i>St(3)</i>	0.329 (0.038)	0.182 (0.004)	-1.638 (0.021)	0.517 (0.005)	0.835 (0.004)	0.767 (0.004)	0.004 (0.072)
		<i>N</i>	0.333 (0.004)	0.289 (0.004)	-1.262 (0.015)	0.519 (0.005)	0.809 (0.005)	0.729 (0.005)	-0.002 (0.058)
		<i>U</i>	0.333 (0.003)	0.323 (0.003)	-1.176 (0.009)	0.521 (0.006)	0.774 (0.005)	0.675 (0.005)	0.014 (0.065)

Table 4.2: Series properties and persistence measures for the covariance stationary GARCH(1,1) model. Note: This table shows the results for the samples of size 10,000, with 5000 pre-sample observations. The reported values are the averages of 100 Monte Carlo replications of the generation process, with the replication standard errors are provided in parentheses. The sample mean, sample median, and sample logarithmic mean of the random sequences  $\{\Psi_t\}_{t=1}^T$ , values of  $J_T$ , and the GPH estimators based on the normalised rank correlations are reported in this table. These results demonstrate the strict stationarity and persistence properties of the squared return and conditional variance processes. More details are provided in Section 4.5.2.

when it is strictly nonstationary, they fall close to unity. In a series of insightful papers, Mikosch and Stărică (2003, 2004) argue that long range dependence of volatility in financial data should be attributed to structural breaks in the unconditional variance, rather than to GARCH-type dynamics. However, it is clear that apparent long range dependence can be observed in the stationary cases simulated here. We would agree with these authors that the evidence of long-range dependence is spurious, in the sense that it is not generated by a fractionally integrated structure, as it is in Table 4.1 for example. However, our diagnosis of the cause does not invoke structural breaks. Rather, we see it as a phenomenon analogous to having an autoregressive root local to unity in a levels process, leading to

$\delta$	Model		$\Psi_t$			$J_T$			$d^R$
	$\beta$	Dist'n	Mean	Median	MeanLog	$\xi_t^2$	$h_t$	$R_t$	$\xi_t^2$
1	0.1	<i>St(3)</i>	1.020 (0.262)	0.262 (0.006)	-1.379 (0.020)	0.588 (0.005)	0.650 (0.005)	0.474 (0.005)	0.002 (0.059)
		<i>N</i>	1.001 (0.015)	0.530 (0.011)	-0.761 (0.014)	0.647 (0.007)	0.699 (0.006)	0.477 (0.004)	0.017 (0.056)
		<i>U</i>	1.001 (0.010)	0.778 (0.015)	-0.495 (0.014)	0.693 (0.008)	0.740 (0.008)	0.495 (0.005)	0.049 (0.056)
	0.5	<i>St(3)</i>	1.068 (0.857)	0.442 (0.010)	-0.769 (0.019)	0.575 (0.006)	0.809 (0.004)	0.693 (0.005)	0.008 (0.060)
		<i>N</i>	0.997 (0.012)	0.768 (0.012)	-0.298 (0.016)	0.627 (0.008)	0.841 (0.006)	0.680 (0.005)	0.061 (0.073)
		<i>U</i>	0.999 (0.008)	0.927 (0.009)	-0.158 (0.009)	0.671 (0.010)	0.866 (0.007)	0.642 (0.005)	0.189 (0.060)
	0.9	<i>St(3)</i>	0.975 (0.098)	0.693 (0.018)	-0.290 (0.031)	0.553 (0.010)	0.953 (0.005)	0.900 (0.004)	0.279 (0.073)
		<i>N</i>	0.998 (0.015)	0.954 (0.016)	-0.049 (0.013)	0.594 (0.022)	0.965 (0.005)	0.873 (0.005)	0.615 (0.074)
		<i>U</i>	0.999 (0.008)	0.992 (0.008)	-0.022 (0.010)	0.629 (0.032)	0.971 (0.006)	0.844 (0.005)	0.729 (0.072)
1.2	0.1	<i>St(3)</i>	1.209 (0.180)	0.319 (0.007)	-1.176 (0.019)	0.607 (0.005)	0.669 (0.005)	0.480 (0.004)	0.004 (0.069)
		<i>N</i>	1.224 (0.016)	0.648 (0.011)	-0.565 (0.015)	0.685 (0.007)	0.739 (0.007)	0.495 (0.005)	0.005 (0.068)
		<i>U</i>	1.223 (0.012)	0.951 (0.018)	-0.294 (0.013)	0.760 (0.009)	0.807 (0.008)	0.517 (0.006)	0.050 (0.067)
	0.5	<i>St(3)</i>	1.396 (0.188)	0.619 (0.013)	-0.428 (0.024)	0.617 (0.007)	0.843 (0.006)	0.708 (0.005)	0.054 (0.068)
		<i>N</i>	1.400 (0.020)	1.080 (0.018)	0.037 (0.015)	0.840 (0.028)	0.952 (0.010)	0.714 (0.006)	0.574 (0.094)
		<i>U</i>	1.399 (0.010)	1.299 (0.013)	0.178 (0.009)	0.994 (0.003)	0.998 (0.001)	0.779 (0.025)	0.999 (0.025)
	0.9	<i>St(3)</i>	3.025 (0.381)	2.090 (0.055)	0.810 (0.034)	0.999 (0.001)	1.000 (0.001)	1.000 (0.001)	0.959 (0.099)
		<i>N</i>	2.999 (0.044)	2.867 (0.048)	1.045 (0.016)	1.000 (0)	1.000 (0)	1.000 (0)	0.991 (0.047)
		<i>U</i>	2.998 (0.026)	2.976 (0.028)	1.078 (0.009)	1.000 (0)	1.000 (0)	1.000 (0)	1.019 (0.010)

Table 4.3: Series properties and persistence measures for the nonstationary GARCH(1,1) model. Note: The simulation results in this table were obtained by following a similar simulation procedure to that in the Table 4.2, except the DGP here considers the nonstationary GARCH(1,1) model. The strict stationarity, persistence of the squared returns and conditional variance are significantly different from those in Table 4.2. Further discussions can be seen in Section 4.5.2.

$\alpha$	Model		$\Psi_t$			$\xi_t^2$	$J_T$		$d^R$ $\xi_t^2$
	$d$	Dist'n	Mean	Median	MeanLog		$h_t$	$R_t$	
0.8	0.9	<i>St(3)</i>	0.805 (0.140)	0.218 (0.005)	-1.418 (0.027)	0.571 (0.005)	0.634 (0.007)	0.656 (0.016)	0.018 (0.073)
		<i>N</i>	0.800 (0.011)	0.412 (0.008)	-0.898 (0.016)	0.614 (0.005)	0.655 (0.006)	0.608 (0.007)	0.030 (0.073)
		<i>U</i>	0.801 (0.008)	0.622 (0.012)	-0.678 (0.012)	0.642 (0.006)	0.669 (0.007)	0.570 (0.006)	0.018 (0.068)
	0.5	<i>St(3)</i>	0.797 (0.091)	0.422 (0.012)	-0.738 (0.030)	0.556 (0.006)	0.785 (0.012)	0.738 (0.020)	0.172 (0.093)
		<i>N</i>	0.800 (0.011)	0.614 (0.010)	-0.432 (0.016)	0.577 (0.006)	0.754 (0.008)	0.656 (0.007)	0.157 (0.078)
		<i>U</i>	0.798 (0.007)	0.710 (0.009)	-0.345 (0.010)	0.585 (0.006)	0.722 (0.007)	0.599 (0.005)	0.134 (0.070)
	0.1	<i>St(3)</i>	0.765 (0.079)	0.611 (0.030)	-0.400 (0.056)	0.523 (0.007)	0.873 (0.017)	0.809 (0.012)	0.224 (0.101)
		<i>N</i>	0.778 (0.010)	0.730 (0.010)	-0.279 (0.014)	0.524 (0.005)	0.797 (0.010)	0.732 (0.006)	0.143 (0.066)
		<i>U</i>	0.779 (0.006)	0.763 (0.007)	-0.266 (0.008)	0.525 (0.005)	0.753 (0.013)	0.666 (0.005)	0.149 (0.062)

Table 4.4: Series properties and persistence measures for the covariance stationary HY/FIGARCH model. Note: The simulation procedure and table structure are same as that in the Table 4.2, except the DGP here considers the covariance stationary HY/FIGARCH model. More details are provided in Section 4.5.2.

Ornstein-Uhlenbeck-type dynamics which are easily confused with long memory in finite samples. However, the analogy is necessarily a loose one in view of the special features of the volatility process which we have detailed in Section 4.3.

## 4.6 Implications for volatility forecasting

This section investigates the implications for volatility forecasting and proposes a new forecasting method. The standard recursion for a minimum mean squared error (MSE) forecast, with  $\xi_{T+j}^2$  for  $j > 0$  replaced by its (assumed) conditional expectation, always used to do the volatility forecasting two or more steps ahead based on the ARCH/GARCH type models. For instant, Poon (2005) page 39, and also the Eviews (2013), page 218, for a practical implementation.

$\alpha$	Model		$\Psi_t$			$\xi_t^2$	$J_T$	$R_t$	$d^R$
	$d$	Dist'n	Mean	Median	MeanLog		$h_t$		$\xi_t^2$
1	0.9	<i>St(3)</i>	0.980 (0.120)	0.271 (0.005)	-1.199 (0.020)	0.588 (0.005)	0.656 (0.007)	0.673 (0.016)	0.043 (0.073)
		<i>N</i>	1.002 (0.013)	0.516 (0.011)	-0.674 (0.016)	0.648 (0.007)	0.702 (0.010)	0.648 (0.014)	0.105 (0.083)
		<i>U</i>	1.001 (0.008)	0.777 (0.013)	-0.453 (0.014)	0.695 (0.009)	0.741 (0.011)	0.619 (0.011)	0.166 (0.062)
	0.5	<i>St(3)</i>	0.993 (0.117)	0.528 (0.016)	-0.507 (0.035)	0.575 (0.009)	0.818 (0.017)	0.760 (0.023)	0.291 (0.086)
		<i>N</i>	0.999 (0.014)	0.767 (0.011)	-0.207 (0.014)	0.622 (0.019)	0.832 (0.024)	0.696 (0.015)	0.410 (0.087)
		<i>U</i>	0.999 (0.009)	0.887 (0.012)	-0.122 (0.009)	0.651 (0.022)	0.838 (0.026)	0.635 (0.010)	0.468 (0.085)
	0.1	<i>St(3)</i>	0.995 (0.306)	0.769 (0.047)	-0.172 (0.067)	0.536 (0.013)	0.893 (0.022)	0.820 (0.016)	0.352 (0.107)
		<i>N</i>	0.976 (0.013)	0.914 (0.012)	-0.057 (0.013)	0.537 (0.006)	0.855 (0.021)	0.744 (0.006)	0.304 (0.058)
		<i>U</i>	0.973 (0.007)	0.953 (0.007)	-0.042 (0.009)	0.538 (0.006)	0.839 (0.024)	0.682 (0.006)	0.344 (0.060)
1.2	0.9	<i>St(3)</i>	1.183 (0.129)	0.326 (0.007)	-1.013 (0.023)	0.606 (0.005)	0.682 (0.009)	0.693 (0.020)	0.083 (0.089)
		<i>N</i>	1.198 (0.015)	0.617 (0.012)	-0.496 (0.015)	0.702 (0.027)	0.780 (0.033)	0.716 (0.034)	0.305 (0.116)
		<i>U</i>	1.201 (0.010)	0.932 (0.015)	-0.271 (0.014)	0.830 (0.041)	0.895 (0.033)	0.751 (0.022)	0.562 (0.101)
	0.5	<i>St(3)</i>	1.211 (0.242)	0.631 (0.021)	-0.322 (0.033)	0.620 (0.032)	0.871 (0.030)	0.792 (0.024)	0.431 (0.121)
		<i>N</i>	1.193 (0.018)	0.916 (0.015)	-0.023 (0.014)	0.929 (0.034)	0.979 (0.013)	0.721 (0.010)	0.966 (0.047)
		<i>U</i>	1.193 (0.010)	1.062 (0.013)	0.063 (0.008)	0.977 (0.010)	0.992 (0.005)	0.683 (0.009)	1.004 (0.013)
	0.1	<i>St(3)</i>	1.148 (0.104)	0.917 (0.041)	0.010 (0.057)	0.574 (0.047)	0.925 (0.026)	0.827 (0.011)	0.454 (0.130)
		<i>N</i>	1.167 (0.015)	1.094 (0.014)	0.125 (0.013)	0.643 (0.029)	0.956 (0.013)	0.788 (0.014)	0.643 (0.058)
		<i>U</i>	1.168 (0.009)	1.144 (0.010)	0.141 (0.009)	0.686 (0.024)	0.964 (0.010)	0.792 (0.019)	0.695 (0.056)

Table 4.5: Series properties and persistence measures for the nonstationary HY/FIGARCH model. Note: The simulation procedure and table structure are same as that in the Table 4.2, except the DGP here considers the covariance non-stationary HY/FIGARCH model. More details are provided in Section 4.5.2.

In other words, if  $h_t$  is defined by the ARCH( $\infty$ ) process (4.2), then:

$$\begin{aligned} h_{t+1} &= \omega + \sum_{j=1}^{\infty} \theta_j \xi_{t-j+1}^2 \\ &= \omega + \theta_1 \xi_t^2 + \theta_2 \xi_{t-1}^2 + \theta_3 \xi_{t-2}^2 + \theta_4 \xi_{t-3}^2 + \dots \\ &= \omega + \theta_1 \xi_t^2 + \sum_{j=2}^{\infty} \theta_j \xi_{t-j+1}^2, \end{aligned}$$

(and if we implicitly assuming the parameters are replaced by appropriate estimates) we would replace  $\xi_t^2$  by  $E_{t-1} \xi_t^2 = h_t$ , and so set<sup>4</sup>:

$$\begin{aligned} \widehat{h}_{t+1|t-1} &= \omega + \theta_1 E_{t-1} \xi_t^2 + \sum_{j=2}^{\infty} \theta_j \xi_{t-j+1}^2 \\ &= \omega + \theta_1 h_t + \sum_{j=2}^{\infty} \theta_j \xi_{t-j+1}^2. \end{aligned} \tag{4.32}$$

The volatility forecast error accordingly has the form:

$$\begin{aligned} f_{t+1|t-1} &= h_{t+1} - \widehat{h}_{t+1|t-1} \\ &= \omega + \theta_1 \xi_t^2 + \sum_{j=2}^{\infty} \theta_j \xi_{t-j+1}^2 - (\omega + \theta_1 h_t + \sum_{j=2}^{\infty} \theta_j \xi_{t-j+1}^2) \\ &= \theta_1 \xi_t^2 - \theta_1 h_t \\ &= \theta_1 h_t (z_t^2 - 1). \end{aligned} \tag{4.33}$$

Then, since:

$$\begin{aligned} h_{t+2} &= \omega + \sum_{j=1}^{\infty} \theta_j \xi_{t-j+2}^2 \\ &= \omega + \theta_1 \xi_{t+1}^2 + \theta_2 \xi_t^2 + \theta_3 \xi_{t-1}^2 + \theta_4 \xi_{t-2}^2 + \dots \end{aligned}$$

Similarly, for the 3-step ahead:

$$\begin{aligned} \widehat{h}_{t+2|t-1} &= \omega + \theta_1 \widehat{h}_{t+1|t-1} + \theta_2 h_t + \theta_3 \xi_{t-1}^2 + \theta_4 \xi_{t-2}^2 + \dots \\ &= \omega + \theta_1 \widehat{h}_{t+1|t-1} + \theta_2 h_t + \sum_{j=2+1}^{\infty} \theta_j \xi_{t-j+2}^2 \\ &= \omega + \theta_1 \widehat{h}_{t+1|t-1} + \theta_2 h_t + \sum_{j=2+1}^{\infty} \theta_j \xi_{t-j+2}^2. \end{aligned}$$

---

<sup>4</sup>We call this expression the two-step volatility forecast since  $h_t$  itself is of course the one-step forecast.

Thus, in the general  $k$ -step ahead case:

$$h_{t+k} = \omega + \sum_{j=1}^k \theta_j \xi_{t-j+k}^2 + \sum_{j=k+1}^{\infty} \theta_j \xi_{t-j+k}^2,$$

and:

$$\widehat{h}_{t+k|t-1} = \omega + \sum_{j=1}^{k-1} \theta_j \widehat{h}_{t-j+k|t-1} + \theta_k h_t + \sum_{j=k+1}^{\infty} \theta_j \xi_{t-j+k}^2, \quad (4.34)$$

and the volatility forecast error as:

$$\begin{aligned} f_{t+k|t-1} &= h_{t+k} - \widehat{h}_{t+k|t-1} & (4.35) \\ &= \sum_{j=1}^k \theta_j \xi_{t-j+k}^2 - \sum_{j=1}^{k-1} \theta_j \widehat{h}_{t-j+k|t-1} - \theta_k h_t \\ &= \sum_{j=1}^k \theta_j \xi_{t-j+k}^2 - \sum_{j=1}^{k-1} \theta_j (h_{t-j+k} - f_{t-j+k|t-1}) - \theta_k h_t \\ &= \sum_{j=1}^k \theta_j \xi_{t-j+k}^2 - \sum_{j=1}^{k-1} \theta_j h_{t-j+k} - \theta_k h_t + \sum_{j=1}^{k-1} \theta_j f_{t-j+k|t-1} \\ &= \sum_{j=1}^{k-1} \theta_j f_{t-j+k|t-1} + \sum_{j=1}^k \theta_j (\xi_{t-j+k}^2 - h_{t-j+k}) \\ &= \sum_{j=1}^{k-1} \theta_j f_{t-j+k|t-1} + \sum_{j=1}^k \theta_j h_{t-j+k} (z_{t-j+k}^2 - 1). \end{aligned}$$

For example, consider the GARCH(1,1) model in (4.30) which rearranges as:

$$\begin{aligned} h_{t+1} &= \gamma + \alpha \xi_t^2 + \beta h_t \\ &= \omega(1 - \beta) + (\delta - \beta) z_t^2 h_t + \beta h_t \\ &= \omega(1 - \beta) + [(\delta - \beta) z_t^2 + \beta] h_t. \end{aligned}$$

If  $z_t^2$  is replaced by  $E_{t-1}(z_t^2) = 1$  to construct the forecast, then (4.35) reduces to:

$$\begin{aligned} f_{t+1|t-1} &= [(\delta - \beta) z_t^2 + \beta] h_t - \delta h_t \\ &= (\delta - \beta) h_t (z_t^2 - 1). \end{aligned}$$

The problem with this formulation, as the preceding analysis demonstrates, is that due to the skewness of the distribution of  $z_t^2$ , the mean may not be the



best measure of central tendency. The persistence of the process, and hence its forecastability, will be exaggerated by this choice. In effect, the problem is closely allied to that of forecasting in model (4.25) by using  $S$  as the forward projection for unobserved  $\Psi_t$ .  $S$  is not the value that  $\Psi_t$  is close to with highest probability, and hence the one that will deliver an accurate projection with high probability. The majority of volatility forecasts will be "overshoots", balanced by a smaller number of more extreme "undershoots". The forecast is unbiased in the sense  $E(f_{t+k|t-1}) = 0$  when this expectation is defined, but this condition excludes the IGARCH and FIGARCH and other nonstationary cases. Even if the mean squared forecast error is defined, in the context, it is not clear that the MSE is an appropriate loss function.

We investigated this issue experimentally with the results reported in Table 4.6 and 4.7 for the GARCH(1,1) and pure HY/FIGARCH models respectively. We studied the distribution of errors in the two-step forecasts constructed under different assumptions about the appropriate measure of central tendency of the shocks, denoted by  $M$  in the definition:

$$f_{t+1|t-1} = \theta_1 h_t (z_t^2 - M). \quad (4.36)$$

The median absolute values (MAVs) of the variables defined in (4.36) were computed for six choices of  $M$ . In the tables, the minimum value of the MAV in each row is indicated in boldface. Note that in only two of these cases does  $M$  exceed 0.5 and in both, the difference from the adjacent lower value is minimal. The rule that  $M = 0.1$  gives the best result for the Student(3) case,  $M = 0.3$  for the Gaussian case and  $M = 0.5$  for the uniform case appears to hold quite generally. The implication may be that future volatility is significantly overstated by conventional procedures.

We can reasonably assume that the optimal  $M$  values are those closest to the modes of the respective distributions. While estimating the mode of an empirical distribution is not a straightforward procedure, constructing medians is easy and the medians of our squared normalised distributions, estimated from samples of size 10,000, are 0.176 for the Student(3), 0.423 for the Gaussian and 0.763 for the uniform distribution. In default of a more precise analysis, a rough and ready rule of thumb would be to estimate the MAV-minimising  $M$  by 2/3 times the sample median of the squared normalised residuals. This corresponds to computing the  $k$ -step volatility forecasts by the recursion:

$$\widehat{h}_{t+k|t-1} = \omega + \frac{2}{3} \text{Median}(z_t^2) \sum_{j=1}^k \theta_j \widehat{h}_{t-j+k|t-1} + \sum_{j=k+1}^{\infty} \theta_j \xi_{t-j+k}^2, \quad (4.37)$$

where  $\widehat{h}_{t|t-1} = h_t$ .

A more extensive simulation study than the present one would be needed to confirm this recommendation. We do note, however, that the rule would apply successfully in both the covariance stationary and the covariance nonstationary cases that have been simulated here. Although  $h_t$  has the interpretation of a conditional variance only in the stationary case, note that the problem we highlight is not connected with the non-existence of moments. It is entirely a matter of adopting a minimum MSE estimator of a highly skewed distribution, such that the outcome is overestimated in a substantially higher proportion of cases than it is underestimated.

$\delta$	Model		$M$					
	$\beta$	Dist'n	1	0.9	0.7	0.5	0.3	0.1
0.8	0.1	<i>St(3)</i>	0.070	0.063	0.049	0.035	0.020	<b>0.010</b>
		<i>N</i>	0.078	0.070	0.054	0.041	<b>0.032</b>	0.047
		<i>U</i>	0.091	0.083	0.071	<b>0.066</b>	0.085	0.121
	0.4	<i>St(3)</i>	0.171	0.153	0.118	0.084	0.052	<b>0.028</b>
		<i>N</i>	0.204	0.185	0.148	0.115	<b>0.090</b>	0.111
		<i>U</i>	0.232	0.216	0.184	<b>0.162</b>	0.176	0.246
	0.7	<i>St(3)</i>	0.073	0.065	0.050	0.036	0.022	<b>0.010</b>
		<i>N</i>	0.076	0.069	0.055	0.041	<b>0.028</b>	0.034
		<i>U</i>	0.076	0.070	0.058	0.046	<b>0.045</b>	0.065
1.0	0.1	<i>St(3)</i>	0.094	0.084	0.064	0.045	0.028	<b>0.015</b>
		<i>N</i>	0.122	0.110	0.088	0.069	<b>0.061</b>	0.085
		<i>U</i>	0.184	0.174	<b>0.160</b>	0.161	0.193	0.257
	0.5	<i>St(3)</i>	0.338	0.303	0.234	0.168	0.108	<b>0.069</b>
		<i>N</i>	0.693	0.637	0.536	0.446	<b>0.386</b>	0.410
		<i>U</i>	1.198	1.150	1.076	<b>1.034</b>	1.067	1.248
	0.9	<i>St(3)</i>	0.246	0.221	0.173	0.127	0.085	<b>0.054</b>
		<i>N</i>	1.118	1.033	0.876	0.734	0.621	<b>0.592</b>
		<i>U</i>	2.352	2.267	2.109	1.984	<b>1.932</b>	2.103

Table 4.6: 2MAV 2-step forecast error in GARCH(1, 1), against  $M$ . Note: In this table, the minimum value of the MAV in each row is indicated in boldface. In only one of these cases does  $M$  exceed 0.5, the difference from the adjacent lower value is minimal. The rule that  $M = 0.1$ ,  $M = 0.3$  and  $M = 0.5$  give the best results for the Student(3) case, the Gaussian case, and the uniform case appear to hold quite generally in GARCH(1, 1) models, respectively. The implication may be that future volatility is significantly overstated by conventional procedures.

## 4.7 Conclusion

In this chapter, we have investigated the dynamics of certain conditional volatility models with a view to understanding their propensity to predict persistent patterns of high or low volatility. Understanding how persistence depends on the various model characteristic, while intriguing and often counterintuitive, is perhaps a matter of mainly theoretical interest. However, there is also an important message here for practitioners. Conventional forecasting methodologies that are optimal under the assumption of symmetrically distributed shocks may be viewed

$\alpha$	Model		$M$					
	$d$	Dist'n	1	0.9	0.7	0.5	0.3	0.1
0.8	0.9	<i>St(3)</i>	0.038	0.034	0.026	0.018	0.011	<b>0.005</b>
		<i>N</i>	0.043	0.039	0.030	0.023	<b>0.017</b>	0.026
		<i>U</i>	0.050	0.046	0.039	<b>0.036</b>	0.048	0.068
	0.5	<i>St(3)</i>	0.156	0.140	0.108	0.078	0.049	<b>0.027</b>
		<i>N</i>	0.216	0.196	0.157	0.122	<b>0.094</b>	0.114
		<i>U</i>	0.244	0.226	0.192	<b>0.166</b>	0.182	0.257
	0.1	<i>St(3)</i>	0.053	0.047	0.036	0.026	0.016	<b>0.007</b>
		<i>N</i>	0.055	0.049	0.039	0.029	<b>0.019</b>	0.024
		<i>U</i>	0.053	0.049	0.041	0.032	<b>0.031</b>	0.045
1.0	0.9	<i>St(3)</i>	0.050	0.045	0.034	0.024	0.014	<b>0.008</b>
		<i>N</i>	0.073	0.066	0.053	0.041	<b>0.036</b>	0.050
		<i>U</i>	0.116	0.110	<b>0.101</b>	<b>0.101</b>	0.120	0.161
	0.5	<i>St(3)</i>	0.300	0.269	0.208	0.150	0.097	<b>0.060</b>
		<i>N</i>	1.382	1.271	1.060	0.874	<b>0.734</b>	0.751
		<i>U</i>	3.862	3.688	3.372	<b>3.164</b>	3.198	3.838
	0.1	<i>St(3)</i>	0.429	0.384	0.298	0.216	0.138	<b>0.076</b>
		<i>N</i>	0.663	0.603	0.488	0.379	<b>0.279</b>	0.314
		<i>U</i>	0.658	0.614	0.526	0.440	<b>0.419</b>	0.599

Table 4.7: MAV 2-step forecast error in HY/FIGARCH, against  $M$ . Note: in this table, the minimum value of the MAV in each row is indicated in boldface. The simulation procedure and table structure are same as that in the Table 4.6, except the DGP here considers the HY/FIGARCH model. This table also shows that future volatility is significantly overstated by conventional procedures.

as overstating the degree of future volatility. This is, of course, an issue essentially of the preferred choice of loss function. Practitioners may validly elect to favour the unbiasedness and minimum MSE properties over minimising the MAV. They should nonetheless not overlook the fact the usual rationale for the former criterion implicitly assumes a Gaussian framework, and is arguably inappropriate in the context of predicting volatility.

## 4.8 Appendix C

### 4.8.1 Proof of Proposition 4.3.1

**Proof of Proposition 4.3.1.** First consider the case of where  $\{\psi_{jt}\}$  is replaced by  $\{\psi_j\}$ , a nonstochastic sequence of coefficients. Then:

$$h_t = \omega + \sum_{j=1}^{\infty} \psi_j h_{t-j}, \quad (4.38)$$

with  $\omega > 0$  and  $\psi_j \geq 0$  for all  $j \geq 1$  has a stable, positive solution if and only if this is true of the equation:

$$h_t^* = \omega + \left( \sum_{j=1}^{\infty} \psi_j \right) h_{t-1}^*. \quad (4.39)$$

Stable solutions of (4.38) and (4.39), if they exist, are both of the form:

$$\frac{\omega}{1 - \sum_{j=1}^{\infty} \psi_j} > 0, \quad (4.40)$$

implying in both cases the necessary and sufficient condition:

$$\sum_{j=1}^{\infty} \psi_j < 1. \quad (4.41)$$

Next, consider the stochastic sequence  $\{\psi_{jt}\}$ . Let this be randomly drawn at date  $t_0$ , as the functional of the random sequence  $\{z_{t_0-j}, j > 0\}$ , and then a let step be taken according to either (4.18) or (4.25). Call this in either case a convergent step if  $\sum_{j=1}^{\infty} \psi_{jt_0} = \Psi_{t_0} < 1$ . That is, if the process is allowed to continue with this same fixed drawing, the sequence of steps so generated must approach the particular solution:

$$\bar{h}_0 = \frac{\omega}{1 - \Psi_{t_0}}. \quad (4.42)$$

This is a drawing from the common distribution of stable solutions, which are almost surely finite.

Suppose that every step taken is convergent, in this sense. Then, the sequence is always moving so as to reduce its distance from some point in the distribution of stable solutions. It therefore cannot diverge. More generally, let each step have a certain fixed probability of being convergent. The probability that the sequence diverges can be reduced to zero by setting this probability high enough. This is, from elementary considerations, a sufficient condition for  $\{h_t^*\}$  to be finite almost surely.

To show that the same condition is sufficient for  $\{h_t\}$  generated by:

$$h_t = \omega + \sum_{j=1}^{\infty} \psi_{jt} h_{t-j},$$

to be finite almost surely, first note that since:

$$\begin{aligned} h_t^* &= \omega + \Psi_{t_0} h_{t-1}^* \\ &= \bar{h}_0(1 - \Psi_{t_0}) + \Psi_{t_0} h_{t-1}^*. \end{aligned}$$

Subtracting  $h_{t-1}^*$  on the two sides of above equation, then:

$$h_t^* - h_{t-1}^* = \bar{h}_0(1 - \Psi_{t_0}) + \Psi_{t_0} h_{t-1}^* - h_{t-1}^*,$$

then the step defined by (4.25) can be written for given  $\Psi_{t_0}$  in the form:

$$\Delta h_t^* = (\Psi_{t_0} - 1)(h_{t-1}^* - \bar{h}_0). \quad (4.43)$$

In this representation, the condition for a convergent step is that  $\Delta h_t^*$  and  $h_{t-1}^* - \bar{h}_0$  have different signs. Now write the BN form of:

$$h_t = \omega + \Psi_t h_{t-1} + R_t$$

in the equivalent representation, as:

$$\Delta h_t = (\Psi_{t_0} - 1)(h_{t-1}^* - \bar{h}_0) + R_t^0, \quad (4.44)$$

where the remainder, like  $\Psi_{t_0}$ , is specified for the particular shock sequence  $\{z_{t_0-j}, j > 0\}$  as:

$$R_t^0 = \sum_{k=1}^{\infty} \psi_{kt_0}^* \Delta h_{t-k}. \quad (4.45)$$

In this case,  $\Psi_{t_0} < 1$  does not imply  $\Delta h_t(h_{t-1} - \bar{h}_0) < 0$  since the sign of  $\Delta h_t$  also depends on  $R_t^0$ . For the case  $\bar{h}_0 < h_{t-1}$ , consider the circumstances in which  $R_t^0 > 0$ . Rearrangement of the sum (4.45) leads to:

$$\begin{aligned} R_t^0 &= -\theta_2 z_{t_0-2}^2 (h_{t-1} - h_{t-2}) - \theta_3 z_{t_0-3}^2 (h_{t-1} - h_{t-3}) - \dots \\ &= -\sum_{k=2}^{\infty} \theta_k z_{t_0-k}^2 (h_{t-1} - h_{t-k}), \end{aligned} \quad (4.46)$$

so that a necessary condition for  $R_t^0 > 0$  is that  $h_{t-1} < h_{t-k}$  for at least one value of  $k > 1$ . This shows that with  $\Psi_{t_0} < 1$  a sequence  $\{h_t\}$  generated by (4.44) can never diverge, and is almost surely finite. Conversely, if  $\bar{h}_0 > h_{t-1}$  the necessary condition for  $R_t^0 < 0$  is  $h_{t-1} > h_{t-k}$  for at least one  $k > 1$ , although this case is not critical to the property  $P(h_t < \infty) = 1$ . This completes the proof. ■

## 4.8.2 Proof of Proposition 4.3.2

**Proof of Proposition 4.3.2.** The solution of (4.25) is:

$$h_t^* = \omega \left( 1 + \sum_{m=1}^{\infty} \prod_{k=0}^{m-1} \Psi_{t-k} \right).$$

Since  $\sum_{j=0}^{\infty} \theta_j < \infty$  and the sequence  $\{\sum_{j=1}^m \theta_j z_{t-j}^2, m \geq 1\}$  is monotone,  $\Psi_t$  is a measurable function of  $\{z_s, -\infty < s < t\}$  by (e.g.) Davidson (1994), Theorems 3.25 and 3.26. The sequence  $\{\Psi_t, -\infty < t < \infty\}$  is therefore strictly stationary

and ergodic<sup>5</sup>. It follows by the ergodic theorem that:

$$\frac{1}{m} \sum_{k=0}^{m-1} \ln \Psi_{t-k} \xrightarrow{a.s.} E(\ln \Psi_t) = \zeta.$$

Hence, with probability one,

$$\limsup_{m \rightarrow \infty} e^{-m\zeta} \prod_{k=0}^{m-1} \Psi_{t-k} < \infty,$$

for  $-\infty < t < \infty$ . There therefore exists  $N < \infty$  such that:

$$h^* = h_{1t}^* + O(e^{N\zeta}),$$

with probability 1, where:

$$h_{1t}^* = \omega \left( 1 + \sum_{m=1}^N \prod_{k=0}^{m-1} \Psi_{t-k} \right).$$

The remainder term can be made as small as desired by taking  $N$  large enough, and  $h_{1t}^*$  is a measurable function of  $\{z_s, -\infty < s < t\}$  by (e.g.) Davidson (1994) Theorem 3.25. Strict stationarity and ergodicity of  $\{h_t^*, -\infty < t < \infty\}$  follows, completing the proof. ■

---

<sup>5</sup>Nelson (1990) cites Theorem 3.5.8 of Stout (1974) in support of a comparable assertion to this one. While the conditions do not precisely correspond, Phillips (1988) Section 1.15 provides a concise proof for the general case of doubly infinite sequences.



## CHAPTER 5

### CONCLUSION

The exponential conditional heteroscedasticity-type volatility models have been widely applied in financial time series analysis. However, there are some mysteries surrounding the statistical properties of these models. This thesis mainly investigated the memory and persistence of the EGARCH-type and ARCH( $\infty$ ) processes and the asymptotic theory of the QMLE in the ARMA models with EGARCH-type errors. This chapter presents the contributions and practical implications of this study, and also some suggestions for further research.

In this thesis, both Chapter 2 and Chapter 4 investigated the memory and persistence of the relevant (exponential) conditional heteroscedasticity-type models. The Chapter 2 explored the memory properties of the EGARCH-type models by applying the concept of NED on an independent process. It demonstrated that the EARCH( $\infty$ ) process can capture geometric memory, hyperbolic memory and even long memory under certain conditions, and proposed the HYEGARCH and FIEGARCH(DL) models. Among the EGARCH-type models, the EGARCH( $p, q$ ) process has geometric memory, the HYEGARCH process has hyperbolic memory, and the FIEGARCH(DL) process can capture long memory in volatility. Chapter 4 investigated the persistence of stationary and nonstationary ARCH( $\infty$ ) processes. It provided a simple sufficient condition for strict stationarity in the ARCH( $\infty$ )-type models by applying the BN decomposition and proposed the  $J_t$ -statistic to measure the persistence of the ARCH( $\infty$ ) process. Here, persistence is defined in terms of the (in)frequency of crossings of the median in successive steps in this study. A process is necessarily nonstationary if it crosses the median at most a finite number of times in a realisation of length  $T$ , as  $T$  tends to  $\infty$ ; it is a se-

rially independent process if it crosses the median with probability  $1/2$  at each step, by construction. Simulations were applied to investigate the persistence, covariance stationary, strictly stationary and nonstationary cases of the ARCH-type processes. The results showed that persistence depends on both the properties of the lag coefficients of the relevant models and the distribution of the driving shocks.

The research findings in Chapter 2 and Chapter 4 have several important contributions to make to both theoretical and empirical research. First, an important implication of the results is their ability to distinguish the persistence properties between conditional mean and conditional variance equations. The literature on the persistence of the conditional heteroscedasticity models is confused since several conditional heteroscedasticity models have been introduced to capture high persistence in volatility by following a similar idea to the integrated and fractional models in levels. However, the persistence properties of conditional variance and conditional mean equations are not parallel. Practitioners should correctly understand the memory properties of the model which they want to adopt. In addition, Chapter 4 also provided a new persistent measure  $J_t$ -statistic which can be widely used by practitioners to measure the memory properties of the square returns and volatility processes. Therefore, this study is useful to help econometricians and practitioners to determine the persistence properties of the conditional mean and conditional variance models. Further, econometricians and practitioners can select an appropriate model to capture the high persistence in volatility and understand how long the effect of shocks on volatility can persist.

Moreover, Chapter 2 investigated the limit theorems of the EGARCH-type models by applying the concept of NED on an independent process. More specifi-

cally, this chapter established the FCLT for the partial sum process of the EGARCH model under similar condition to the study of Lee (2014) and also constructed the FCLT and fractional FCLT for the partial sum of the processes  $\{\ln h_t - \omega\}$  in the HYGARCH and FIEGARCH processes, respectively.

Research on the FCLT or fractional FCLT for the partial sum of the relevant return or volatility processes has important applications in economics and financial data analysis. Firstly, these research findings contribute to the econometric literature on the limit theorems of exponential conditional heteroscedasticity-type models and the application of the concept of NED on an independent process. Secondly, these results are helpful for answering some interesting empirical questions. For example, does standard inference work for the return series, which have highly persistent volatility, and log-volatility processes, and do these processes satisfy the FCLT or fractional FCLT? In other words, can we do conventional tests when we have a large amount of conditional volatility involved? The answers to these questions are important for empirical studies. It is well known that the return processes are usually uncorrelated, but the square or the absolute value of the returns and volatility processes are highly persistent. Establishing the FCLT and fractional FCLT of the relevant return and volatility processes would be useful for answering these empirical questions. Thirdly, the establishments of FCLTs or fractional FCLTs are normally required for the statistical inference in the conditional heteroscedasticity-type models. For instance, in empirical studies, the CUSUM (cumulative sum) or MOSUM (moving sum) statistics are widely used for detecting change-point in the mean and volatility conditional heteroscedasticity models, see e.g. Hörmann (2008), Andreou and Ghysels (2009), and Lee (2014). The (fractional) FCLTs in the partial sums of the relevant processes of conditional heteroscedasticity-type models provide a foundation for establishing the asymp-

otic distribution of the CUSUM or MOSUM statistics. Also, the FCLTs are required for establishing the asymptotic distribution of the Dickey-Fuller (DF) statistics for the unit root test in an AR model with conditional heteroscedasticity errors, see e.g. Andreou and Ghysels (2009) and Lee (2014).

Furthermore, Chapter 3 focused on dealing with parameter estimation in the ARMA model with the EGARCH-type errors. The QMLE is one of the most popular estimation methods for the conditionally heteroscedastic time series models. However, the asymptotic properties of the QMLE in these models have not been fully explored. To investigate the asymptotic theory of the QMLE in EGARCH-type processes, the invertibility is an important issue. Thus, Chapter 3 first investigated the invertibility of the EGARCH(1, 1) error by using the methods of Martinet and McAleer (2015), and established the CAN of the QMLE in the ARMA(1, 1)-EGARCH(1, 1) processes. This chapter also proved the consistency of the QMLE in the HYEGARCH process with the assumption that the invertibility can be satisfied in this process, and provided an investigation into the asymptotic normality of the QMLE in the HYEGARCH process and the asymptotic properties of the QMLE in the FIEAGARCH(DL) model.

Research on the asymptotic theory of the QMLE in the ARMA models with EGARCH-type errors in Chapter 3, is essential for statistical inferences on the parameters in these models. This study is important for developing the theoretical properties of these models and improving the econometric literature on parameter estimation in ARMA models with exponential conditional heteroscedasticity errors. On the other hand, it is also important for empirical research. These models have been widely used in analysing the volatility of financial data. To investigate the returns and volatility of financial time series by using these models, it is essential

to establish the asymptotic theory of the QMLE in these models and to make statistical inferences. Practitioners should be clear whether the estimation method can provide consistency and asymptotic normality estimators for these parameters or not. Although the asymptotic properties of the QMLE in the HYGARCH and FIEGARCH processes have not been fully established, the Monte Carlo simulation results showed that the QMLEs are consistent estimators in these processes.

Last but not least, the fourth chapter also investigated forecasting methodologies in the ARCH( $\infty$ ) process. Alternative forecasting method was considered for volatility forecasting. This study argued that the conventional forecasting methodologies may overstate the degree of future volatility under the assumption of shocks with symmetrical distribution. This is, of course, an issue essentially of the preferred choice of the loss function. This finding is highly important for empirical studies of volatility forecasting. It is worth noting for practitioners that the minimising MAV is more appropriate than the minimum MSE. Practitioners also need to be aware that the latter criterion implicitly assumes a Gaussian framework, and is arguably inappropriate in the context of predicting volatility.

Based on these studies, there are several topics of interest for future studies. Firstly, for the statistical inference on the HYGARCH and FIEGARCH models, it is vital to understand the limit theorem of the return and conditional variance processes. However, the FCLT or fractional FCLT for the partial sum of these processes has not been fully established in this thesis. For further study, the exponential BM might be useful for establishing the limit theorems for the returns and conditional variance processes following HY/FIEGARCH models. Secondly, it is worth noting for future studies that the method of Martinet and McAleer (2015) can probably also be applied to derive the conditions for the invertibility

of HY/FIEGARCH processes. Meanwhile, based on the discussion of the asymptotic theory of the HY/FIEGARCH models in chapter 3, the method of Beran and Schützner (2009) is likely to be useful for establishing the asymptotic theory of the HY/FIEGARCH(DL) models. This is worth considering in further investigations. Moreover, following a similar procedure to Chapter 3, the asymptotic theory for ARMA( $r, s$ )-EGARCH( $p, q$ ) and vector ARMA-EGARCH models would also be worth investigating in further studies, since these models are becoming more and more popular. Thirdly, an investigation into volatility forecasting using the minimising MAV criterion is worth further consideration.

## BIBLIOGRAPHY

- Amemiya, T. (1985). *Advanced econometrics*. Harvard University Press.
- Andreou, E. and Ghysels, E. (2009). Structural breaks in financial time series. In Mikosch, T., Kreiß, J.-P., Davis, R. A., and Andersen, T. G., editors, *Handbook of financial time series*, pages 839–870. Springer.
- Andrews, D. W. (1984). Non-strong mixing autoregressive processes. *Journal of Applied Probability*, 21(4):930–934.
- Andrews, D. W. (1988). Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory*, 4(3):458–467.
- Aue, A., Berkes, I., and Horváth, L. (2006). Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli*, 12(4):583–608.
- Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73(1):5–59.
- Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30.
- Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman and Hall/CRC press, New York.
- Beran, J., Feng, Y., Ghosh, S., and Kulik, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer Berlin Heidelberg.
- Beran, J. and Schützner, M. (2009). On approximate pseudo-maximum likelihood estimation for LARCH-processes. *Bernoulli*, 15(4):1057–1081.

- Berkes, I., Hörmann, S., and Horváth, L. (2008). The functional central limit theorem for a family of GARCH observations with applications. *Statistics and Probability letters*, 78(16):2725–2730.
- Berkes, I., Horváth, L., and Kokoszka, P. (2003). GARCH processes: Structure and estimation. *Bernoulli*, 9(2):201–227.
- Beveridge, S. and Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics*, 7(2):151–174.
- Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley and Sons, Inc. New York.
- Billingsley, P. (1999). *Convergence of Probability Measures, second edition*. John Wiley and Sons, Inc. New York.
- Black, F. (1976). Studies of stock price volatility changes. In *Proceedings of the 1976 American Statistical Association, Business and Economical Statistics Section*, pages 177–181. American Statistical Association, Washington DC.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.
- Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis*, 9(2):121–131.
- Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). ARCH modeling in finance: a review of the theory and empirical evidence. *Journal of Econometrics*, 52(1):5–59.



- Bollerslev, T. and Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73(1):151–184.
- Bougerol, P. (1993). Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization*, 31(4):942–959.
- Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, 52(1):115 – 127.
- Breidt, F., Crato, N., and Lima, P. D. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83(1-2):325–348.
- Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory*, 18(1):17–39.
- Chung, K. L. (1968). *A course in probability theory*. New York: Harcourt Brace and World.
- Dahl, C. and Iglesias, E. M. (2008). The limiting properties of the QMLE in a general class of asymmetric volatility models. *Available at SSRN 1155404*.
- Davidson, J. (1992). A central limit theorem for globally nonstationary near-epoch dependent functions of mixing processes. *Econometric Theory*, 8(3):313–329.
- Davidson, J. (1993). The central limit theorem for globally nonstationary near-epoch dependent functions of mixing processes: the asymptotically degenerate case. *Econometric Theory*, 9(3):402–412.
- Davidson, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians*. Oxford University Press, UK.
- Davidson, J. (2000). *Econometric theory*. Blackwell.

- Davidson, J. (2002). Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. *Journal of Econometrics*, 106(2):243–269.
- Davidson, J. (2004). Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business and Economic Statistics*, 22(1):16–29.
- Davidson, J. (2006). Asymptotic methods and functional central limit theorems. In Mills, T. C. and Patterson, K., editors, *Palgrave Handbooks of Econometrics*. Palgrave-Macmillan.
- Davidson, J. (2009). When is a time series  $I(0)$ ? In Castle, J. and Shephard, N., editors, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, chapter 13, pages 322–342. Oxford University Press, Oxford.
- Davidson, J. and De Jong, R. M. (2000). The functional central limit theorem and weak convergence to stochastic integrals II: fractionally integrated processes. *Econometric Theory*, 16(5):643–666.
- Davidson, J. and Li, X. (2015). Strict stationarity, persistence and volatility forecasting in ARCH ( $\infty$ ) processes. *Journal of Empirical Finance*, <http://dx.doi.org/10.1016/j.jempfn.2015.08.010>.
- Davydov, Y. A. (1970). The invariance principle for stationary processes. *Theory of Probability & Its Applications*, 15(3):487–498.
- De Jong, R. M. and Davidson, J. (2000). The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent processes. *Econometric Theory*, 16(5):621–642.

- Deb, P. (1996). Finite sample properties of maximum likelihood and quasi-maximum likelihood estimators of EGARCH models. *Econometric Reviews*, 15(1):51–68.
- Demos, A. and Kyriakopoulou, D. (2010). Bias correction of ML and QML estimators in the EGARCH(1, 1) model. *Preprint*.
- Demos, A. and Kyriakopoulou, D. (2013). Asymptotic normality of the QML estimator of the EGARCH(1,1) model. *Available at SSRN-id2236055, accessed on 24 August 2015*.
- Ding, Z. and Granger, C. W. (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics*, 73(1):185–215.
- Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.
- Donsker, M. D. (1951). An invariance principle for certain probability limit theorems. *Memoris of the American Mathematical Society*, (6):1–12.
- Douc, R., Roueff, F., and Soulier, P. (2008). On the existence of some ARCH( $\infty$ ) processes. *Stochastic Processes and Their Applications*, 118(5):755–761.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.
- Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5(1):1–50.
- Eviews (2013). *Eviews8 User Guide*. IHS Global Inc. Irvine, CA.

- Francq, C. and Zakořian, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637.
- Francq, C. and Zakořian, J.-M. (2009). A tour in the asymptotic theory of GARCH estimation. In Mikosch, T., Kreiß, J.-P., Davis, R. A., and Andersen, T. G., editors, *Handbook of Financial Time Series*. Springer.
- Gallant, A. R. and White, H. (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Basil Blackwell, Oxford.
- Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4):221–238.
- Giraitis, L., Kokoszka, P., and Leipus, R. (2000a). Stationary ARCH models: Dependence structure and central limit theorem. *Econometric Theory*, 16(1):3–22.
- Giraitis, L., Koul, H. L., and Surgailis, D. (2012). *Large sample inference for long memory processes*. Imperial College Press, London.
- Giraitis, L., Leipus, R., Robinson, P. M., and Surgailis, D. (2004). LARCH, leverage, and long memory. *Journal of Financial Econometrics*, 2(2):177–210.
- Giraitis, L., Leipus, R., and Surgailis, D. (2007). Recent advances in ARCH modelling. In Teysnière, G. and Kirman, A. P., editors, *Long Memory in Economics*, pages 3–38. Springer.
- Giraitis, L., Leipus, R., and Surgailis, D. (2009). ARCH ( $\infty$ ) models and long memory properties. In Mikosch, T., Kreiß, J.-P., Davis, R. A., and Andersen, T. G., editors, *Handbook of Financial Time Series*, pages 71–84. Springer Berlin Heidelberg.

- Giraitis, L., Robinson, P. M., and Surgailis, D. (2000b). A model for long memory conditional heteroscedasticity. *The Annals of Applied Probability*, 10(3):1002–1024.
- Giraitis, L. and Surgailis, D. (2002). ARCH-type bilinear models with double long memory. *Stochastic Processes and their Applications*, 100(1-2):275–300.
- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801.
- Granger, C. and Andersen, A. (1978). On the invertibility of time series models. *Stochastic Processes and their Applications*, 8(1):87–92.
- Granger, C. W. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14(2):227–238.
- Granger, C. W. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29.
- Grublytė, I., Surgailis, D., and Škarnulis, A. (2015). Quasi-MLE for quadratic ARCH model with long memory. *arXiv preprint arXiv:1509.06422*.
- Hall, P. and Heyde, C. (1980). *Martingale Limit Theory and Applications*. Academic Press, New York.
- Halunga, A. G. and Orme, C. D. (2009). First-order asymptotic theory for parametric misspecification tests of GARCH models. *Econometric Theory*, 25(2):364–410.
- Hansen, B. (1991). GARCH(1, 1) processes are near epoch dependent. *Economics Letters*, 36(2):181–186.

- Harvey, A. C. (2002). Long memory in stochastic volatility. In Satchell, S. and Knight, J., editors, *Forecasting Volatility in the Financial Markets, Second edition*, pages 307–320. Butterworth-Heinemann.
- He, C., Malmsten, H., and Teräsvirta, T. (2008). Higher-order dependence in the general power ARCH process and the role of power parameter. In *Recent Advances in Linear Models and Related Areas: Essays in Honour of Helge Toutenburg*, pages 231–251. Physica-Verlag HD, Heidelberg.
- He, C. and Teräsvirta, T. (1999a). Fourth moment structure of the GARCH (p,q) process. *Econometric Theory*, 15(6):824–846.
- He, C. and Teräsvirta, T. (1999b). Properties of moments of a family of GARCH processes. *Journal of Econometrics*, 92(1):173–192.
- He, C., Teräsvirta, T., and Malmsten, H. (2002). Moment structure of a family of first-order exponential GARCH models. *Econometric Theory*, 18(4):868–885.
- Hörmann, S. (2008). Augmented GARCH sequences: dependence structure and asymptotics. *Bernoulli*, 14(2):543–561.
- Hosking, J. R. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.
- Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theory of Probability and Its Applications*, 7(4):349–382.
- Johansen, S. and Nielsen, M. Ø. (2012). A necessary moment condition for the fractional functional central limit theorem. *Econometric Theory*, 28(3):671–679.
- Karanasos, M. and Kim, J. (2003). Moments of the ARMA–EGARCH model. *Econometrics Journal*, 6(1):146–166.

- Kazakevičius, V. and Leipus, R. (2002). On stationarity in the ARCH ( $\infty$ ) model. *Econometric Theory*, 18(1):1–16.
- Kazakevičius, V. and Leipus, R. (2003). A new theorem on the existence of invariant distributions with applications to ARCH processes. *Journal of Applied Probability*, 40(1):147–162.
- Kazakevičius, V., Leipus, R., and Viano, M.-C. (2004). Stability of random coefficient ARCH models and aggregation schemes. *Journal of Econometrics*, 120(1):139–158.
- Kesten, H. and Spitzer, F. (1984). Convergence in distribution of products of random matrices. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(4):363–386.
- Kokoszka, P. and Leipus, R. (2000). Change-point estimation in ARCH models. *Bernoulli*, 6(3):513–539.
- Kyriakopoulou, D. (2015). Asymptotic normality of the QML estimator of the EGARCH(1,1) model. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2236055>.
- Lee, O. (2013). The functional central limit theorem for ARMA–GARCH processes. *Economics Letters*, 121(3):432–435.
- Lee, O. (2014). Functional central limit theorems for augmented GARCH(p,q) and FIGARCH processes. *Journal of the Korean Statistical Society*, 43(3):393–401.
- Lee, S.-W. and Hansen, B. E. (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory*, 10(1):29–52.

- Ling, S. and Li, W. K. (1997). On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. *Journal of the American Statistical Association*, 92(439):1184–1194.
- Ling, S. and McAleer, M. (2002). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, 106(1):109–117.
- Ling, S. and McAleer, M. (2003). Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory*, 19(2):pp. 280–310.
- Lopes, S. R. and Prass, T. S. (2014). Theoretical results on fractionally integrated exponential generalized autoregressive conditional heteroskedastic processes. *Physica A: Statistical Mechanics and its Applications*, 401:278–307.
- Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. *Econometrica*, 64(3):575–596.
- Marinucci, D. and Robinson, P. M. (2000). Weak convergence of multivariate fractional processes. *Stochastic Processes and their applications*, 86(1):103–120.
- Martinet, G. G. and McAleer, M. (2015). On the invertibility of EGARCH( $p, q$ ). Technical report, Tinbergen Institute Discussion Paper.
- McLeish, D. L. (1975). A maximal inequality and dependent strong laws. *The Annals of probability*, 3(5):829–839.
- Mikosch, T. and Stărică, C. (2003). Long-range dependence effects and ARCH modeling. In Doukhan, P., Oppenheim, G., and Taqqu, M. S., editors, *Theory and applications of long-range dependence*, pages 439–460. Birkhauser, Boston.



- Mikosch, T. and Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *The Review of Economics and Statistics*, 86(1):378–390.
- Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, 6(3):318–334.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370.
- Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*, 10(2):229–235.
- Perez, A. and Zaffaroni, P. (2008). Finite-sample properties of maximum likelihood and whittle estimators in EGARCH and FIEGARCH models. *Quantitative and Qualitative Analysis in Social Sciences*, 2(1):78–97.
- Phillips, P. C. (1988). *Lectures on Stationary and Nonstationary Time Series*. Institute of Advanced Studies, Vienna. (Text at <http://korora.econ.yale.edu/phillips/teach/notes/1988-lectures.pdf>).
- Phillips, P. C. and Solo, V. (1992). Asymptotics for linear processes. *The Annals of Statistics*, 20(2):971–1001.
- Poon, S.-H. (2005). *A Practical Guide to Forecasting Financial Market Volatility*. John Wiley & Sons Ltd., Chichester.
- Prass, T. S. (2008). Análise e estimação de medidas de risco em processos FIEGARCH, Universidade Federal do Rio Grande do Sul.
- Robinson, P. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, 47(1):67–84.

- Ruiz, E. and Veiga, H. (2008). Modelling long-memory volatilities with leverage effect: A-LMSV versus FIEGARCH. *Computational Statistics & Data Analysis*, 52(6):2846–2862.
- Saadi, S., Gandhi, D., and Dutta, S. (2006). Testing for nonlinearity & modeling volatility in emerging capital markets: the case of Tunisia. *International Journal of Theoretical and Applied Finance*, 9(7):1021–1050.
- Schützner, M. (2009). *Asymptotic Statistical Theory for Long Memory Volatility Models*. PhD thesis, University Konstanz.
- Stock, J. H. (1994). Unit roots, structural breaks and trends. In Engle, R. and McFadden, D., editors, *Handbook of econometrics*, volume 4, pages 2740–2843. Amsterdam: Elsevier.
- Stout, W. F. (1974). *Almost sure convergence*. Academic Press, New York.
- Straumann, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models*. Lecture Notes in Statistics 181. Springer Berlin Heidelberg, New York.
- Straumann, D. and Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics*, 34(5):2449–2495.
- Surgailis, D. and Viano, M.-C. (2002). Long memory properties and covariance structure of the EGARCH model. *ESAIM: Probability and Statistics*, 6:311–329.
- Taqqu, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31(4):287–302.

- Tong, H. (1993). *Non-Linear Time Series, A Dynamical System Approach*. Oxford Science Publications, Oxford.
- von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the  $r$ th absolute moment of a sum of random variables,  $1 \leq r \leq 2$ . *The Annals of Mathematical Statistics*, 36(1):299–303.
- Weiss, A. A. (1986). Asymptotic theory for ARCH models: Estimation and testing. *Econometric Theory*, 2(1):107–131.
- White, H. (1984). *Asymptotic Theory for Econometricians*. Academic press, New York.
- Wintenberger, O. (2013). Continuous invertibility and stable QML estimation of the EGARCH(1,1) model. *Scandinavian Journal of Statistics*, 40(4):846–867.
- Wooldridge, J. M. and White, H. (1988). Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory*, 4(2):210–230.
- Wu, W. B. and Shao, X. (2006). Invariance principles for fractionally integrated nonlinear processes. In *Recent Developments in Nonparametric Inference and Probability*, volume 50, pages 20–30. Institute of Mathematical Statistics.
- Zaffaroni, P. (2009). Whittle estimation of EGARCH and other exponential volatility models. *Journal of Econometrics*, 151(2):190–200.
- Zakořian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955.