

Statistical Inference for Lévy-Driven Ornstein-Uhlenbeck Processes

Ibrahim Abdelrazeq

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Mathematics ¹

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

© Ibrahim Abdelrazeq, Ottawa, Canada, 2014

¹The Ph.D. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Abstract

When an Ornstein-Uhlenbeck (or CAR(1)) process is observed at discrete times $0, h, 2h, \dots, [T/h]h$, the unobserved driving process can be approximated from the observed process. Approximated increments of the driving process are used to test the assumption that the process is Lévy-driven. Asymptotic behavior of the test statistic at high sampling frequencies is developed assuming that the model parameters are known. The behavior of the test statistics using an estimated parameter is also studied. If it can be concluded that the driving process is Lévy, the empirical process of the approximated increments can then be used to carry out more precise tests of goodness-of-fit. For example, one can test whether the driving process can be modeled as a Brownian motion or a gamma process. In each case, performance of the proposed test is illustrated through simulation.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Prof. Gail Ivanoff and Dr. Rafal Kulik. Prof. Ivanoff, who has been my guiding beacon, for all of her help, guidance and leadership. I thank her for the patience, effort, and encouragement that she so graciously put forth to help me in this endeavor. I would also like to thank Dr. Kulik. His innovative insights inspired my research. His positive thinking, discussions, ideas, and feedback were undeniably vital.

Our weekly meetings were invaluable throughout this process. Their support, patience, valuable comments and suggestions helped me to build, enhance and complete this work. I feel privileged by their mentorship. Their dedication, direction and important advice helped me to develop an experience consistent with my long term research and academic goals.

I would also like to thank the rest of the faculty of the Ottawa-Carleton Institute of Mathematics and Statistics as well as my fellow students and colleagues for all of their support and encouragement over the course of my doctoral studies at the University of Ottawa.

I would like to give my special thanks to my wife Maha whose love and care enabled me to complete this work.

The financial support of the Department of Mathematics and Statistics and the University of Ottawa are gratefully acknowledged.

Dedication

To my mother and my father I dedicate this work.

Contents

List of Figures	viii
1 Introduction	1
2 Preliminaries	7
2.1 Second-Order Stationarity and Strict Stationarity	8
2.2 Lévy processes	9
2.2.1 Examples of Second-Order Lévy Processes	10
2.3 Example of a Second-Order Non-Lévy Processes	12
2.4 CAR(1) models	13
2.5 The Sampled CAR(1) Process	17
2.6 Recovering the driving process	19
2.7 Approximation of Lévy increments using the inversion formula	20
3 Asymptotics for the sampled process	24
3.1 Discrete approximation of L	25
3.2 Case (I): h is fixed, $T = Nh$, and $N \rightarrow \infty$	27
3.3 Case (II) and Case (III): $h = 1/M$, and $M \rightarrow \infty$	32
3.3.1 Approximation of Y and L	33
3.3.2 Estimated unit increments	36
3.3.3 Asymptotics for the Sample Mean	40

3.3.4	Asymptotics for the Sample Covariances	42
3.4	CLT for the Lévy-driven CAR(1) process	47
4	Inference based on the sampled process	49
4.1	Test Statistics	49
4.2	Simulation Study	50
4.2.1	Brownian motion driven CAR(1) process	50
4.2.2	Gamma-driven CAR(1) process	56
4.2.3	Poisson-driven CAR(1) process	60
4.2.4	An alternative case	65
5	Asymptotics for the sampled process with estimated parameter	69
5.1	Estimation of a	70
5.2	Test statistics: asymptotic normality	73
5.2.1	Preliminary lemmas	74
5.2.2	Asymptotics for Sample Mean and Sample Variance	82
5.2.3	Asymptotics for Sample Covariances	84
5.2.4	Proof of Theorem 5.2.1	88
5.3	Simulation Studies	89
5.3.1	Brownian motion driven CAR(1) process	89
5.3.2	Gamma-driven CAR(1) process	90
5.3.3	An alternative case	91
6	Goodness-of-Fit Test	96
6.1	Introduction and Main Results	97
6.2	Proof of Theorem 6.1.1	100
6.3	Test Statistics	111
6.3.1	Kolmogorov-Smirnov Test	111

7	Brief Summary and Extensions	118
7.1	Future Work and Extensions	118

List of Figures

4.1	Simulated standard Brownian motion for $N = 100$ and $K = 5000$	51
4.2	Simulated Brownian motion driven CAR(1) process Y for $\{N = 100, K = 5000, \mu = 0, \eta = 1, \sigma = 1, a = 1\}$ and Y_{Sampled} for $M = 500$	52
4.3	True Increments: $\Delta_1 B_n$, Recovered Increments: $\Delta_1 \widehat{B}_n^{(M)}$	53
4.4	(True-Recovered) Increments: $\Delta_1 B_n - \Delta_1 \widehat{B}_n^{(M)}$	53
4.5	The sample autocorrelation function for the True Increments $\Delta_1 B_n$ (True) and for the Recovered Increment $\Delta_1 \widehat{B}_n^{(M)}$ (Recovered)	54
4.6	Simulated gamma Process for $\{N = 100, K = 5000, \mu = 1, \eta = 1\}$	56
4.7	Simulated gamma-driven CAR(1) process for $\{N = 100, K = 5000, \mu = 1, \eta = 1, \sigma = 1, a = 1\}$ and Y_{Sampled} for $M = 500$	57
4.8	True Increments: $\Delta_1 G_n$, Recovered Increments: $\Delta_1 \widehat{G}_n^{(M)}$	57
4.9	(True-Recovered) Increments: $\Delta_1 G_n - \Delta_1 \widehat{G}_n^{(M)}$ ($M = 500$)	58
4.10	The sample autocorrelation function for the True Increments $\Delta_1 G_n$ (True) and for the Recovered Increments $\Delta_1 \widehat{G}_n^{(M)}$ (Recovered).	58
4.11	(True-Recovered) Increments with and without peaks for small $M = 100$	60
4.12	Simulated Poisson Process for $\{N = 100, K = 5000, \lambda = 1\}$	61
4.13	Simulated gamma-driven CAR(1) process for $\{N = 100, K = 5000, \lambda = 1, \sigma = 1, a = 1\}$ and Y_{Sampled} for $M = 500$	62
4.14	True Increments: $\Delta_1 P_n$, Recovered Increments: $\Delta_1 \widehat{P}_n^{(M)}$ ($M = 500$)	62

4.15 (True-Recovered) Increments: $\Delta_1 P_n - \Delta_1 \widehat{P}_n^{(M)}$ ($M = 500$)	63
4.16 The sample autocorrelation function for the True Increments $\Delta_1 P_n$ (True) and for the Recovered Increments $\Delta_1 \widehat{P}_n^{(M)}$ (Recovered).	63
4.17 (True-Recovered) Increments with and without peaks for small $M =$ 100	64
4.18 Simulated fractional Brownian motion for $N = 100$, $K = 5000$ and $H = 0.8$	65
4.19 Simulated fractional Brownian motion driven CAR(1) process for $N = 100$, $K = 5000$ and Y_{Sampled} for $M = 500$	66
4.20 True Increments: $\Delta_1 B_{H,n}$, Recovered Increments: $\Delta_1 \widehat{B}_{H,n}^{(M)}$	66
4.21 (True-Recovered) Increments: $\Delta_1 B_{H,n} - \Delta_1 \widehat{B}_{H,n}^{(M)}$	67
4.22 The sample autocorrelation function for the True Increments $\Delta_1 B_{H,n}$ (True) and for the Recovered Increment $\Delta_1 \widehat{B}_{H,n}^{(M)}$ (Recovered).	67
4.23 The power function for the test based on $W_{\Delta_1 B_H}(1)$ is the solid line and the one based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the dashed line	68
5.1 The power function for the test based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the solid line and the one based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the dashed line. Fix $a = 5$	92
5.2 The power function for the test based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the solid line and the one based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the dashed line. Fix $H = 0.8$	92
5.3 The power function for the test based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the solid line and the one based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the dashed line. Fix $H = 0.65$	95
6.1 The power functions for the test based on $KS_{\Delta_1 G_n}$ is the solid line and the one based on $KS_{\Delta_1 \widehat{G}_n^{(M)}}$ is the dashed line	115

6.2 Histograms of the Kolmogorov-Smirnov test values based on the true gamma process increments and parameters and the one that based on the estimated gamma process increments and bootstrapped estimators respectively. $R = 1000$, $N = 100$, $M = 500$, $\eta = 1$ and $\mu = 10$	117
--	-----

Chapter 1

Introduction

In keeping with the Nobel prize winning work of Black, Scholes [8] and Merton [36] that proposed a continuous-time stochastic model for option pricing, researchers have long recognized that continuous-time models are needed to represent the reality of the economy and financial markets. A continuous time model reflects the natural evolution of a process indexed by an interval $I \subseteq \mathbb{R}_+$ and the choice of an appropriate stochastic model is essential. Thus, it is critical to test how well the proposed model represents the observed behavior of the process - in other words, we must assess the “goodness-of-fit” of our model. However, in reality most frequently the continuous time process can only be observed at discrete times. Reconciling the discrete data with the continuous model is the principal motivation for this work. The recent availability of high-frequency (or tick-by-tick transaction) data of various financial markets allows us to closely approximate the behavior of the continuous time process.

Univariate continuous-time autoregressive moving average (CARMA) processes are the continuous time analogue of the widely employed discrete-time ARMA process. CARMA(p, q) processes are the solutions of linear stochastic differential equations of

the form

$$D^p Y(t) + a_1 D^{p-1} Y(t) + \dots + a_p Y(t) = b_0 DL(t) + b_1 D^2 L(t) + \dots + b_q D^{q+1} L(t). \quad (1.0.1)$$

They were introduced in [20] in a Gaussian setting and generalized in [9] to include Lévy driving processes. Further extensions include multivariate CARMA or fractionally integrated CARMA models (see e.g. [35] and [13]).

The probabilistic properties of CARMA processes have received considerable attention. However, there has been little development in statistical inference for such models or in particular, goodness-of-fit. This thesis takes the first steps in developing rigorous statistical techniques for assessing goodness-of-fit of CARMA models, and complements recent work by Brockwell and Schlemm [14]. As pointed out in [14], if one decides to model a continuous time process using the CARMA framework, three main problems arise: a) the choice of the orders p and q ; b) estimation of the model coefficients a_i and b_j ; c) choosing an appropriate model for the Lévy driving process.

Here we will focus on analyzing b) and c) for the CARMA(1,0) (equivalently, CAR(1)) process. Because (under general assumptions) the CARMA(p, q) process can be expressed as a sum of dependent CAR(1) processes (cf. [12]), we expect that in the future, our results will be useful in analyzing the more general model.

Formally, the Lévy-driven Ornstein-Uhlenbeck process (equivalently CAR(1) or CARMA(1,0) process, see (1.0.1)) is the stationary process, Y , that satisfies the stochastic differential equation

$$dY(t) = -aY(t)dt + \sigma dL(t), \quad a, \sigma > 0. \quad (1.0.2)$$

This model was proposed by Barndorff-Nielsen and Shephard [4] as a continuous time stochastic volatility model (also known as the BNS model) by replacing the Brownian motion driving process (equivalently, noise) in the classical Ornstein-Uhlenbeck

model (c.f. [43]) by a more general Lévy process L . Lévy-driven CAR(1) processes not only allow one to consider non-Gaussian driving processes such as the gamma process, but can also incorporate discontinuities, or jumps, in the stochastic volatility models considered in many areas of application see ([26],[30], [31], [33],[34], for example). This model has become a very popular way to describe moderate and high frequency financial data, see ([17], [5], [4] and [3]).

There are several papers that discuss estimation of the coefficients for CAR(1) models; see e.g. [11] and [44] and the references therein. In [11] and [14] the authors address the third issue, namely estimation of the parameters of a specified family of Lévy processes, assuming that the order and the coefficients of the model are known. In [28] the authors consider nonparametric estimation of the underlying Lévy measure. Their method utilizes the Markovian structure of the CAR(1) process and its mixing properties. Sample characteristics (sample mean, sample autocovariances) of the discretely sampled process Y are examined in [19] for a general Lévy driven moving average model that includes the CAR(1) process studied here.

However, before one selects a parametric family of Lévy processes and/or estimates the model coefficients, one should verify whether it is reasonable to assume that the driving process is Lévy. From the point of view of exploratory data analysis, the first step would be to plot the sample covariances of the driving process at various lags. This procedure assumes a priori that the underlying driving process has finite second moment.

The driving process L is unobservable and cannot be directly recovered if the CAR(1) process Y is sampled at discrete times. Thus, as in [11], the driving process can only be estimated and inference must be performed with noisy data. Hence, the first topic addressed in this thesis is the development of statistical inference techniques for the

sample covariances of the (approximately) recovered driving process L , assuming that the second moments of the driving process are finite. (Note that this is in contrast to [19], where it is the sample covariances of Y that are of interest.) Initially, we assume that the model coefficients a and σ are known. We go beyond exploratory data analysis by providing a formal test of the hypothesis of uncorrelated increments. Our test statistic is shown to be asymptotically normal. En route, we prove several results of independent interest, including finding a precise bound on the approximation error of the unit increments of the recovered process, as well as providing an elementary proof of a central limit theorem for the integrated CAR(1) process Y .

Subsequently, we explore the performance of the test statistic when the model coefficient a is unknown and must be estimated. Without loss of generality, the parameter σ can be assumed to be one since it can be incorporated into a reparametrization of the Lévy noise. Due to the complex relationship between the parameter a and the (approximate) recovered increments of L , the choice of a suitable estimator is not straightforward. We will show the consistency and asymptotic normality of our proposed estimator and then demonstrate its effect on the asymptotic behaviour of the test statistic.

The final contribution of this thesis is the following: If the hypothesis of independent increments of the driving process of a CAR(1) model has not been rejected, we then consider a test of goodness-of-fit for the unobserved Lévy driving process. Using the empirical process defined by the estimated unit increments of the driving process L , we provide a test of the composite hypothesis that the driving process belongs to a specific class of Lévy processes, such as Brownian motion or gamma processes. There are two main challenges that arise here: first, the hypothesis is composite, and so estimators of the parameters defining the distribution of L must be substituted for the theoretical values in the usual empirical process. As is well known, even when

exact values of L are available, the resulting limiting distribution no longer leads to distribution-free tests. Second, we must use estimated values of L both to define the empirical distribution and to calculate estimators. The first issue is resolved by using a simple yet powerful technique proposed by Burke and Gombay in [16]: if the parameters of the distribution are estimated using a single bootstrap sample drawn from the original observations, the limit of the empirical process is the same as when the correct theoretical values of the parameters are used. The second problem can be resolved with a high sampling frequency.

We proceed as follows. After introducing some preliminaries in Chapter 2, we use the inversion formula of Pham [42], that represents the unobserved driving process L in terms of the continuously observed CAR(1) process Y . The same strategy was employed in [11] and in a multivariate setting in ([14], Theorem 4.3). Since Y is observed at discrete times, as noted above the driving process L cannot be recovered exactly and a trapezoidal approximation is used to replace an unobservable integral.

In Chapter 3, we are able to provide a uniform bound on the approximation error in the unit increments of the recovered driving process (see Lemmas 3.3.5 and 3.3.7 as well as Theorem 3.3.1). Our Lemma 3.3.5 can be compared to Theorem 5.7 in [14]. Although the result in the latter paper holds for more general multivariate CARMA models, our bound is more precise and is uniform with respect to N , the length of time that the process is observed. As a consequence, we can derive central limit theorems for partial sums and sample covariances (see Theorem 3.3.1 together with Corollary 3.3.3, and Theorem 3.3.11). In a brief digression, as an important by-product, we prove a central limit theorem for the integrated CAR(1) process Y (Theorem 3.4.2). The significance of this result is that the proof is quite elementary and does not require any mixing arguments.

We turn to the problem of verifying the assumption that the driving process L is Lévy in Chapter 4. We propose an appropriate test statistic to test the hypothesis that the driving process has uncorrelated increments and in Corollary 4.1.1 we prove that it is asymptotically $N(0, 1)$ under the null hypothesis. Several simulation studies illustrate the behaviour of the statistic under both the hypothesis and alternative.

In Chapter 5 we scrutinize the effect of using an estimated value of the model parameter a in the test statistics that were defined in Chapter 5. Theorem 5.2.6 exhibits somewhat different asymptotic behavior for the resulting test statistics. Once again we illustrate with simulations.

In Chapter 6, we propose more precise tests of goodness-of-fit based on the empirical processes of the estimated unit increments of L . Theorem 6.1.1 shows that the bootstrap technique of [16] leads to test statistics that are asymptotically distribution-free. We give simulation studies that illustrate the performance of the classic Kolmogorov-Smirnov test.

In the conclusion, Chapter 7, we briefly summarize our results and propose directions for future research.

Chapter 2

Preliminaries

In this chapter we introduce CAR(1) models and describe their elementary properties. In order to do this, we recall the notions of strict and second-order stationarity, as well as the definition of Lévy processes, followed by examples such as Brownian motion or gamma process (Sections 2.1-2.2). We also give an example of a non-Lévy process that will be used to show power of some tests (Section 2.3).

Next, we introduce second-order CAR(1) models (Section 2.4). The material presented there is taken from references such as [11]. However, we evaluate moment properties of the CAR(1) models for completeness.

In Section 2.5 we introduce a sampled process (with fixed frequency h). We show that the sampled process has an AR(1) structure. This is known in the literature (see [19]).

In Section 2.6 we introduce one of the most important tools of this thesis, namely the inversion formula that allows us to express the unobserved driving noise in terms of the observed CAR(1) process. The formula is taken from [42], [12]. However, we provide a proof for completeness. We note in passing that a different strategy to recover the unobserved driving process was employed in [23] for the general CARMA(p, q) model. Without assuming any particular values of p and q , increments of the Lévy driving

process are estimated via renormalized recovered noise from the Wold representation of the sampled CARMA sequence. The recovered process is shown to be L_2 -consistent under the assumption that the CARMA process is invertible.

Using the inversion formula, in Section 2.7, we approximate the unobserved increments of the driving process by discretely sampled CAR(1) process. We study the moment properties of the sampled increments and we show inconsistency in case when sampling frequency is fixed (see Proposition 2.7.1).

We remark that the contents of Sections 2.1-2.6 are based on existing literature ([10], [42], [11] and [38]), while Section 2.7 is new.

We finish with a comment regarding notation. In general, we use the notation $Y(t)$ when the time parameter t is continuous, and Y_t when the time parameter t is discrete.

2.1 Second-Order Stationarity and Strict Stationarity

We begin with an arbitrary T -indexed stochastic process $X \equiv \{X(t), t \in T\}$, where $T \equiv [0, \infty)$.

Definition 2.1.1 (The Autocovariance Function $\gamma_X(\cdot, \cdot)$). *If X is a process with a finite variance, we define the autocovariance function $\gamma_X(\cdot, \cdot)$ as*

$$\gamma_X(r, s) = \text{Cov}(X(r), X(s)) = \mathbb{E}[(X(r) - \mathbb{E}[X(r)])(X(s) - \mathbb{E}[X(s)])] \quad r, s \in T.$$

Definition 2.1.2 (Second-Order Stationarity). *The process X is said to be second-order stationary if*

- (i) $\mathbb{E}[X^2(t)] < \infty$ for all $t \in T$,
- (ii) $\mathbb{E}[X(t)] = m$ for all $t \in T$,
- (iii) $\gamma_X(r, s) = \gamma(r + t, s + t)$ for all $r, s, t \in T$.

Remark 2.1.3. If X is second-order stationary then $\gamma_X(r, s) = \gamma_X(r - s, 0)$ for all $r, s \in T$. Therefore, it is convenient to redefine the autocovariance function of a second-order stationary process as a function of only one variable:

$$\gamma(h) \equiv \gamma_X(h, 0) = \text{Cov}(X(t+h), X(t)) \quad \text{for all } t, h \in T.$$

Definition 2.1.4 (Strict Stationarity). *The process X is said to be strictly stationary if $(X(t_1), \dots, X(t_k))$ and $(X(t_1+h), \dots, X(t_k+h))$ have the same joint distribution for all integers $t_1, \dots, t_k \in T$, $k \geq 1$ and $h > 0$.*

Note that any strictly stationary sequence with a finite second moment is second-order stationary.

2.2 Lévy processes

Suppose we are given a stochastic base $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq \infty}, P)$, where \mathcal{F}_0 contains all the P -null sets of \mathcal{F} and (\mathcal{F}_t) is right-continuous.

Definition 2.2.1 (Lévy process). *A process $L \equiv \{L(t), t \geq 0\}$ is an (\mathcal{F}_t) -adapted Lévy process if $L(t) \in \mathcal{F}_t \forall t \geq 0$ and*

- $L(0) = 0$ a.s.,
- $L(t)$ has independent increments, i.e., $L(t) - L(s)$ is independent of \mathcal{F}_s , for any $0 \leq s < \infty$,
- $L(t)$ has stationary increments, i.e., $L(t+s) - L(s)$ has the same distribution as $L(t)$, for any $s, t > 0$,
- $L(t)$ is stochastically continuous, i.e. $\forall \epsilon > 0$ and $\forall t \geq 0$,

$$\lim_{s \rightarrow t} P(|L(t) - L(s)| > \epsilon) = 0,$$

- $L(t)$ has càdlàg (right continuous with left limits) sample paths.

Lévy-Khintchine representation

For a Lévy process L , we can characterize the distribution of $L(t)$ by its characteristic function,

$$\phi_{L(t)}(\theta) = \mathbb{E} [e^{i\theta L(t)}],$$

$\theta \in \mathbb{R}$, which satisfies the following relation:

$$\phi_{L(t)}(\theta) = \exp \left(i\theta mt - \frac{1}{2}\theta^2 b^2 t + t \int_{\mathbb{R}_0} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|<1}) \nu(dx) \right) \quad (2.2.1)$$

for some $m, b \in \mathbb{R}$, where ν is a measure defined over the Borel subsets of $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$ such that:

$$\int_{\mathbb{R}_0} \min(x^2, 1) \nu(dx) < \infty.$$

The measure ν is called the Lévy measure of the process L . A wealth of distributions for L is attainable by a suitable choice of the measure ν . The triplet (m, b, ν) is called the Lévy triplet. It describes the distribution of the Lévy process completely.

Second-Order Lévy Processes

Definition 2.2.2 (Second-Order Lévy Process). *We define L to be a second-order Lévy process if L is a Lévy process and $\mathbb{E}[L^2(1)] < \infty$. If $\mu = \mathbb{E}[L(1)]$ and $\eta^2 = \text{Var}[L(1)]$, then by the independence and stationarity of the increments of L we have*

$$\begin{aligned} \mathbb{E}[L(t)] &= \mu t, & t \geq 0, \\ \text{Var}(L(t)) &= \eta^2 t, & t \geq 0. \end{aligned} \quad (2.2.2)$$

2.2.1 Examples of Second-Order Lévy Processes

Brownian Motion

If ν in the Lévy triplet is the zero measure then equation (2.2.1) becomes:

$$\phi_{L(t)}(\theta) = \exp\left(i\theta mt - \frac{1}{2}\theta^2 b^2 t\right).$$

Then $L(t) \sim N(mt, b^2 t)$ and hence, L is a Brownian motion with drift. We denote L by B and we have $B(1) \sim N(m, b^2)$, and μ and η in equation (2.2.2) are m and b respectively.

Poisson Process

If we let $m = b = 0$ in the Lévy-Khintchine decomposition and $\nu = \lambda\delta(1)$ for $\lambda > 0$ where $\delta(1)$ denotes the Dirac measure with support on $\{1\}$, then equation (2.2.1) becomes

$$\phi_{L(t)}(\theta) = \exp(-\lambda t (1 - e^{i\theta})).$$

Hence, $L(t)$ follows a Poisson(λt) distribution. We denote L by P , $P(1) \sim \text{Poisson}(\lambda)$. In equation (2.2.2) we have $\mu = \eta^2 = \lambda$.

Gamma Process

If we consider the Lévy triplet with $b = 0$, $\nu(dx) = \alpha \frac{1}{x} e^{-\frac{x}{\beta}} \mathbf{1}_{\{x>0\}} dx$, and $m = \int_0^1 x\nu(dx)$, then equation (2.2.1) becomes:

$$\begin{aligned} \phi_{L(t)}(\theta) &= \exp\left(i\theta t \int_0^1 x\nu(dx) + t \int_{\mathbb{R}_0} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|<1}) \nu(dx)\right) \\ &= \exp\left(t \int_0^\infty (e^{i\theta x} - 1) \alpha \frac{1}{x} e^{-\frac{x}{\beta}} dx\right) = \exp\left(-\alpha t \int_0^\infty \frac{e^{-\frac{x}{\beta}} - e^{(i\theta - \frac{1}{\beta})x}}{x} dx\right) \\ &= \exp\left(-\alpha t \ln\left(\frac{i\theta - \frac{1}{\beta}}{-\frac{1}{\beta}}\right)\right) \\ &= \frac{1}{(1 - i\theta\beta)^{\alpha t}}, \end{aligned}$$

where in the third equality we used Frullani's integral, see §21.16 in [27].

Hence, $L(t)$ follows a $\Gamma(\alpha t, \beta)$ distribution and we denote L by G . Then in equation

(2.2.2) we have

$$\mu = \alpha\beta \quad \text{and} \quad \eta^2 = \alpha\beta^2 = \mu\beta.$$

Consequently,

$$\beta = \frac{\eta^2}{\mu} \quad \text{and} \quad \alpha = \frac{\mu^2}{\eta^2}$$

and we have

$$G(1) \sim \Gamma(\alpha, \beta) = \Gamma\left(\frac{\mu^2}{\eta^2}, \frac{\eta^2}{\mu}\right).$$

2.3 Example of a Second-Order Non-Lévy Processes

Fractional Brownian motion

A continuous-time Gaussian process $B_H \equiv \{B_H(t), t \geq 0\}$ is called fractional Brownian motion with Hurst parameter $H \in (0, 1)$ if it has zero mean and the covariance function

$$\text{Cov}(B_H(s), B_H(t)) = \frac{1}{2} (|t|^{2H} + |s|^{2H} - |t - s|^{2H}). \quad (2.3.1)$$

The fractional Brownian motion, B_H , has the following properties

- B_H has stationary increments, i.e., $B_H(t+s) - B_H(s)$ has the same distribution as $B_H(t)$, for any $s, t > 0$ which is clear from equation (2.3.1)
- The Hurst parameter, $H \in (0, 1)$, associated with the process B_H , describes the process in the sense that when $H = \frac{1}{2}$ then $B_{\frac{1}{2}}$ is a Brownian motion. When $H > \frac{1}{2}$ the increments of the process are positively correlated, and when $H < \frac{1}{2}$ the increments of the process are negatively correlated.

Hence $B_H, H \neq \frac{1}{2}$ is a second-order process ($\mathbb{E}[B_H^2(1)] < \infty$) which has stationary dependent increments.

2.4 CAR(1) models

In what follows, we assume that the process L is càdlàg with stationary increments.

Definition 2.4.1 (CAR(1) process). *A CAR(1) process Y driven by the process L is defined to be the solution of the stochastic differential equation*

$$dY(t) = -aY(t)dt + \sigma dL(t), \quad (2.4.1)$$

where $a, \sigma \in \mathbb{R}_+$ and $Y(0)$ is independent of L . We call the process L the driving process, and if L is a Lévy process then Y is called a Lévy-driven CAR(1) process.

The unique solution (cf. [38], Section 17) for equation (2.4.1) can be written as

$$Y(t) = e^{-at}Y(0) + \sigma \int_0^t e^{-a(t-u)} dL(u), \quad t \geq 0.$$

The function $f(u) = e^{-a(t-u)}$ is deterministic and continuously differentiable. Using an integration by parts formula we can define the CAR(1) process pathwise as:

$$Y(t) = e^{-at}Y(0) + \sigma L(t) - a\sigma \int_0^t e^{-a(t-u)} L(u) du, \quad t \geq 0. \quad (2.4.2)$$

Proposition 2.4.2 (Proposition 1, cf. [11]). *Let Y be a CAR(1) process driven by a second order Lévy process L . If $Y(0)$ is independent of $\{L(t), t > 0\}$, then Y is strictly stationary if and only if $a > 0$ and $Y(0)$ has the distribution of $\sigma \int_0^\infty e^{-au} dL(u)$.*

Lemma 2.4.3. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (2.2.2) holds. Then*

$$\mathbb{E}[Y(0)] = \frac{\mu\sigma}{a}, \quad \text{Var}(Y(0)) = \frac{\sigma^2\eta^2}{2a}, \quad \gamma_Y(s) \equiv \text{Cov}(Y(0), Y(s)) = \frac{\sigma^2\eta^2}{2a} e^{-as}. \quad (2.4.3)$$

Proof: Recall equation (2.4.2). Let

$$\mathcal{I}_t \equiv \int_0^t e^{au} dL(u) = e^{at}L(t) - a \int_0^t e^{au}L(u)du, \quad (2.4.4)$$

then

$$Y(t) = e^{-at}Y(0) + \sigma e^{-at}\mathcal{I}_t = e^{-at}(Y(0) + \sigma\mathcal{I}_t). \quad (2.4.5)$$

First we need to compute both $\mathbb{E}[\mathcal{I}_t]$ and $\mathbb{E}[\mathcal{I}_t^2]$:

$$\mathbb{E}[\mathcal{I}_t] = \mu t e^{at} - a\mu \left(t \frac{e^{at}}{a} - \int_0^t \frac{e^{au}}{a} du \right) = \frac{\mu}{a} (e^{at} - 1).$$

Now to find $\mathbb{E}[\mathcal{I}_t^2]$ we need;

$$\begin{aligned} \mathbb{E}[L(t)L(t+h)] &= \mathbb{E}[L(t)(L(t+h) - L(t) + L(t))] \\ &= \mathbb{E}[L(t)]\mathbb{E}[L(t+h) - L(t)] + \mathbb{E}[L^2(t)] \\ &= \mu^2 ht + \eta^2 t + \mu^2 t^2, \end{aligned}$$

and in general when $u \leq t$ we have:

$$\mathbb{E}[L(u)L(t)] = \mu^2(t-u)u + \eta^2 u + \mu^2 u^2 = (\mu^2 t + \eta^2)u. \quad (2.4.6)$$

To simplify calculations we will work with a centered Lévy process $\tilde{L}(t) = L(t) - \mu t$.

Then,

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^t e^{au} \tilde{L}(u) du \right)^2 \right] &= \mathbb{E} \left[\int_0^t \int_0^t e^{au} e^{av} \tilde{L}(u) \tilde{L}(v) dv du \right] \\ &= 2 \int_0^t e^{au} \left(\int_0^u e^{av} \mathbb{E}[\tilde{L}(v)\tilde{L}(u)] dv \right) du \\ &= 2\eta^2 \int_0^t e^{au} \left(\int_0^u v e^{av} dv \right) du = \frac{2\eta^2}{a} \int_0^t e^{au} \left(u e^{au} - \frac{e^{au} - 1}{a} \right) du \\ &= \frac{2\eta^2}{a} \int_0^t u e^{2au} du - \frac{2\eta^2}{a^2} \int_0^t e^{2au} du + \frac{2\eta^2}{a^2} \int_0^t e^{au} du \\ &= \frac{\eta^2}{a^2} \left(t e^{2at} - \frac{e^{2at} - 1}{2a} \right) - \frac{\eta^2}{a^3} (e^{2at} - 1) + \frac{2\eta^2}{a^3} (e^{at} - 1), \quad (2.4.7) \end{aligned}$$

where in the third equality we used equation (2.4.6) with $\mu = 0$.

Let

$$\tilde{\mathcal{I}}_t \equiv \int_0^t e^{au} d\tilde{L}(u) = e^{at}\tilde{L}(t) - a \int_0^t e^{au}\tilde{L}(u)du. \quad (2.4.8)$$

Using equation (2.4.7) we have

$$\begin{aligned} \mathbb{E} \left[\left(\tilde{\mathcal{I}}_t \right)^2 \right] &= \mathbb{E} \left[\left(e^{at}\tilde{L}(t) - a \int_0^t e^{au}\tilde{L}(u)du \right)^2 \right] \\ &= e^{2at}\mathbb{E} \left[\tilde{L}^2(t) \right] - 2ae^{at} \int_0^t e^{au}\mathbb{E} \left[\tilde{L}(t)\tilde{L}(u) \right] du + a^2\mathbb{E} \left[\left(\int_0^t e^{au}\tilde{L}(u)du \right)^2 \right] \\ &= \eta^2 te^{2at} - 2a\eta^2 e^{at} \int_0^t ue^{au} du + \eta^2 \left(te^{2at} - \frac{e^{2at} - 1}{2a} \right) - \frac{\eta^2}{a} (e^{2at} - 1) \\ &\quad + \frac{2\eta^2}{a} (e^{at} - 1) \\ &= \eta^2 te^{2at} - 2\eta^2 e^{at} \left(te^{at} - \frac{e^{at} - 1}{a} \right) + \eta^2 \left(te^{2at} - \frac{e^{2at} - 1}{2a} \right) - \frac{\eta^2}{a} (e^{2at} - 1) \\ &\quad + \frac{2\eta^2}{a} (e^{at} - 1) \\ &= \frac{2\eta^2 e^{at}}{a} (e^{at} - 1) - \frac{\eta^2}{2a} (e^{2at} - 1) - \frac{\eta^2}{a} (e^{2at} - 1) + \frac{2\eta^2}{a} (e^{at} - 1) \\ &= \frac{2\eta^2}{a} (e^{2at} - e^{at}) - \frac{3\eta^2}{2a} (e^{2at} - 1) + \frac{2\eta^2}{a} (e^{at} - 1) \\ &= \frac{\eta^2}{2a} e^{2at} + \frac{3\eta^2}{2a} - \frac{2\eta^2}{a} = \frac{\eta^2}{2a} (e^{2at} - 1). \end{aligned}$$

Using (2.4.4) and (2.4.8) we conclude the following relationship between integrals \mathcal{I}_t and $\tilde{\mathcal{I}}_t$:

$$\begin{aligned} \mathcal{I}_t &= \int_0^t e^{au} dL(u) = e^{at}L(t) - a \int_0^t e^{au}L(u)du \\ &= e^{at}(\tilde{L}(t) + \mu t) - a \int_0^t e^{au}(\tilde{L}(u) + \mu u)du \\ &= e^{at}\tilde{L}(t) - a \int_0^t e^{au}\tilde{L}(u)du + \mu te^{at} - a\mu \int_0^t ue^{au}du \\ &= \tilde{\mathcal{I}}_t + \mu te^{at} - \mu ue^{au} \Big|_0^t + \mu \int_0^t e^{au}du = \tilde{\mathcal{I}}_t + \frac{\mu}{a} (e^{at} - 1) \end{aligned}$$

$$= \tilde{\mathcal{I}}_t + \mathbb{E}[\mathcal{I}_t]. \quad (2.4.9)$$

Hence

$$\text{Var}(\mathcal{I}_t) = \text{Var}(\tilde{\mathcal{I}}_t) = \mathbb{E}[\tilde{\mathcal{I}}_t^2] = \frac{\eta^2}{2a} (e^{2at} - 1)$$

and consequently

$$\mathbb{E}[\mathcal{I}_t^2] = \text{Var}(\mathcal{I}_t) + \mathbb{E}^2[\mathcal{I}_t] = \frac{\eta^2}{2a} (e^{2at} - 1) + \frac{\mu^2}{a^2} (e^{at} - 1)^2.$$

Since Y is stationary, ($\mathbb{E}[Y(t)] = \mathbb{E}[Y(0)]$), and by taking the expectation for both sides of equation (2.4.5) we have:

$$(1 - e^{-at}) \mathbb{E}[Y(0)] = \sigma e^{-at} \mathbb{E}[\mathcal{I}_t] = \frac{\mu\sigma}{a} e^{-at} (e^{at} - 1),$$

which implies $\mathbb{E}[Y(0)] = \frac{\mu\sigma}{a}$. This concludes the computation of the expected value.

To find $\mathbb{E}[Y^2(0)]$, we start by squaring both sides of equation (2.4.5),

$$Y^2(t) = e^{-2at} Y^2(0) + 2\sigma e^{-2at} Y(0) \mathcal{I}_t + \sigma^2 e^{-2at} \mathcal{I}_t^2.$$

Taking the expectation for both sides, using stationarity and the fact that $Y(0)$ is independent of $\{L(t), t > 0\}$, we have:

$$\begin{aligned} & (1 - e^{-2at}) \mathbb{E}[Y^2(0)] \\ &= 2\sigma e^{-2at} \mathbb{E}[Y(0)] \mathbb{E}[\mathcal{I}_t] + \sigma^2 e^{-2at} \mathbb{E}[\mathcal{I}_t^2] \\ &= \frac{2\mu^2\sigma^2}{a^2} e^{-2at} (e^{at} - 1) + \sigma^2 e^{-2at} \mathbb{E}[\mathcal{I}_t^2] \\ &= \frac{2\mu^2\sigma^2}{a^2} e^{-at} - \frac{2\mu^2\sigma^2}{a^2} e^{-2at} + \sigma^2 e^{-2at} \left(\frac{\eta^2}{2a} (e^{2at} - 1) + \frac{\mu^2}{a^2} (e^{at} - 1)^2 \right) \\ &= \frac{2\mu^2\sigma^2}{a^2} e^{-at} - \frac{2\mu^2\sigma^2}{a^2} e^{-2at} + \frac{\sigma^2\eta^2}{2a} (1 - e^{-2at}) + \frac{\sigma^2\mu^2}{a^2} (1 - 2e^{-at} + e^{-2at}) \\ &= -\frac{2\mu^2\sigma^2}{a^2} e^{-2at} + \frac{\sigma^2\eta^2}{2a} (1 - e^{-2at}) + \frac{\sigma^2\mu^2}{a^2} + \frac{\sigma^2\mu^2}{a^2} e^{-2at} \\ &= \frac{\sigma^2\mu^2}{a^2} (1 - e^{-2at}) + \frac{\sigma^2\eta^2}{2a} (1 - e^{-2at}). \end{aligned}$$

This implies

$$\mathbb{E}[Y^2(0)] = \frac{\sigma^2\mu^2}{a^2} + \frac{\sigma^2\eta^2}{2a} \quad \text{and} \quad \text{Var}(Y(0)) = \frac{\sigma^2\eta^2}{2a}.$$

Finally, we compute the covariance:

$$\begin{aligned} \gamma_Y(s) &= \text{Cov}(Y(t), Y(t+s)) = \text{Cov}(Y(0), Y(s)) = \mathbb{E}[Y(0)Y(s)] - \mathbb{E}^2[Y(0)] \\ &= \mathbb{E}[Y(0)Y(s)] - \frac{\mu^2\sigma^2}{a^2}. \end{aligned}$$

Now

$$\begin{aligned} \mathbb{E}[Y(0)Y(s)] &= \mathbb{E}[Y(0)](e^{-as}(Y(0) + \sigma\mathcal{I}_s)) = e^{-as}(\mathbb{E}[Y^2(0)] + \sigma\mathbb{E}[Y(0)]\mathbb{E}[\mathcal{I}_s]) \\ &= e^{-as}\left(\mathbb{E}[Y^2(0)] + \frac{\mu^2\sigma^2}{a^2}(e^{as} - 1)\right) = e^{-as}\mathbb{E}[Y^2(0)] + \frac{\mu^2\sigma^2}{a^2}(1 - e^{-as}) \\ &= e^{-as}\mathbb{E}\left[Y^2(0) - \frac{\mu^2\sigma^2}{a^2}\right] + \frac{\mu^2\sigma^2}{a^2} = e^{-as}\text{Var}(Y(0)) + \frac{\mu^2\sigma^2}{a^2} \\ &= \frac{\sigma^2\eta^2}{2a}e^{-as} + \frac{\mu^2\sigma^2}{a^2}, \end{aligned}$$

which implies

$$\gamma_Y(s) = e^{-as}\text{Var}(Y(0)) = \frac{\sigma^2\eta^2}{2a}e^{-as}.$$

□

2.5 The Sampled CAR(1) Process

In practice, continuous time processes are usually sampled at discrete times. Here we assume that the CAR(1) process is observed at equally spaced intervals of length h .

To be precise, let Y be a strictly stationary CAR(1) process

$$Y(t) = e^{-at}Y(0) + \sigma \int_0^t e^{-a(t-u)}dL(u).$$

For $0 \leq s < t$ we have:

$$Y(t) = e^{-a(t-s)}Y(s) + \sigma \int_s^t e^{-a(t-u)}dL(u).$$

For $h > 0$ and $n \in \mathbb{Z}_+$ choose $t = nh$ and $s = (n-1)h$. Define $Y_n^{(h)} \equiv Y(nh)$. Then

$$Y_n^{(h)} = e^{-ah}Y_{n-1}^{(h)} + \sigma \int_{(n-1)h}^{nh} e^{-a(nh-u)} dL(u).$$

The sampled process $\{Y_n^{(h)}, n = 0, 1, 2, \dots\}$ can be written as:

$$Y_n^{(h)} = \phi Y_{n-1}^{(h)} + Z_n^{(h)}, \quad n = 0, 1, 2, \dots, \quad (2.5.1)$$

where

$$\phi = e^{-ah}, \quad \text{and} \quad Z_n^{(h)} = \sigma \int_{(n-1)h}^{nh} e^{-a(nh-u)} dL(u). \quad (2.5.2)$$

Now assume that L is a Lévy process. Since L has stationary and independent increments, $\{Z_n^{(h)}, n \geq 1\}$ is an i.i.d. sequence. Hence, the sampled process $\{Y_n^{(h)}, n = 0, 1, 2, \dots\}$ is a discrete-time AR(1) process.

If L is a second-order Lévy process then we can represent the noise $Z_n^{(h)}$ as

$$Z_n^{(h)} = Y_n^{(h)} - \phi Y_{n-1}^{(h)}.$$

Consequently, using Lemma 2.4.3 and stationarity of the process Y we have

$$\mathbb{E}[Z_n^{(h)}] = \mathbb{E}[Y_n^{(h)}] - \phi \mathbb{E}[Y_{n-1}^{(h)}] = (1 - \phi) \mathbb{E}[Y_0] = \frac{\mu\sigma}{a}(1 - \phi). \quad (2.5.3)$$

Similarly we can use $\mathbb{E}[Y_0^2]$ to compute $\mathbb{E}\left[\left(Z_n^{(h)}\right)^2\right]$ as follows:

$$\begin{aligned} \mathbb{E}\left[\left(Z_n^{(h)}\right)^2\right] &= \mathbb{E}\left[\left(Y_n^{(h)} - \phi Y_{n-1}^{(h)}\right)^2\right] = \mathbb{E}[Y_0^2] - 2\phi \mathbb{E}[Y_n^{(h)}Y_{n-1}^{(h)}] + \phi^2 \mathbb{E}[(Y_0)^2] \\ &= (1 + \phi^2) \mathbb{E}[Y_0^2] - 2\phi \mathbb{E}[Y_0Y_h] \\ &= (1 + \phi^2) \left(\frac{\mu^2\sigma^2}{a^2} + \frac{\sigma^2\eta^2}{2a}\right) - 2\phi \left(\frac{\sigma^2\eta^2}{2a}\phi + \frac{\mu^2\sigma^2}{a^2}\right) \\ &= (1 - \phi^2) \frac{\sigma^2\eta^2}{2a} + (1 - \phi)^2 \frac{\mu^2\sigma^2}{a^2}. \end{aligned}$$

Hence,

$$\text{Var}(Z_n^{(h)}) = (1 - \phi^2) \frac{\sigma^2\eta^2}{2a}. \quad (2.5.4)$$

2.6 Recovering the driving process

If the CAR(1) process Y is continuously observed on $[0, T]$ then the following theorem (cf. [11]) provides an inversion formula that represents L in terms of Y . The formula uses an argument of Pham-Din-Tuan (cf. [42]). The proof is provided for completeness.

Theorem 2.6.1 (Inversion Formula). *Let Y be a CAR(1) process satisfying*

$$Y(t) = e^{-at}Y(0) + \sigma \int_0^t e^{-a(t-u)} dL(u).$$

Then

$$L(t) = \sigma^{-1} \left[Y(t) - Y(0) + a \int_0^t Y(s) ds \right] \quad (2.6.1)$$

Proof: Let L be the process defined as in equation (2.6.1). It is enough to show that:

$$e^{-at}Y(0) + \sigma \int_0^t e^{-a(t-u)} dL(u) = Y(t).$$

Now,

$$\begin{aligned} & e^{-at}Y(0) + \sigma \int_0^t e^{-a(t-u)} dL(u) \\ &= e^{-at}Y(0) + \sigma L(t) - a\sigma \int_0^t e^{-a(t-u)} L(u) du \\ &= e^{-at}Y(0) + \left[Y(t) - Y(0) + a \int_0^t Y(s) ds \right] \\ &\quad - a \int_0^t e^{-a(t-u)} \left[Y(u) - Y(0) + a \int_0^u Y(s) ds \right] du \\ &= e^{-at}Y(0) + Y(t) - Y(0) + a \int_0^t Y(s) ds - a \int_0^t e^{-a(t-u)} Y(u) du \\ &\quad + a \int_0^t e^{-a(t-u)} Y(0) du - a^2 \int_0^t \left(\int_0^u e^{-a(t-u)} Y(s) ds \right) du \\ &= e^{-at}Y(0) + Y(t) - Y(0) + a \int_0^t Y(s) ds - a \int_0^t e^{-a(t-u)} Y(u) du + Y(0) - e^{-at}Y(0) \\ &\quad - a^2 \int_0^t \left(\int_s^t e^{-a(t-u)} Y(s) du \right) ds \quad (\text{by Fubini's Theorem}) \\ &= Y(t) + a \int_0^t Y(s) ds - a \int_0^t e^{-a(t-u)} Y(u) du - a \int_0^t Y(s) ds + a \int_0^t e^{-a(t-s)} Y(s) ds \end{aligned}$$

$$= Y(t).$$

□

Note that Theorem 2.6.1 does not require the assumption that L is Lévy.

2.7 Approximation of Lévy increments using the inversion formula

When the CAR(1) process Y is sampled discretely, the driving process L cannot be recovered exactly via the inversion formula (2.6.1). Instead, it is necessary to approximate the increment of L over the sampling intervals.

Recall the inversion formula (2.6.1):

$$L(t) = \sigma^{-1} \left[Y(t) - Y(0) + a \int_0^t Y(s) ds \right].$$

Thus, the increment of L over the interval $((n-1)h, nh)$, $h > 0$ is:

$$\Delta L_n^{(h)} = L(nh) - L((n-1)h) = \sigma^{-1} \left[Y_n^{(h)} - Y_{n-1}^{(h)} + a \int_{(n-1)h}^{nh} Y(u) du \right]. \quad (2.7.1)$$

The above increments require the continuously observed process Y . If the process is observed at discrete times nh , then Brockwell et al. (cf. [11]) replace the integral by a trapezoidal approximation. For now, let us assume that a and σ are known, in which case we have:

$$\Delta \widehat{L}_n^{(h)} := \sigma^{-1} \left[Y_n^{(h)} - Y_{n-1}^{(h)} + ah \frac{Y_n^{(h)} + Y_{n-1}^{(h)}}{2} \right], \quad n = 1, \dots, N. \quad (2.7.2)$$

We will refer to the above equation as *the estimated increments*. The estimated increments have the following properties.

Proposition 2.7.1. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L , $\mathbb{E}[L(t)] = \mu t$, $\text{Var}(L(t)) = \eta^2 t$. Then,*

$$(i) \quad \mathbb{E} \left[\Delta \widehat{L}_n^{(h)} \right] = \mathbb{E} \left[\Delta L_n^{(h)} \right] = \mu h;$$

(ii) *For each $s \geq 0$,*

$$\begin{aligned} \gamma_{\Delta \widehat{L}^{(h)}}(s) &\equiv \text{Cov} \left(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_{s+1}^{(h)} \right) \\ &= \left(\frac{\eta^2}{a} + \frac{a\eta^2 h^2}{4} \right) \phi^s + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) (\phi^{|s-1|} + \phi^{s+1}), \end{aligned}$$

where $\phi = e^{-ah}$.

(iii) *For each $s \geq 0$,*

$$\rho_{\Delta \widehat{L}^{(h)}}(s) \equiv \text{Corr} \left(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_{s+1}^{(h)} \right) = \frac{\left(\frac{\eta^2}{a} + \frac{a\eta^2 h^2}{4} \right) \phi^s + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) (\phi^{|s-1|} + \phi^{s+1})}{\left(\frac{\eta^2}{a} + \frac{a\eta^2 h^2}{4} \right) + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) (2\phi)}.$$

Before we provide a proof, we make several comments:

- We note that although L has independent increments, the non-zero covariance in (ii) appears due to the discretization error introduced by the trapezoidal approximation.
- $\lim_{n \rightarrow \infty} \text{Cov} \left(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_n^{(h)} \right) = 0$, for fixed h .
- For each $s \geq 0$ we have $\lim_{h \rightarrow 0} \text{Cov} \left(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_{s+1}^{(h)} \right) = \frac{\eta^2}{a} - \frac{\eta^2}{2a} - \frac{\eta^2}{2a} = 0$. Hence, as expected the discretization error disappears when $h \rightarrow 0$. However, the rate of convergence to zero is the same for covariances ($s > 0$) and variance ($s = 0$). For this reason, it does not seem possible to test for independence of the increments of L using estimates of $\text{Cov} \left(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_{s+1}^{(h)} \right)$ as $h \rightarrow 0$. An alternative approach is discussed in Section 3.3.4.

Proof: (i) We have from Lemma 2.4.3

$$\mathbb{E} \left[\Delta \widehat{L}_n^{(h)} \right] = \mathbb{E} \left[\Delta \widehat{L}_1^{(h)} \right] = \frac{ah}{\sigma} \mathbb{E} \left[Y_0^{(h)} \right] = \mu h.$$

(ii) For simplicity we consider $\gamma_{\Delta \widehat{L}^{(h)}}(n-1) = \text{Cov}(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_n^{(h)})$. Now

$$\begin{aligned} \text{Cov} \left(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_n^{(h)} \right) &= \mathbb{E} \left[\Delta \widehat{L}_1^{(h)} \Delta \widehat{L}_n^{(h)} \right] - \mathbb{E} \left[\Delta \widehat{L}_1^{(h)} \right] \mathbb{E} \left[\Delta \widehat{L}_n^{(h)} \right] \\ &= \mathbb{E} \left[\Delta \widehat{L}_1^{(h)} \Delta \widehat{L}_n^{(h)} \right] - \mu^2 h^2. \end{aligned}$$

Now,

$$\begin{aligned} &\mathbb{E} \left[\Delta \widehat{L}_1^{(h)} \Delta \widehat{L}_n^{(h)} \right] \\ &= \frac{1}{\sigma^2} \mathbb{E} \left[\left(Y_1^{(h)} - Y_0^{(h)} + ah \frac{Y_1^{(h)} + Y_0^{(h)}}{2} \right) \left(Y_n^{(h)} - Y_{n-1}^{(h)} + ah \frac{Y_n^{(h)} + Y_{n-1}^{(h)}}{2} \right) \right] \\ &= \frac{1}{\sigma^2} \left(\mathbb{E} \left[Y_1^{(h)} Y_n^{(h)} \right] + \mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] - \mathbb{E} \left[Y_1^{(h)} Y_{n-1}^{(h)} \right] - \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right] \right) \\ &\quad + \frac{ah}{2\sigma^2} \left(\mathbb{E} \left[Y_1^{(h)} Y_n^{(h)} \right] + \mathbb{E} \left[Y_1^{(h)} Y_{n-1}^{(h)} \right] - \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right] - \mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] \right) \\ &\quad + \mathbb{E} \left[Y_1^{(h)} Y_n^{(h)} \right] - \mathbb{E} \left[Y_1^{(h)} Y_{n-1}^{(h)} \right] \\ &\quad + \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right] - \mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] + \frac{a^2 h^2}{4\sigma^2} \mathbb{E} \left(\left(Y_1^{(h)} + Y_0^{(h)} \right) \left(Y_n^{(h)} + Y_{n-1}^{(h)} \right) \right) \\ &= \frac{1}{\sigma^2} \left(\mathbb{E} \left[Y_1^{(h)} Y_n^{(h)} \right] + \mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] - \mathbb{E} \left[Y_1^{(h)} Y_{n-1}^{(h)} \right] - \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right] \right) \\ &\quad + \frac{a^2 h^2}{4\sigma^2} \left(\mathbb{E} \left[Y_1^{(h)} Y_n^{(h)} \right] + \mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] + \mathbb{E} \left[Y_1^{(h)} Y_{n-1}^{(h)} \right] + \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right] \right) \\ &= \frac{1}{\sigma^2} \left(2\mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] - \mathbb{E} \left[Y_0^{(h)} Y_{(n-2)}^{(h)} \right] - \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right] \right) \\ &\quad + \frac{a^2 h^2}{4\sigma^2} \left(2\mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] + \mathbb{E} \left[Y_0^{(h)} Y_{n-2}^{(h)} \right] + \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right] \right) \\ &= \left(\frac{2}{\sigma^2} + \frac{a^2 h^2}{2\sigma^2} \right) \mathbb{E} \left[Y_0^{(h)} Y_{n-1}^{(h)} \right] + \left(\frac{a^2 h^2}{4\sigma^2} - \frac{1}{\sigma^2} \right) \mathbb{E} \left[Y_0^{(h)} Y_{(n-2)}^{(h)} \right] \\ &\quad + \left(\frac{a^2 h^2}{4\sigma^2} - \frac{1}{\sigma^2} \right) \mathbb{E} \left[Y_0^{(h)} Y_n^{(h)} \right]. \end{aligned}$$

Using Lemma 2.4.3 we obtain

$$\mathbb{E} \left[\Delta \widehat{L}_1^{(h)} \Delta \widehat{L}_n^{(h)} \right]$$

$$\begin{aligned}
&= \left(\frac{2}{\sigma^2} + \frac{a^2 h^2}{2\sigma^2} \right) \left(\frac{\sigma^2 \eta^2}{2a} e^{-a(n-1)h} + \frac{\mu^2 \sigma^2}{a^2} \right) + \left(\frac{a^2 h^2}{4\sigma^2} - \frac{1}{\sigma^2} \right) \left(\frac{\sigma^2 \eta^2}{2a} e^{-a|n-2|h} + \frac{\mu^2 \sigma^2}{a^2} \right) \\
&\quad + \left(\frac{a^2 h^2}{4\sigma^2} - \frac{1}{\sigma^2} \right) \left(\frac{\sigma^2 \eta^2}{2a} e^{-anh} + \frac{\mu^2 \sigma^2}{a^2} \right) \\
&= \left(\frac{\eta^2}{a} + \frac{a\eta^2 h^2}{4} \right) e^{-a(n-1)h} + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) e^{-a|n-2|h} + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) e^{-anh} \\
&\quad + \frac{\mu^2 \sigma^2}{a^2} \left(\frac{2}{\sigma^2} + \frac{a^2 h^2}{2\sigma^2} \right) + \frac{\mu^2 \sigma^2}{a^2} \left(\frac{a^2 h^2}{2\sigma^2} - \frac{2}{\sigma^2} \right) \\
&= \left(\frac{\eta^2}{a} + \frac{a\eta^2 h^2}{4} \right) e^{-a(n-1)h} + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) e^{-a|n-2|h} + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) e^{-anh} + \mu^2 h^2.
\end{aligned}$$

This implies

$$\begin{aligned}
\text{Cov} \left(\Delta \widehat{L}_1^{(h)}, \Delta \widehat{L}_n^{(h)} \right) &= \left(\frac{\eta^2}{a} + \frac{a\eta^2 h^2}{4} \right) e^{-a(n-1)h} + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) e^{-a|n-2|h} \\
&\quad + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) e^{-anh} \\
&= \left(\frac{\eta^2}{a} + \frac{a\eta^2 h^2}{4} \right) e^{-a(n-1)h} + \left(\frac{a\eta^2 h^2}{8} - \frac{\eta^2}{2a} \right) \left(e^{-a|n-2|h} + e^{-anh} \right).
\end{aligned}$$

This finishes the proof. \square

Chapter 3

Asymptotics for the sampled process

In this chapter, we consider the asymptotic properties of the sample characteristics (mean, variance, covariance) of the recovered driving process under different discretization scenarios. Assume that Y is observed at times $0, h, 2h, \dots, [T/h]h$ on the interval $[0, T]$.

- **Case (I)**: h is fixed, $T = Nh$, and $N \rightarrow \infty$.
- **Case (II)**: $T = N$ remains constant, $h = \frac{1}{M}$, and $M \rightarrow \infty$.
- **Case (III)**: $T = N$, $h = \frac{1}{M}$, and $N \wedge M \rightarrow \infty$.

We will see that Case I can be handled with classical time series techniques (see e.g. [10], [19]), whereas Cases II and III require a new approach, which is our main contribution (see also [1]).

In Section 3.1 we introduce the three scenarios. In Section 3.2 we obtain asymptotic normality of the sample mean and sample covariances in Case I. Proofs use standard time series techniques.

Section 3.3 deals with Cases (II) and (III). The first main result is Theorem 3.3.1 that provides the approximation of the true process L . The proof relies on Lemma 3.3.2 in which we approximate the continuous process Y via its sampled version. Then, we apply the inversion formula. As a corollary, we obtain the central limit theorem for the estimated process L (Corollary 3.3.3). Next, we provide a uniform bound on the approximation error in the unit increments of the recovered driving process (see Lemmas 3.3.5 and 3.3.7). Our Lemma 3.3.5 can be compared to Theorem 5.7 in [14]. Although the result in the latter paper holds for more general multivariate CARMA models, our bound is more precise and is uniform with respect to N , the length of time that the process is observed. As a consequence, we can derive central limit theorems for sample covariances (see Theorem 3.3.11).

In Section 4, in a brief digression, as an important by-product, we prove a central limit theorem for the integrated CAR(1) process Y (Theorem 3.4.2). The significance of this result is that the proof is quite elementary and does not require any mixing arguments.

3.1 Discrete approximation of L

Recall equation (2.2.2):

$$\mathbb{E}[L(t)] = \mu t \quad \text{and} \quad \text{Var}(L(t)) = \eta^2 t, \quad t \geq 0, \quad (3.1.1)$$

and the estimated increments (cf. (2.7.2)):

$$\begin{aligned} \Delta \widehat{L}_n^{(h)} &= \sigma^{-1} \left[Y_n^{(h)} - Y_{n-1}^{(h)} + ah \frac{Y_n^{(h)} + Y_{n-1}^{(h)}}{2} \right] \\ &= \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) Y_n^{(h)} + \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) Y_{n-1}^{(h)}, \quad n = 1, \dots, N. \end{aligned} \quad (3.1.2)$$

Since $L(Nh) = \sum_{n=1}^N \Delta L_n^{(h)}$, the estimated value $\widehat{L}^{(h)}(Nh)$ of $L(Nh)$ from the dis-

cretely observed process $Y_n^{(h)}$ can be written as a partial sum:

$$\begin{aligned}
\widehat{L}^{(h)}(Nh) &\equiv \sum_{n=1}^N \Delta \widehat{L}_n^{(h)} = \frac{1}{\sigma} \sum_{n=1}^N \left[Y_n^{(h)} - Y_{n-1}^{(h)} + ah \frac{Y_n^{(h)} + Y_{n-1}^{(h)}}{2} \right] \\
&= \frac{1}{\sigma} \left(Y_N^{(h)} - Y_0^{(h)} \right) + \frac{ah}{2\sigma} \sum_{n=1}^N Y_n^{(h)} + \frac{ah}{2\sigma} \sum_{n=1}^N Y_{n-1}^{(h)} \\
&= \frac{1}{\sigma} \left(Y_N^{(h)} - Y_0^{(h)} \right) + \frac{ah}{2\sigma} \sum_{n=1}^N Y_n^{(h)} + \frac{ah}{2\sigma} \left(Y_N^{(h)} - Y_N^{(h)} + Y_0^{(h)} + \sum_{n=2}^N Y_{n-1}^{(h)} \right) \\
&= \frac{ah}{\sigma} \sum_{n=1}^N Y_n^{(h)} + \left(\frac{1}{\sigma} - \frac{ah}{2\sigma} \right) \left(Y_N^{(h)} - Y_0^{(h)} \right). \tag{3.1.3}
\end{aligned}$$

Sample mean and covariance

We define $\overline{\Delta \widehat{L}^{(h)}}$, the sample mean of the estimated increments $\Delta \widehat{L}_n^{(h)}$, $n = 1, \dots, N$, as

$$\overline{\Delta \widehat{L}^{(h)}} \equiv \frac{1}{N} \sum_{n=1}^N \Delta \widehat{L}_n^{(h)} = \frac{\widehat{L}^{(h)}(Nh)}{N},$$

and $\widehat{\gamma_{\Delta \widehat{L}^{(h)}}}(k)$, the sample covariances of $\Delta \widehat{L}_n^{(h)}$, $n = 1, 2, \dots, N$, at lag $k \geq 0$, as

$$\widehat{\gamma_{\Delta \widehat{L}^{(h)}}}(k) = \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta \widehat{L}_{n+k}^{(h)} - \overline{\Delta \widehat{L}^{(h)}} \right) \left(\Delta \widehat{L}_n^{(h)} - \overline{\Delta \widehat{L}^{(h)}} \right), \quad 0 \leq k \leq N. \tag{3.1.4}$$

Furthermore, let $\widehat{\rho_{\Delta \widehat{L}^{(h)}}}(k)$ be the sample correlation which is defined as

$$\widehat{\rho_{\Delta \widehat{L}^{(h)}}}(k) = \frac{\widehat{\gamma_{\Delta \widehat{L}^{(h)}}}(k)}{\widehat{\gamma_{\Delta \widehat{L}^{(h)}}}(0)}.$$

Three discretization scenarios

We will consider three different discrete sampling scenarios when investigating the asymptotic behaviour of the sample mean $\overline{\Delta \widehat{L}^{(h)}}$ and sample covariances $\widehat{\gamma_{\Delta \widehat{L}^{(h)}}}(k)$. Assume that Y is observed at times $0, h, 2h, \dots, [T/h]h$ on the interval $[0, T]$.

- **Case (I):** h is fixed, $T = Nh$, and $N \rightarrow \infty$.
- **Case (II):** $T = N$ remains constant, $h = \frac{1}{M}$, and $M \rightarrow \infty$.
- **Case (III):** $T = N$, $h = \frac{1}{M}$, and $N \wedge M \rightarrow \infty$.

3.2 Case (I): h is fixed, $T = Nh$, and $N \rightarrow \infty$

As we noticed in Proposition 2.7.1, the discretization with a fixed frequency h introduces an error leading to estimated recovered increments with non-zero covariance. Nevertheless, one can still obtain some relevant limiting results that can be used for estimation of the model parameters.

Proposition 3.2.1. *Consider the second order stationary Lévy-driven CAR(1) model Y such that (3.1.1) holds. Then*

- (i) $\overline{\Delta \widehat{L}^{(h)}} \xrightarrow{p} \frac{ah}{\sigma} \mathbb{E}[Y(0)] = \mu h$ as $N \rightarrow \infty$;
- (ii) $\sqrt{N} \left(\overline{\Delta \widehat{L}^{(h)}} - \mu h \right) = \frac{1}{\sqrt{N}} \left(\widehat{L}^{(h)}(Nh) - \mu Nh \right) \xrightarrow{d} N \left(0, \frac{ah^2 \eta^2}{2} \frac{1+\phi}{1-\phi} \right)$, $\phi = e^{-ah}$, as $N \rightarrow \infty$.

Remark 3.2.2. Before we provide a proof, we indicate how this result can be used in statistical inference. When $h > 0$ is known, the central limit theorem for the sample mean allows us to construct confidence intervals for μ based on the recovered Lévy process. Nuisance parameters η, a appear only in the variance of the limiting normal distribution and in principle can be bootstrapped without estimating the parameters a, η . On the other hand, when a confidence interval for μ is based on the process Y then Lemma 2.4.3 indicates a non-identifiability issue.

Proof: We start with part (i). From equation (3.1.3), the sum of Lévy increments $\sum_{n=1}^N \Delta L_n^{(h)}$ defining $L(Nh)$ can be replaced with the sum of AR(1) random variables plus a term that becomes negligible under norming by N . We obtain

$$\overline{\Delta \widehat{L}^{(h)}} = \frac{ah}{\sigma} \frac{1}{N} \sum_{n=1}^N Y_n^{(h)} + \left(\frac{1}{\sigma} - \frac{ah}{2\sigma} \right) \frac{Y_N^{(h)} - Y_0^{(h)}}{N}. \quad (3.2.1)$$

By ergodicity of $Y_n^{(h)}$, $n \geq 1$, and Lemma 2.4.3 we have

$$\frac{1}{N} \sum_{n=1}^N Y_n^{(h)} \xrightarrow{p} \mathbb{E} \left[Y_0^{(h)} \right] = \frac{\mu\sigma}{a}.$$

The result follows since the second term in the equation (3.2.1) converges to zero in probability as $N \rightarrow \infty$ since $\mathbb{E} \left[\left(Y_N^{(h)} - Y_0^{(h)} \right)^2 \right] < \infty$.

We proceed with the proof of (ii). Since $Y_n^{(h)}$ is a strictly stationary AR(1) process (c.f. (2.5.1)), it can be written as (cf. [10])

$$\begin{aligned} Y_n^{(h)} &= \sum_{j=0}^{\infty} \phi^j Z_{n-j}^{(h)} = \sum_{j=0}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} \left[Z_{n-j}^{(h)} \right] + \mathbb{E} \left[Z_{n-j}^{(h)} \right] \right) \\ &= \sum_{j=0}^{\infty} \phi^j \mathbb{E} \left[Z_{n-j}^{(h)} \right] + \sum_{j=0}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} \left[Z_{n-j}^{(h)} \right] \right), \end{aligned}$$

where $Z_n^{(h)}$ is defined as in equation (2.5.2) if $n \geq 0$ and if $n < 0$ we define,

$$Z_n^{(h)} = \sigma \int_{(n-1)h}^{nh} e^{-a(nh-u)} dM(-u-),$$

where $M \equiv \{M(t), t \geq 0\}$ is a second Lévy process independent of L and with the same distribution.

Now by equation (2.5.3)

$$\begin{aligned} Y_n^{(h)} &= \sum_{j=0}^{\infty} \phi^j \frac{\mu\sigma}{a} (1 - \phi) + \sum_{j=0}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} \left[Z_{n-j}^{(h)} \right] \right) \\ &= \frac{\mu\sigma}{a} + \sum_{j=0}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} \left[Z_{n-j}^{(h)} \right] \right) \\ &= \frac{\mu\sigma}{a} + \sum_{j=-\infty}^{\infty} \psi_j \left(Z_{n-j}^{(h)} - \mathbb{E} \left[Z_{n-j}^{(h)} \right] \right), \end{aligned} \tag{3.2.2}$$

where

$$\psi_j = \begin{cases} 0 & \text{if } j < 0 \\ \phi^j & \text{if } j \geq 0 \end{cases}.$$

Note that (cf. (2.5.4)):

$$\left(Z_{n-j}^{(h)} - \mathbb{E} \left[Z_{n-j}^{(h)} \right] \right) \sim \text{IID} \left(0, \text{Var} \left(Z_{n-j}^{(h)} \right) \right) = \text{IID} \left(0, (1 - \phi^2) \frac{\sigma^2 \eta^2}{2a} \right).$$

Using Theorem 7.1.2 in [10] we have:

$$\frac{1}{\sqrt{N}} \sum_{n=1}^N (Y_n^{(h)} - \mathbb{E}[Y_0]) \xrightarrow{d} N(0, \beta^2), \quad (3.2.3)$$

where

$$\begin{aligned} \beta^2 &= \sum_{j=-\infty}^{\infty} \gamma_Y(jh) = (1 - \phi^2) \frac{\sigma^2 \eta^2}{2a} \left(\sum_{j=-\infty}^{\infty} \psi_j \right)^2 \\ &= (1 - \phi^2) \frac{\sigma^2 \eta^2}{2a} \left(\frac{1}{1 - \phi} \right)^2 = \left(\frac{1 + \phi}{1 - \phi} \right) \frac{\sigma^2 \eta^2}{2a}. \end{aligned}$$

Hence,

$$\frac{ah}{\sigma} \frac{1}{\sqrt{N}} \sum_{n=1}^N \left(Y_n^{(h)} - \frac{\mu\sigma}{a} \right) \xrightarrow{d} N\left(0, \frac{a^2 h^2}{\sigma^2} \beta^2\right).$$

The second term (multiplied by \sqrt{N}) in (3.2.1) converges to zero in probability as $N \rightarrow \infty$, and so by Slutsky's Theorem we have:

$$\frac{1}{\sqrt{N}} \sum_{n=1}^N \left(\Delta \widehat{L}_n^{(h)} - \mu h \right) \xrightarrow{d} N\left(0, \frac{a^2 h^2}{\sigma^2} \beta^2\right) = N\left(0, \frac{ah^2 \eta^2}{2} \frac{1 + \phi}{1 - \phi}\right).$$

□

We state a result for the sample covariances as well. However, from an applied point of view this result is not particularly useful, since the theoretical covariances of the estimated increments do not vanish due to the discretization error.

Proposition 3.2.3. *Consider the second order stationary CAR(1) model Y such that (3.1.1) holds. Then for $k \geq 1$ we have*

- (i) $\widehat{\gamma_{\Delta \widehat{L}^{(h)}}}(k) \xrightarrow{p} \gamma_{\Delta \widehat{L}^{(h)}}(k)$ as $N \rightarrow \infty$;
- (ii) $\sqrt{N} \left(\widehat{\rho_{\Delta \widehat{L}^{(h)}}}(k) - \rho_{\Delta \widehat{L}^{(h)}}(k) \right) \xrightarrow{d} N(0, W_k^2)$ as $N \rightarrow \infty$, where

$$W_k^2 = \sum_{j=1}^{\infty} \left(\rho_{\Delta \widehat{L}^{(h)}}(j+k) + \rho_{\Delta \widehat{L}^{(h)}}(j-k) - 2\rho_{\Delta \widehat{L}^{(h)}}(k)\rho_{\Delta \widehat{L}^{(h)}}(j) \right)^2.$$

Proof: (i): We decompose the sample covariances in the standard way:

$$\begin{aligned}
& \widehat{\gamma_{\Delta\widehat{L}_n^{(h)}}}(k) \\
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta\widehat{L}_{n+k}^{(h)} - \overline{\Delta\widehat{L}^{(h)}} \right) \left(\Delta\widehat{L}_n^{(h)} - \overline{\Delta\widehat{L}^{(h)}} \right) \\
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta\widehat{L}_{n+k}^{(h)} \Delta\widehat{L}_n^{(h)} - \Delta\widehat{L}_{n+k}^{(h)} \overline{\Delta\widehat{L}^{(h)}} - \Delta\widehat{L}_n^{(h)} \overline{\Delta\widehat{L}^{(h)}} + \overline{\Delta\widehat{L}^{(h)}}^2 \right) \\
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta\widehat{L}_{n+k}^{(h)} \Delta\widehat{L}_n^{(h)} - \overline{\Delta\widehat{L}^{(h)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta\widehat{L}_{n+k}^{(h)} - \overline{\Delta\widehat{L}^{(h)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta\widehat{L}_n^{(h)} \\
&\quad + \left(\overline{\Delta\widehat{L}^{(h)}} \right)^2.
\end{aligned} \tag{3.2.4}$$

We use the formula (3.1.2) to simplify the first term of equation (3.2.4):

$$\begin{aligned}
& \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta\widehat{L}_{n+k}^{(h)} \Delta\widehat{L}_n^{(h)} \\
&= \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right)^2 \frac{1}{N-k} \sum_{n=1}^{N-k} Y_{n+k}^{(h)} Y_n^{(h)} + \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right)^2 \frac{1}{N-k} \sum_{n=1}^{N-k} Y_{n-1}^{(h)} Y_{n+k-1}^{(h)} \\
&\quad + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \frac{1}{N-k} \sum_{n=1}^{N-k} Y_{n+k-1}^{(h)} Y_n^{(h)} \\
&\quad + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \frac{1}{N-k} \sum_{n=1}^{N-k} Y_{n+k}^{(h)} Y_{n-1}^{(h)}.
\end{aligned} \tag{3.2.5}$$

Using ergodicity of $Y_n^{(h)}$, $n \geq 1$, we have

$$\frac{1}{N-k} \sum_{n=1}^{N-k} Y_{n+k}^{(h)} Y_n^{(h)} \xrightarrow{p} \mathbb{E} \left[Y_0^{(h)} Y_k^{(h)} \right].$$

A similar argument can be applied to the other three terms in (3.2.5). Hence, as $N \rightarrow \infty$,

$$\frac{1}{N-k} \sum_{n=1}^{N-k} \Delta\widehat{L}_{n+k}^{(h)} \Delta\widehat{L}_n^{(h)}$$

$$\begin{aligned}
& \xrightarrow{p} \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right)^2 \mathbb{E} \left[Y_0^{(h)} Y_k^{(h)} \right] + \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right)^2 \mathbb{E} \left[Y_0^{(h)} Y_k^{(h)} \right] \\
& \quad + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \mathbb{E} \left[Y_0^{(h)} Y_{k-1}^{(h)} \right] + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \mathbb{E} \left[Y_0^{(h)} Y_{k+1}^{(h)} \right] \\
& = \left(\left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right)^2 + \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right)^2 \right) \mathbb{E} \left[Y_0^{(h)} Y_k^{(h)} \right] \\
& \quad + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \mathbb{E} \left[Y_0^{(h)} Y_{k-1}^{(h)} \right] + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \mathbb{E} \left[Y_0^{(h)} Y_{k+1}^{(h)} \right] \\
& \stackrel{(*)}{=} \left(\frac{2}{\sigma^2} + \frac{a^2 h^2}{2\sigma^2} \right) \left(\frac{\sigma^2 \eta^2}{2a} e^{-ahk} + \frac{\mu^2 \sigma^2}{a^2} \right) + \left(\frac{a^2 h^2}{4\sigma^2} - \frac{1}{\sigma^2} \right) \left(\frac{\sigma^2 \eta^2}{2a} e^{-ah|k-1|} + \frac{\mu^2 \sigma^2}{a^2} \right) \\
& \quad + \left(\frac{a^2 h^2}{4\sigma^2} - \frac{1}{\sigma^2} \right) \left(\frac{\sigma^2 \eta^2}{2a} e^{-ah(k+1)} + \frac{\mu^2 \sigma^2}{a^2} \right) \\
& = \left(\frac{2}{\sigma^2} + \frac{a^2 h^2}{2\sigma^2} \right) \frac{\sigma^2 \eta^2}{2a} e^{-ahk} + \left(\frac{a^2 h^2}{4\sigma^2} - \frac{1}{\sigma^2} \right) \frac{\sigma^2 \eta^2}{2a} (e^{-ah|k-1|} + e^{-ah(k+1)}) + \mu^2 h^2 \\
& = \left(\frac{\eta^2}{a} + \frac{ah^2 \eta^2}{4} \right) e^{-ahk} + \left(\frac{ah^2 \eta^2}{8} - \frac{\eta^2}{2a} \right) (e^{-ah|k-1|} + e^{-ah(k+1)}) + \mu^2 h^2 \\
& \stackrel{(**)}{=} \gamma_{\Delta \widehat{L}^{(h)}}(k) + \mu^2 h^2;
\end{aligned}$$

(in (*) we used Lemma 2.4.3 and in (**)) Proposition 2.7.1).

Using the continuous mapping theorem in conjunction with Proposition 3.2.1 and Slutsky's theorem we have:

- $\overline{\Delta \widehat{L}^{(h)}}^2 \xrightarrow{p} \mu^2 h^2,$
- $\overline{\Delta \widehat{L}^{(h)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta \widehat{L}_{n+k}^{(h)} \xrightarrow{p} \mu^2 h^2,$
- $\overline{\Delta \widehat{L}^{(h)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta \widehat{L}_n^{(h)} \xrightarrow{p} \mu^2 h^2.$

Putting everything together we have

$$\widehat{\gamma_{\Delta \widehat{L}^{(h)}}}(k) \xrightarrow{p} \gamma_{\Delta \widehat{L}^{(h)}}(k) + \mu^2 h^2 - \mu^2 h^2 - \mu^2 h^2 + \mu^2 h^2.$$

This finishes the proof of (i).

(ii): Using equation (3.2.2) we can rewrite equation (3.1.2) in the following form:

$$\begin{aligned}
\Delta \widehat{L}_n^{(h)} &= \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left(\frac{\mu\sigma}{a} + \sum_{j=0}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} \left[Z_{n-j}^{(h)} \right] \right) \right) \\
&\quad + \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \left(\frac{\mu\sigma}{a} + \sum_{j=0}^{\infty} \phi^j \left(Z_{n-1-j}^{(h)} - \mathbb{E} \left[Z_{n-1-j}^{(h)} \right] \right) \right) \\
&= \left(\frac{\mu}{a} + \frac{\mu h}{2} \right) + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) \left((Z_n^{(h)} - \mathbb{E} [Z_n^{(h)}]) + \sum_{j=1}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} [Z_{n-j}^{(h)}] \right) \right) \\
&\quad + \left(\frac{\mu h}{2} - \frac{\mu}{a} \right) + \left(\frac{ah}{2\sigma} - \frac{1}{\sigma} \right) \frac{1}{\phi} \sum_{j=1}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} [Z_{n-j}^{(h)}] \right) \\
&= \mu h + \sum_{j=1}^{\infty} \phi^j \left(Z_{n-j}^{(h)} - \mathbb{E} [Z_{n-j}^{(h)}] \right) \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} + \frac{ah}{2\phi} - \frac{1}{\sigma\phi} \right) \\
&\quad + \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) (Z_n^{(h)} - \mathbb{E} [Z_n^{(h)}]) \\
&= \mu h + \sum_{j=-\infty}^{\infty} \tilde{\psi}_j \left(Z_{n-j}^{(h)} - \mathbb{E} [Z_{n-j}^{(h)}] \right), \tag{3.2.7}
\end{aligned}$$

where

$$\tilde{\psi}_j = \begin{cases} 0 & \text{if } j < 0 \\ \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} \right) & \text{if } j = 0 \\ \phi^j \left(\frac{1}{\sigma} + \frac{ah}{2\sigma} + \frac{ah}{2\phi} - \frac{1}{\sigma\phi} \right) & \text{if } j \geq 1. \end{cases}$$

Since we have:

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty \quad \text{and} \quad \sum_{j=-\infty}^{\infty} \psi_j^2 |j| < \infty,$$

then the result (ii) follows by using Theorem 7.2.2 in [10]. \square

3.3 Case (II) and Case (III): $h = 1/M$, and $M \rightarrow \infty$

We start with some notation that will be used throughout this section. Here, we assume that $h = 1/M$. As in equation (3.1.3) we consider

$$\widehat{L}_N^{(M)} \equiv \widehat{L}^{(1/M)} \left(NM \frac{1}{M} \right) = \sum_{n=1}^{NM} \Delta \widehat{L}_n^{(1/M)}. \quad (3.3.1)$$

Note that $\widehat{L}_N^{(M)}$ has a different meaning than $\widehat{L}_N^{(h)}$ in (3.1.3). Indeed, in (3.1.3) we computed the sum of N observations taken at time points $h, 2h, \dots, Nh$. In other words, $\widehat{L}_N^{(h)}$ approximates the Lévy process at time point Nh . Here, we have $N \times M$ observations sampled at $1/M, 2/M, \dots, N$. In other words, $\widehat{L}_N^{(M)}$ approximates $L(N)$. As in (3.1.3), we will represent $\widehat{L}_N^{(M)}$ in terms of the sampled process Y . Clearly, $Y_n^{(1/M)} = Y \left(\frac{n}{M} \right)$. Hence,

$$\begin{aligned} \widehat{L}_N^{(M)} &= \frac{a}{M\sigma} \sum_{n=1}^{NM} Y_n^{(1/M)} + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) \left(Y_{NM}^{(1/M)} - Y_0^{(1/M)} \right) \\ &= \frac{a}{M\sigma} \sum_{n=1}^{NM} Y_{\frac{n}{M}} + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) (Y_N - Y_0). \end{aligned} \quad (3.3.2)$$

3.3.1 Approximation of Y and L

In Case I we analyzed the behaviour of the partial sum $\widehat{L}^{(h)}(Nh)$ by representing it in terms of an AR(1) model. Here, we take a different route. We approximate the estimated Lévy process by the true process L using the bound given in Theorem 3.3.1 below. This result is applicable in Case (II) and Case (III) since the sampling error will converge to zero. This approach could not be used in the previous case, due to the sampling error coming from a fixed h .

Theorem 3.3.1. *Consider Case (II) and Case (III). Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (3.1.1) holds. If $\widehat{L}_N^{(M)}$ is defined as in equation (3.3.2) then for every $N, M \in \mathbb{Z}_+$,*

$$\left\| \frac{1}{\sqrt{N}} \left(L(N) - \widehat{L}_N^{(M)} \right) \right\|_{L_2} \leq \eta \sqrt{N} \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{\sqrt{a}\eta}{2M\sqrt{N}}.$$

Consequently, the bound converges to 0 as $N \rightarrow \infty$ and $N/M \rightarrow 0$.

The proof of Theorem 3.3.1 relies on the following uniform bound on the difference between integrals of Y and the corresponding discretely observed process.

Lemma 3.3.2. *Under the assumptions of Theorem 3.3.1, for every $N, M \in \mathbb{Z}_+$,*

$$\left\| \frac{1}{\sqrt{N}} \left(\int_0^N Y(s) ds - \frac{1}{M} \sum_{n=1}^{NM} Y_{\frac{n}{M}} \right) \right\|_{L_2} \leq \sqrt{N} \frac{\sigma \eta}{\sqrt{a}} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}}. \quad (3.3.3)$$

The bound converges to 0 as $N \rightarrow \infty$ and $N/M \rightarrow 0$.

Proof: In what follows we will also use the following notation. Let

$$Y_N^{(M)}(s) \equiv \sum_{i=1}^{NM} Y_{\frac{i}{M}} 1_{\left(\frac{i-1}{M} < s \leq \frac{i}{M}\right)}. \quad (3.3.4)$$

Then

$$\begin{aligned} \left\| \frac{1}{\sqrt{N}} \left(\int_0^N Y(s) ds - \frac{1}{M} \sum_{n=1}^{NM} Y_{\frac{n}{M}} \right) \right\|_{L_2}^2 &= \frac{1}{N} \mathbb{E} \left[\left(\int_0^N Y(s) ds - \frac{1}{M} \sum_{n=1}^{NM} Y_{\frac{n}{M}} \right)^2 \right] \\ &= \frac{1}{N} \mathbb{E} \left[\left(\int_0^N Y(s) ds - \int_0^N Y_N^{(M)}(s) ds \right)^2 \right] \\ &= \frac{1}{N} \mathbb{E} \left[\int_0^N \left(Y(s) - Y_N^{(M)}(s) \right)^2 ds \right] \\ &\leq \mathbb{E} \int_0^N \left(Y(s) - Y_N^{(M)}(s) \right)^2 ds, \end{aligned}$$

by the Cauchy-Buniakowski-Schwarz inequality. We have

$$\begin{aligned} \int_0^N \mathbb{E} \left[\left(Y(s) - Y_N^{(M)}(s) \right)^2 \right] ds &= \sum_{i=1}^{NM} \int_{\frac{i-1}{M}}^{\frac{i}{M}} \mathbb{E} \left[\left(Y(s) - Y_{\frac{i}{M}} \right)^2 \right] ds \\ &= \sum_{i=1}^{NM} \int_{\frac{i-1}{M}}^{\frac{i}{M}} \text{Var} \left(Y(s) - Y_{\frac{i}{M}} \right) ds \\ &= \sum_{i=1}^{NM} \int_{\frac{i-1}{M}}^{\frac{i}{M}} \left(2\text{Var} Y_0 - 2\text{Cov}(Y_0, Y_{\frac{i}{M}-s}) \right) ds. \end{aligned}$$

Hence, using (2.4.3)

$$\int_0^N \mathbb{E} \left[\left(Y(s) - Y_N^{(M)}(s) \right)^2 \right] ds = \sum_{i=1}^{NM} \int_{\frac{i-1}{M}}^{\frac{i}{M}} \left(2 \frac{\sigma^2 \eta^2}{2a} - 2 \frac{\sigma^2 \eta^2}{2a} e^{-a(\frac{i}{M}-s)} \right) ds$$

$$\begin{aligned}
&= \frac{\sigma^2 \eta^2}{a} \sum_{i=1}^M \int_{\frac{i-1}{M}}^{\frac{i}{M}} \left(1 - e^{-a(\frac{i}{M}-s)}\right) ds \\
&\leq \frac{\sigma^2 \eta^2}{a} \sum_{i=1}^{NM} \int_{\frac{i-1}{M}}^{\frac{i}{M}} \left(1 - e^{-\frac{a}{M}}\right) ds \\
&= N \frac{\sigma^2 \eta^2}{a} \left(1 - e^{-\frac{a}{M}}\right).
\end{aligned}$$

This finishes the proof. \square

Proof of Theorem 3.3.1:

Recall equation (3.3.2) and the representation of $L(N)$ (cf. (2.6.1)):

$$L(N) = \frac{1}{\sigma} \left[Y(N) - Y(0) - \int_0^N Y(s) ds \right].$$

Using the bound from Lemma 3.3.2 we have

$$\begin{aligned}
&\left\| \frac{1}{\sqrt{N}} \left(L(N) - \widehat{L}_N^{(M)} \right) \right\|_{L_2} \\
&= \frac{1}{\sqrt{N}} \left\| \frac{1}{\sigma} \left[Y(N) - Y(0) + a \int_0^N Y(s) ds \right] - \frac{a}{M\sigma} \sum_{n=1}^{NM} Y_{\frac{n}{M}} + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) (Y_N - Y_0) \right\|_{L_2} \\
&= \frac{1}{\sqrt{N}} \left\| \frac{a}{\sigma} \int_0^N Y(s) ds - \frac{a}{M\sigma} \sum_{n=1}^{NM} Y_{\frac{n}{M}} - \frac{a}{2M\sigma} (Y_N - Y_0) \right\|_{L_2} \\
&\leq \frac{1}{\sqrt{N}} \frac{a}{\sigma} \left\| \int_0^N Y(s) ds - \frac{1}{M} \sum_{n=1}^{NM} Y_{\frac{n}{M}} \right\|_{L_2} + \frac{1}{\sqrt{N}} \left\| \frac{a}{2M\sigma} (Y_N - Y_0) \right\|_{L_2} \\
&\leq \sqrt{N} \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}}\right)^{\frac{1}{2}} + \frac{a}{2M\sqrt{N}\sigma} \text{Var}^{\frac{1}{2}}(Y_N - Y_0) \\
&\leq \sqrt{N} \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}}\right)^{\frac{1}{2}} + \frac{a}{2M\sqrt{N}\sigma} \left(\frac{\sigma^2 \eta^2}{a} \right)^{\frac{1}{2}} \leq \sqrt{N} \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}}\right)^{\frac{1}{2}} + \frac{\sqrt{a} \eta}{2M\sqrt{N}}.
\end{aligned}$$

\square

Since, as $N \rightarrow \infty$, $N^{-1/2}(L(N) - N\mu)$ converges to a normal distribution with variance η^2 , the following corollary is immediate in Case (III).

Corollary 3.3.3. *Consider scenario (III). Under the assumptions of Theorem 3.3.1 we have $N \rightarrow \infty$ and $N/M \rightarrow 0$,*

$$\frac{1}{\sqrt{N}} \left(\widehat{L}_N^{(M)} - N\mu \right) \xrightarrow{d} \mathcal{N}(0, \eta^2).$$

3.3.2 Estimated unit increments

Inference on the covariance structure of L will require estimated increments of L over intervals of fixed length. Therefore, we look at finer properties of the estimated increments over unit intervals. Recall now notation (2.7.1). In the analogy we define

$$\Delta_1 L_n \equiv \Delta L_n^{(1)} = L_n - L_{n-1} \quad \text{and} \quad \Delta_1 \widehat{L}_n^{(M)} \equiv \widehat{L}_n^{(M)} - \widehat{L}_{n-1}^{(M)}. \quad (3.3.5)$$

We note that the latter notation indicates the increments over interval $(n-1, n]$, when the sampling frequency is M , as opposed to (2.7.1) where the increment over $((n-1)h, nh]$ is considered.

Using the inversion formula we have

$$\begin{aligned} \Delta_1 L_n &= L_n - L_{n-1} \\ &= \frac{1}{\sigma} \left[Y(n) - Y(0) + a \int_0^n Y(s) ds \right] - \frac{1}{\sigma} \left[Y(n-1) - Y(0) + a \int_0^{n-1} Y(s) ds \right] \\ &= \frac{1}{\sigma} \left[Y(n) - Y(n-1) + a \int_{n-1}^n Y(s) ds \right]. \end{aligned} \quad (3.3.6)$$

Furthermore, we represent

$$\Delta_1 \widehat{L}_1^{(M)} = \widehat{L}_1^{(M)} - \widehat{L}_0^{(M)} = \widehat{L}_1^{(M)}$$

and

$$\begin{aligned} \Delta_1 \widehat{L}_n^{(M)} &= \widehat{L}_n^{(M)} - \widehat{L}_{n-1}^{(M)} \\ &= \sum_{i=1}^{nM} \Delta \widehat{L}_{\frac{i}{M}} - \sum_{i=1}^{(n-1)M} \Delta \widehat{L}_{\frac{i}{M}} = \sum_{i=(n-1)M+1}^{nM} \Delta \widehat{L}_{\frac{i}{M}} \end{aligned}$$

$$= \frac{a}{M\sigma} \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}} + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) (Y_n - Y_{n-1}). \quad (3.3.7)$$

That is, the estimated increment over $(n-1, n]$ is represented as the sum of estimated increments over the small intervals $((n-1), (n-1) + \frac{1}{M}] , \dots$.

Properties of the estimated increments

We will prove some basic properties of the estimated increments $\Delta_1 \widehat{L}_n^{(M)}$, $n \geq 1$ defined in (3.3.5).

Remark 3.3.4. Since Y is a strictly stationary process with a finite second moment then it can be easily shown using (3.3.7) that $\Delta_1 \widehat{L}_n^{(M)}$, $n \geq 1$, is a second-order strictly stationary sequence.

In the next Lemmas we show how closely the estimated increments $\Delta_1 \widehat{L}_n^{(M)}$ approximate the true increments $\Delta_1 L_n$, $n \geq 1$. The estimate in (i) will be used to develop the asymptotic behavior of partial sums and sample covariances.

Lemma 3.3.5. *Consider Case (II) and Case (III). Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (3.1.1) holds. If $\Delta_1 L_n$, $\Delta_1 \widehat{L}_n^{(M)}$ are defined as in equation (3.3.5), then*

- (i) $\forall n \in \mathbb{N}$, $\left\| \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right\|_{L_2} \leq \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{\eta \sqrt{a}}{2M} (1 - e^{-a})^{\frac{1}{2}}$;
- (ii) $\forall n \in \mathbb{N}$, $\left\| \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right\|_{L_2} \rightarrow 0$ as $M \rightarrow \infty$.

In order to prove Theorem 3.3.5 we first prove the following uniform approximation.

Lemma 3.3.6. *Under the assumptions of Lemma 3.3.5*

$$\left\| \frac{1}{M} \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}} - \int_{n-1}^n Y(s) ds \right\|_{L_2} \leq \frac{\sigma \eta}{\sqrt{a}} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}}.$$

Proof: Recall Lemma 3.3.2 , for every $N, M \in \mathbb{Z}_+$,

$$\left\| \frac{1}{\sqrt{N}} \left(\int_0^N Y(s) ds - \frac{1}{M} \sum_{i=1}^{NM} Y_{\frac{i}{M}} \right) \right\|_{L_2} \leq \sqrt{N} \frac{\sigma \eta}{\sqrt{a}} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}}.$$

Now by letting $N = 1$ we have

$$\left\| \left(\int_0^1 Y(s) ds - \frac{1}{M} \sum_{i=1}^M Y_{\frac{i}{M}} \right) \right\|_{L_2} \leq \frac{\sigma \eta}{\sqrt{a}} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}}.$$

The result extends to all n by stationarity. \square

Proof of Lemma 3.3.5: (i)

By using equations (3.3.7) and (3.3.6) we compute,

$$\left(\Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right) = \frac{a}{\sigma} \left(\frac{1}{M} \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}} - \int_{n-1}^n Y(s) ds \right) - \frac{a}{2M\sigma} (Y_n - Y_{n-1}). \quad (3.3.8)$$

By taking the L_2 norm and using Lemma 3.3.6 we get:

$$\begin{aligned} & \left\| \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right\|_{L_2} \\ & \leq \frac{a}{\sigma} \left\| \frac{1}{M} \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}} - \int_{n-1}^n Y(s) ds \right\|_{L_2} + \left\| \frac{a}{2M\sigma} (Y_n - Y_{n-1}) \right\|_{L_2} \\ & \leq \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{a}{2M\sigma} \mathbb{E}^{\frac{1}{2}} [(Y_1 - Y_0)^2] \\ & = \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{a}{2M\sigma} (\text{Var} (Y_1 - Y_0))^{\frac{1}{2}} \\ & = \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \left(\frac{a}{2M\sigma} (2\text{Var} (Y_0) - 2\text{Cov}(Y_0, Y_1)) \right)^{\frac{1}{2}}. \end{aligned}$$

Using Lemma 2.4.3 we have

$$\begin{aligned} \left\| \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right\|_{L_2} & \leq \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{a}{2M\sigma} \left(2 \frac{\sigma^2 \eta^2}{2a} - 2 \frac{\sigma^2 \eta^2}{2a} e^{-a} \right)^{\frac{1}{2}} \\ & = \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{\eta \sqrt{a}}{2M} (1 - e^{-a})^{\frac{1}{2}}. \end{aligned}$$

(ii) The result is a consequence of (i), since

$$\eta\sqrt{a} \left(1 - e^{-\frac{a}{M}}\right)^{\frac{1}{2}} + \frac{\sqrt{a}\eta}{2M} (1 - e^{-a})^{\frac{1}{2}} \rightarrow 0 \text{ as } M \rightarrow \infty.$$

□

The following approximation lemma can be viewed as a corollary of Lemma 3.3.5.

Lemma 3.3.7. *Under the assumptions of Lemma 3.3.5, for $k \geq 0$,*

$$\begin{aligned} & \left\| \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_{n+k} \Delta_1 L_n \right\|_{L_1} \\ & \leq \left(\eta + \mathbb{E}^{\frac{1}{2}} \left[\Delta_1 \widehat{L}_1^{(M)} \right]^2 \right) \left(\eta\sqrt{a} \left(1 - e^{-\frac{a}{M}}\right)^{\frac{1}{2}} + \frac{\eta\sqrt{a}}{2M} (1 - e^{-a})^{\frac{1}{2}} \right). \end{aligned}$$

The bound converges to 0 as $M \rightarrow \infty$, uniformly in n and k .

Proof: By Lemma 3.3.5 (i) we have

$$\begin{aligned} & \mathbb{E} \left| \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_{n+k} \Delta_1 L_n \right| \\ & = \mathbb{E} \left| \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_{n+k} \Delta_1 L_n + \Delta_1 L_{n+k} \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_{n+k} \Delta_1 \widehat{L}_n^{(M)} \right| \\ & = \mathbb{E} \left| \Delta_1 L_{n+k} \left(\Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right) + \Delta_1 \widehat{L}_n^{(M)} \left(\Delta_1 \widehat{L}_{n+k}^{(M)} - \Delta_1 L_{n+k} \right) \right| \\ & \leq \mathbb{E} \left| \Delta_1 L_{n+k} \left(\Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right) \right| + \mathbb{E} \left| \Delta_1 \widehat{L}_n^{(M)} \left(\Delta_1 \widehat{L}_{n+k}^{(M)} - \Delta_1 L_{n+k} \right) \right| \\ & \leq \mathbb{E}^{\frac{1}{2}} \left[(\Delta_1 L_{n+k})^2 \right] \mathbb{E}^{\frac{1}{2}} \left[\left(\Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right)^2 \right] \\ & \quad + \mathbb{E}^{\frac{1}{2}} \left[\left(\Delta_1 \widehat{L}_n^{(M)} \right)^2 \right] \mathbb{E}^{\frac{1}{2}} \left[\left(\Delta_1 \widehat{L}_{n+k}^{(M)} - \Delta_1 L_{n+k} \right)^2 \right] \\ & = \eta \left\| \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right\|_{L_2} + \mathbb{E}^{\frac{1}{2}} \left[\left(\Delta_1 \widehat{L}_1^{(M)} \right)^2 \right] \left\| \Delta_1 \widehat{L}_{n+k}^{(M)} - \Delta_1 L_{n+k} \right\|_{L_2} \\ & \leq \left(\eta + \mathbb{E}^{\frac{1}{2}} \left[\left(\Delta_1 \widehat{L}_1^{(M)} \right)^2 \right] \right) \left(\eta\sqrt{a} \left(1 - e^{-\frac{a}{M}}\right)^{\frac{1}{2}} + \frac{\eta\sqrt{a}}{2M} (1 - e^{-a})^{\frac{1}{2}} \right). \end{aligned}$$

The bound converges to 0 by application of Lemma 3.3.5 (ii). □

Remark 3.3.8. We note that the bounds in Lemma 3.3.5 (i) and Lemma 3.3.7 are independent of N , so that convergence is uniform in N . As a consequence, we can use the above estimates both in Case (II) and (III).

Corollary 3.3.9. *Under the assumptions of Lemma 3.3.5,*

- (i) $\forall n \in \mathbb{N}, \mathbb{E} \left[\Delta_1 \widehat{L}_n^{(M)} \right] = \mu;$
- (ii) $\forall n \in \mathbb{N}, \text{Var} \left(\Delta_1 \widehat{L}_n^{(M)} \right) \rightarrow \text{Var} (\Delta_1 L_n) \text{ as } M \rightarrow \infty;$
- (iii) $\forall n > 1, \text{Cov} \left(\Delta_1 \widehat{L}_1^{(M)}, \Delta_1 \widehat{L}_n^{(M)} \right) \rightarrow 0.$

Proof:

By Remark 3.3.4 without loss of generality it is enough to prove (i) and (ii) for $n = 1$ only.

(i): Using equation (3.3.7) and Lemma 2.4.3 we have

$$\mathbb{E} \left[\Delta_1 \widehat{L}_1^{(M)} \right] = \frac{a}{M\sigma} \sum_{i=1}^M \mathbb{E} \left[Y_{\frac{i}{M}} \right] + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) \mathbb{E} [Y_1 - Y_0] = \mu.$$

(ii) is an immediate consequence of Lemma 3.3.5 (ii).

(iii) is an immediate consequence of Lemma 3.3.7. □

3.3.3 Asymptotics for the Sample Mean

The next result is a simple corollary of Theorem 3.3.1 and Corollary 3.3.3. We note that the law of large numbers requires $N \wedge M \rightarrow \infty$ only, while the central limit theorem needs $N \rightarrow \infty$ and $N/M \rightarrow 0$.

Corollary 3.3.10. *Consider scenario (III). Under the assumptions of Lemma 3.3.5. Let $\overline{\Delta_1 \widehat{L}^{(M)}}$ and $\widehat{\eta}^2$ be the sample mean and sample variance, respectively, of $\Delta_1 \widehat{L}_n^{(M)}$, $n = 1, \dots, N$: i.e.*

$$\overline{\Delta_1 \widehat{L}^{(M)}} \equiv \frac{1}{N} \sum_{n=1}^N \Delta_1 \widehat{L}_n^{(M)} \quad \text{and} \quad \widehat{\eta}^2 \equiv \frac{1}{N} \sum_{n=1}^N \left(\Delta_1 \widehat{L}_n^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right)^2.$$

Then,

- (i) $\overline{\Delta_1 \widehat{L}^{(M)}} \xrightarrow{p} \mu$ as $N \wedge M \rightarrow \infty$.
- (ii) $\sqrt{N} \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \mu \right) \xrightarrow{d} N(0, \eta^2)$ as $N \rightarrow \infty$ and $N/M \rightarrow 0$.
- (iii) $\widehat{\eta}^2 \xrightarrow{p} \eta^2$ as $N \wedge M \rightarrow \infty$.

Proof of (i), (ii): Note that

$$\overline{\Delta_1 \widehat{L}^{(M)}} = \frac{1}{N} \sum_{n=1}^{NM} \Delta \widehat{L}_{\frac{n}{M}}^{(M)} = \frac{1}{N} \widehat{L}_N^{(M)}.$$

By Theorem 3.3.1,

$$\begin{aligned} \left\| \frac{1}{N} \widehat{L}_N^{(M)} - \mu \right\|_{L_2} &\leq \frac{1}{\sqrt{N}} \left\| \frac{1}{\sqrt{N}} (L(N) - \widehat{L}_N^{(M)}) \right\|_{L_2} + \left\| \frac{L(N)}{N} - \mu \right\|_{L_2} \\ &\leq \eta \sqrt{a} (1 - e^{-a/M})^{1/2} + \frac{\sqrt{a}\eta}{2MN} + \left\| \frac{L(N)}{N} - \mu \right\|_{L_2}. \end{aligned}$$

Let $N \wedge M \rightarrow \infty$ and (i) follows. Part (ii) is an immediate consequence of Corollary 3.3.3.

Proof of (iii): We have

$$\begin{aligned} \mathbb{E} \left| \frac{1}{N} \sum_{n=1}^N \left(\Delta_1 \widehat{L}_n^{(M)} \right)^2 - \frac{1}{N} \sum_{n=1}^N \left(\Delta_1 L_n \right)^2 \right| &\leq \frac{1}{N} \sum_{n=1}^N \mathbb{E} \left| \left(\Delta_1 \widehat{L}_n^{(M)} \right)^2 - \left(\Delta_1 L_n \right)^2 \right| \\ &\rightarrow 0 \text{ as } M \rightarrow \infty \text{ (by Lemma 3.3.7, } k = 0\text{)}. \end{aligned}$$

Now since,

$$\frac{1}{N} \sum_{n=1}^N \left(\Delta_1 L_n \right)^2 \xrightarrow{p} \mathbb{E} [L_1]^2 \text{ as } N \rightarrow \infty$$

then

$$\frac{1}{N} \sum_{n=1}^N \left(\Delta_1 \widehat{L}_n^{(M)} \right)^2 \xrightarrow{p} \mathbb{E} [L_1]^2 = \eta^2 + \mu^2 \text{ as } N \wedge M \rightarrow \infty;$$

consequently the result follows by (i). □

3.3.4 Asymptotics for the Sample Covariances

In this section we consider the sample covariances of the estimated unit increments, defined by

$$\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k) \equiv \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta_1 \widehat{L}_{n+k}^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right) \left(\Delta_1 \widehat{L}_n^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right),$$

where

$$\overline{\Delta_1 \widehat{L}^{(M)}} = \frac{1}{N} \sum_{n=1}^N \Delta_1 \widehat{L}_n^{(M)}.$$

and $k > 1$.

Again, we note that the sample covariance here has a different meaning than the one in (3.1.4). There, “lag k ” means that we look at dependence between $\Delta \widehat{L}_n^{(h)}$ and $\Delta \widehat{L}_{n+k}^{(h)}$ ($n \geq 1$), that is the estimated increment over $((n-1)h, nh]$ and the estimated increment over $((n+k-1)h, (n+k)h]$. In the present setting we look at dependence between the estimated increments over $((n-1), n]$ and $((n+k-1), (n+k)]$, when the CAR(1) process is sampled at frequency $h = 1/M$, $M \rightarrow \infty$.

Theorem 3.3.11. *Consider scenario (III). Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (3.1.1) holds. Then*

- (i) $\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k) \xrightarrow{p} 0$ as $N \wedge M \rightarrow \infty \quad \forall k \geq 1$;
- (ii) $\sqrt{N} \widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k) \xrightarrow{d} N(0, \eta^4)$ as $N \rightarrow \infty$ and $N/M \rightarrow 0 \quad \forall k \geq 1$.

Remark 3.3.12.

- (i) For the sample covariances we can assume without loss of generality that $\mathbb{E} \left[\Delta_1 \widehat{L}_1^{(M)} \right] = \mu = 0$ because in case $\mu \neq 0$ we can consider $\Delta_1 \widetilde{L}_n^{(M)} = \Delta_1 \widehat{L}_n^{(M)} - \mu$. Then

$$\begin{aligned} & \widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k) \\ &= \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta_1 \widehat{L}_{n+k}^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right) \left(\Delta_1 \widehat{L}_n^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta_1 \widehat{L}_{n+k}^{(M)} - \mu - \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \mu \right) \right) \left(\Delta_1 \widehat{L}_n^{(M)} - \mu - \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \mu \right) \right) \\
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta_1 \widetilde{L}_{n+k}^{(M)} - \overline{\Delta_1 \widetilde{L}^{(M)}} \right) \left(\Delta_1 \widetilde{L}_n^{(M)} - \overline{\Delta_1 \widetilde{L}^{(M)}} \right) \\
&= \widehat{\gamma_{\Delta_1 \widetilde{L}^{(M)}}}(k).
\end{aligned}$$

(ii) L is a Lévy process, so it has stationary independent increments. If $X_n = \Delta L_{n+k} \Delta L_n$, $k \geq 1$, $n \geq 1$, then X_n is a strictly stationary k -dependent sequence with mean zero (assuming $\mu = 0$) and autocovariance function:

$$\begin{aligned}
\gamma_X(n-1) &= \text{Cov}(\Delta L_{1+k} \Delta L_1, \Delta L_{n+k} \Delta L_n) \\
&= \mathbb{E}[\Delta L_{1+k} \Delta L_1 \Delta L_{n+k} \Delta L_n] \\
&= \begin{cases} \mathbb{E}[(\Delta L_{1+k})^2] \mathbb{E}[(\Delta L_1)^2] & \text{if } n = 1 \\ 0 & \text{if } 1 < n \leq k \\ \mathbb{E}[(\Delta L_{1+k})^2] \mathbb{E}[\Delta L_1 \Delta L_{2k+1}] & \text{if } n = k + 1 \\ 0 & \text{if } n > k + 1 \end{cases}.
\end{aligned}$$

This implies

$$\gamma_X(n) = \begin{cases} \eta^4 & \text{if } n = 0 \\ 0 & \text{if } n > 0 \end{cases}.$$

Using Theorem 6.4.2 in [10] we have:

$$\frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} X_n = \frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta L_{n+k} \Delta L_n \xrightarrow{d} N(0, \eta^4).$$

Proof of Theorem 3.3.11: (i): We decompose the sample covariance as

$$\begin{aligned}
&\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k) \\
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta_1 \widehat{L}_{n+k}^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right) \left(\Delta_1 \widehat{L}_n^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \left(\Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 \widehat{L}_{n+k}^{(M)} \overline{\Delta_1 \widehat{L}^{(M)}} - \Delta_1 \widehat{L}_n^{(M)} \overline{\Delta_1 \widehat{L}^{(M)}} + \overline{\Delta_1 \widehat{L}^{(M)}}^2 \right) \\
&= \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \\
&\quad - \overline{\Delta_1 \widehat{L}^{(M)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_n^{(M)} + \left(\overline{\Delta_1 \widehat{L}^{(M)}} \right)^2. \tag{3.3.9}
\end{aligned}$$

For the last three terms in equation (3.3.9) we use Corollary 3.3.10 (i) with $\mu = 0$.

Then we have as $N \wedge M \rightarrow \infty$ that:

- $\left(\overline{\Delta_1 \widehat{L}^{(M)}} \right)^2 \xrightarrow{p} 0$;
- $\overline{\Delta_1 \widehat{L}^{(M)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \xrightarrow{p} 0$, since $\overline{\Delta_1 \widehat{L}^{(M)}} \xrightarrow{p} 0$ and $\frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \xrightarrow{p} 0$;
- $\overline{\Delta_1 \widehat{L}^{(M)}} \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_n^{(M)} \xrightarrow{p} 0$, since $\overline{\Delta_1 \widehat{L}^{(M)}} \xrightarrow{p} 0$ and $\frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_n^{(M)} \xrightarrow{p} 0$.

For the first term of equation (3.3.9) we have for all $n \in \mathbb{N}$,

$$\begin{aligned}
&\left\| \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n \right\|_{L_1} \\
&\leq \frac{1}{N-k} \sum_{n=1}^{N-k} \left\| \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_{n+k} \Delta_1 L_n \right\|_{L_1} \\
&\leq \left(\eta + \mathbb{E}^{\frac{1}{2}} \left[\left(\Delta_1 \widehat{L}_1^{(M)} \right)^2 \right] \right) \left(\eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{\eta \sqrt{a}}{2M} \left(1 - e^{-a} \right)^{\frac{1}{2}} \right) \tag{3.3.10}
\end{aligned}$$

by Lemma 3.3.7. The bound converges to 0 as $M \rightarrow \infty$. Hence,

$$\frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n - \frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} \xrightarrow{p} 0 \text{ as } N \wedge M \rightarrow \infty.$$

Using Remark 3.3.12 (ii) we have

$$\frac{1}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n \xrightarrow{p} 0 \text{ as } N \rightarrow \infty.$$

This completes the proof of (i).

(ii): We use again the decomposition (3.3.9). The last three terms of equation (3.3.9) multiplied by \sqrt{N} are negligible as $N \rightarrow \infty$ and $N/M \rightarrow 0$ by a similar argument as in part (i), now using Corollary 3.3.10 (ii).

For the first term of equation (3.3.9) multiplied by \sqrt{N} , we claim that

$$\frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n \xrightarrow{p} 0 \text{ as } N \rightarrow \infty \text{ and } N/M \rightarrow 0. \quad (3.3.11)$$

Indeed by using equation (3.3.10) we have,

$$\begin{aligned} & \left\| \frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n \right\|_{L_1} \\ & \leq \sqrt{N} \left(\eta + \mathbb{E}^{\frac{1}{2}} \left[\Delta_1 \widehat{L}_1^{(M)} \right]^2 \right) \left(\eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{\eta \sqrt{a}}{2M} \left(1 - e^{-a} \right)^{\frac{1}{2}} \right) = O \left(\sqrt{\frac{N}{M}} \right). \end{aligned} \quad (3.3.12)$$

The bound converges to 0 since $N/M \rightarrow 0$. This completes the proof of the claim (3.3.11). Now, we are ready to finish the proof of (ii). Using Remark 3.3.12, (ii) we have that:

$$\frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n \xrightarrow{d} N(0, \eta^4) \text{ as } N \rightarrow \infty.$$

Now,

$$\frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)}$$

$$\begin{aligned}
&= \frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 \widehat{L}_{n+k}^{(M)} \Delta_1 \widehat{L}_n^{(M)} - \frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n + \frac{\sqrt{N}}{N-k} \sum_{n=1}^{N-k} \Delta_1 L_{n+k} \Delta_1 L_n \\
&\xrightarrow{d} N(0, \eta^4),
\end{aligned}$$

as $N \rightarrow \infty$ and $N/M \rightarrow 0$. So,

$$\sqrt{N} \widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(k) \xrightarrow{d} N(0, \eta^4) \text{ as } N \rightarrow \infty \text{ and } N/M \rightarrow 0,$$

completing the proof of (ii). □

3.4 CLT for the Lévy-driven CAR(1) process

This section is a brief digression from the central topic of our thesis. Here we use our estimator of the driving process L to prove a central limit theorem for the integrated CAR(1) process Y . The significance of this result is that the proof is quite elementary and does not require any mixing arguments. In Corollary 3.3.3, we have proven a CLT for the estimated process $\widehat{L}_N^{(M)}$ by showing its closeness to the true Lévy process $L(N)$. In this section, a similar approach leads to a CLT for the sampled process $(Y_n^{(1/M)})$ by using the CLT for $\widehat{L}_N^{(M)}$, see Theorem 3.4.2. We note that results similar to those in Corollary 3.4.1 and Theorem 3.4.2 can be obtained using different methods. In [44], mixing properties of the CAR(1) model are utilized, while sample means and sample covariances for the discretely sampled process Y are considered in [19]. In our Scenario I (fixed h) asymptotic normality of the sample mean and sample covariances are proven under appropriate moment assumptions.

Corollary 3.4.1. *Consider scenario (III). Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (3.1.1) holds. Then*

(i) *As $N \wedge M \rightarrow \infty$,*

$$\frac{1}{NM} \sum_{n=1}^{NM} (Y_{\frac{n}{M}} - \mathbb{E}[Y_0]) \xrightarrow{p} 0$$

(ii) *As $N \rightarrow \infty$ and $N/M \rightarrow 0$*

$$\frac{1}{\sqrt{N}} \frac{1}{M} \sum_{n=1}^{NM} (Y_{\frac{n}{M}} - \mathbb{E}[Y_0]) \xrightarrow{d} N\left(0, \frac{\eta^2 \sigma^2}{a^2}\right).$$

Proof: From (3.3.7) we have

$$(\widehat{L}_N^{(M)} - N\mu) = \left(\frac{a}{M\sigma} \sum_{n=1}^{NM} Y_{\frac{n}{M}} - N\mu \right) + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) (Y_N - Y_0).$$

Therefore, since Y is stationary, (i) follows by Corollary 3.3.10 (i) and (ii) follows by Corollary 3.3.3. \square

In fact, Corollary 3.4.1 gives us a very simple proof of the CLT for the integrated Lévy-driven CAR(1) process Y :

Theorem 3.4.2. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (3.1.1) holds. Then*

$$\lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \int_0^N (Y(s) - \mathbb{E}[Y_0]) ds \stackrel{d}{=} N \left(0, \frac{\eta^2 \sigma^2}{a^2} \right).$$

Proof: If we center equation (3.3.3), we have, for every N

$$\begin{aligned} & \left\| \frac{1}{\sqrt{N}} \int_0^N (Y(s) - \mathbb{E}[Y_0]) ds - \frac{1}{\sqrt{N}} \frac{1}{M} \sum_{n=1}^{NM} (Y_{\frac{n}{M}} - \mathbb{E}[Y_0]) \right\|_{L_2} \\ & \leq \sqrt{N} \frac{\sigma \eta}{\sqrt{a}} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} = O\left(\sqrt{N/M}\right). \end{aligned}$$

Let

$$A_N = \frac{1}{\sqrt{N}} \int_0^N (Y(s) - \mathbb{E}[Y_0]) ds$$

and

$$B_{N,M} = \frac{1}{\sqrt{N}} \frac{1}{M} \sum_{n=1}^{NM} (Y_{\frac{n}{M}} - \mathbb{E}[Y_0]).$$

Arguing as in Theorem 25.4 of [7], for any $y' < x < y''$, with $y'' - x < \epsilon$, $x - y' < \epsilon$,

$$\begin{aligned} P(B_{N,M} \leq y') - P(|A_N - B_{N,M}| \geq \epsilon) & \leq P(A_N \leq x) \\ & \leq P(B_{N,M} \leq y'') + P(|A_N - B_{N,M}| \geq \epsilon) \end{aligned}$$

for all N, M . Therefore, letting $N \rightarrow \infty$ and choosing M such that $N/M \rightarrow 0$, by Corollary 3.4.1,

$$P(W \leq y') \leq \liminf_{N \rightarrow \infty} P(A_N \leq x) \leq \limsup_{N \rightarrow \infty} P(A_N \leq x) \leq P(W \leq y''),$$

where W is normal with mean zero and variance $\eta^2 \sigma^2 / a^2$. Since ϵ is arbitrary, the result follows. □

Chapter 4

Inference based on the sampled process

We now return to the central topic of our thesis. In this chapter, we develop a test of hypothesis that the driving process L in our CAR(1) model has uncorrelated increments. Recall equation (2.2.2)

$$\mathbb{E}[L(t)] = \mu t \quad \text{and} \quad \text{Var}(L(t)) = \eta^2 t, \quad t \geq 0. \quad (4.0.1)$$

In the following sections we formulate the test statistic as well as study the performance of the test under both the null hypothesis (sections 4.2.1, 4.2.2, 4.2.3) and the alternative hypothesis (section 4.2.4).

In summary, the performance is very good for Brownian motion and the Poisson process, but slightly worse for the gamma process.

4.1 Test Statistics

Let $(Y_n^{(1/M)}, n = 1, \dots, NM)$ be a discretely sampled stochastic process. If Y is a CAR(1) model driven by a process L we can use the estimated increments to test H_0 that L has uncorrelated increments, which will be true if L is a Lévy process. We

reject H_0 for a large absolute value of the statistic $W_{\Delta_1 \widehat{L}^{(M)}}(k)$ for a specified value of k , where $W_{\Delta_1 \widehat{L}^{(M)}}(k)$ is defined as follows:

Corollary 4.1.1. *Consider scenario (III). Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (4.0.1) holds. If $\widehat{\eta}^2$ and $\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k)$ are the sample variance and covariance of $\widehat{\Delta_1 \widehat{L}^{(M)}}$ respectively, then,*

$$W_{\Delta_1 \widehat{L}^{(M)}}(k) \equiv \sqrt{N} \frac{\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k)}{\widehat{\eta}^2} \xrightarrow{d} N(0, 1) \text{ as } N \rightarrow \infty \text{ and } \frac{N}{M} \rightarrow 0.$$

Proof: By Theorem 3.3.11 we have shown that

$$\sqrt{N} \widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k) \xrightarrow{d} N(0, \eta^4) \text{ as } N \rightarrow \infty \text{ and } \frac{N}{M} \rightarrow 0.$$

Consequently,

$$\sqrt{N} \frac{\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k)}{\eta^2} \xrightarrow{d} N(0, 1).$$

Now, $\frac{\eta^2}{\widehat{\eta}^2} \xrightarrow{p} 1$ by Corollary 3.3.10, hence

$$\begin{aligned} \sqrt{N} \frac{\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k)}{\widehat{\eta}^2} &= \sqrt{N} \frac{\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(k)}{\eta^2} \frac{\eta^2}{\widehat{\eta}^2} \\ &\xrightarrow{d} N(0, 1) \text{ by Slutsky's theorem.} \end{aligned}$$

□

Under H_0 , for large N, M and N/M small we have

$$\alpha \approx P(|W_{\Delta_1 \widehat{L}^{(M)}}(k)| > z_{\alpha/2}), \quad (4.1.1)$$

where $z_{\alpha/2}$ is the critical Z -value where $Z \sim N(0, 1)$.

4.2 Simulation Study

4.2.1 Brownian motion driven CAR(1) process

The Brownian motion B is defined in Section 2.2.1. For a large K , we simulate an i.i.d. sequence (noise) $Z_{\frac{i}{K}} \sim N(0, \frac{1}{K})$, $i = 1, 2, \dots, NK$. We approximate the

driving process $B(t)$ as,

$$B(t) = \begin{cases} 0 & \text{if } t = 0 \\ \sum_{i=1}^n Z_{\frac{i}{K}} & \text{if } \frac{n-1}{K} < t \leq \frac{n}{K}, \end{cases}$$

which will be treated as an approximation to a standard Brownian motion on $[0, N]$.

Figure 4.1 shows a simulated driving process B_t .

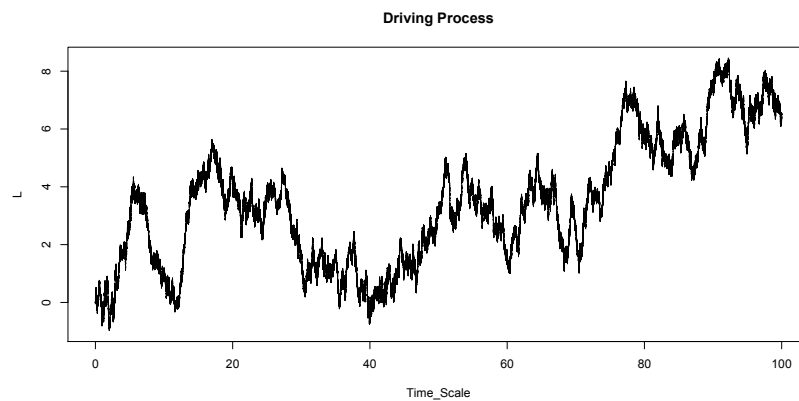


Figure 4.1: Simulated standard Brownian motion for $N = 100$ and $K = 5000$.

In order to simulate Y_t , Brownian motion-driven CAR(1) process, we look at its definition through its stochastic differential equation

$$dY(t) = -aY(t) dt + \sigma dB(t)$$

which by Euler's scheme can be approximated by the difference equation:

$$Y_t - Y_{t-\frac{1}{K}} = -aY_{t-\frac{1}{K}} \frac{1}{K} + \sigma \left(B_t - B_{t-\frac{1}{K}} \right),$$

where t is of the form $t = i/K$, $i = 1, \dots, NM$. Using the simulated noise $Z_{\frac{i}{K}}$ we have,

$$Y_t = \left(1 - \frac{a}{K}\right) Y_{t-\frac{1}{K}} + \sigma Z_{\frac{i}{K}}.$$

Figure 4.2 shows a simulated sample path Y_t with $K = 5000$ and $N = 100$ and Y_{Sampled} which are the values of Y_t observed at the times $\{0, \frac{1}{M}, \frac{2}{M}, \dots, N\}$, with $M = 500$.

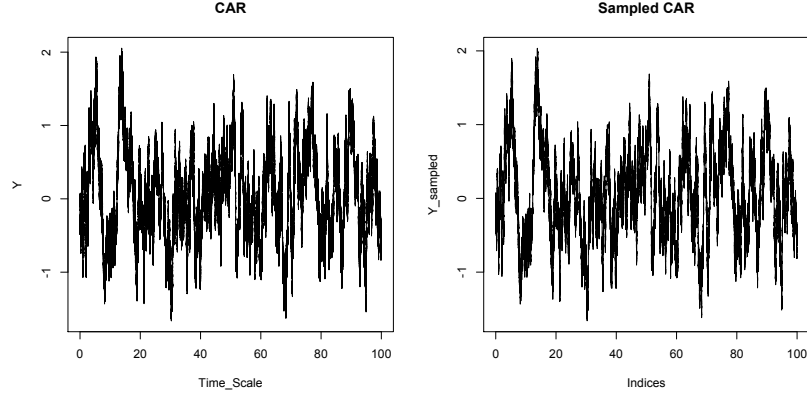


Figure 4.2: Simulated Brownian motion driven CAR(1) process Y for $\{N = 100, K = 5000, \mu = 0, \eta = 1, \sigma = 1, a = 1\}$ and Y_{Sampled} for $M = 500$

We recall (3.3.7):

$$\Delta_1 \widehat{L}_n^{(M)} = \frac{a}{M\sigma} \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}} + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) (Y_n - Y_{n-1}).$$

We compute the estimates of the recovered increments $\Delta_1 \widehat{B}_n^{(M)}$ over the intervals $(i-1, i], i = 1, \dots, 1 = 100$, using Y_{Sampled} illustrated in Figure 4.2.

In Figure 4.3 below, we compare the estimated increments with the true increments $\Delta_1 B_n = B_n - B_{n-1}$.

To show that $\Delta_1 B_n$ and $\Delta_1 \widehat{B}_n^{(M)}$ are not identical, we display the differences in Figure 4.4.

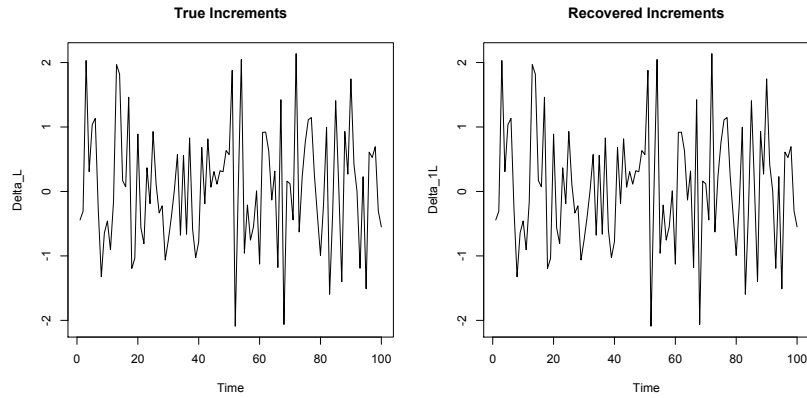


Figure 4.3: True Increments: $\Delta_1 B_n$, Recovered Increments: $\Delta_1 \widehat{B}_n^{(M)}$

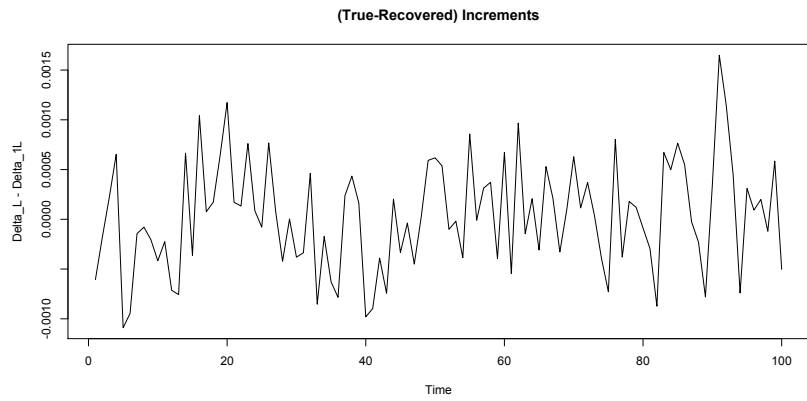


Figure 4.4: (True-Recovered) Increments: $\Delta_1 B_n - \Delta_1 \widehat{B}_n^{(M)}$

In Figure 4.5, we compute the sample autocorrelations function for both the recovered increments ($\Delta_1 \widehat{B}_n^{(M)}$) and the true increments ($\Delta_1 B_n$). The sample autocorrelations are very similar for the true and estimated increments, in agreement with equation (3.3.12). Figure 4.5 also reflects the 0 correlation of the true increments and the asymptotic 0 correlation of the estimated increments. This is in agreement with Theorem 3.3.11 and Remark 3.3.12.

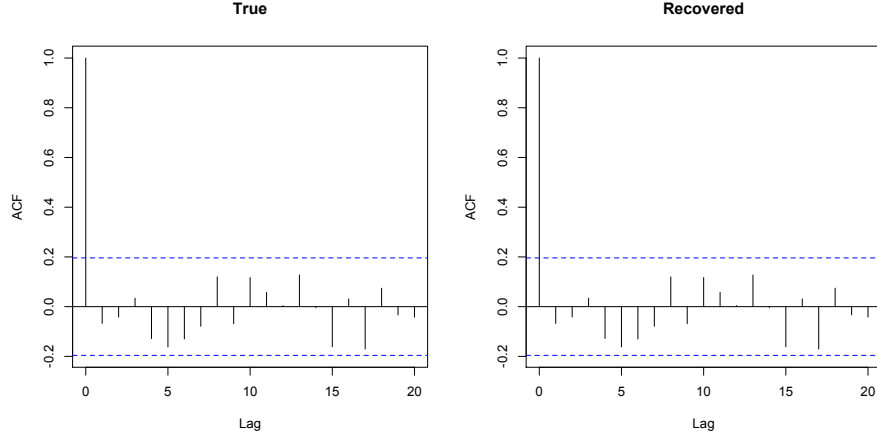


Figure 4.5: The sample autocorrelation function for the True Increments $\Delta_1 B_n$ (True) and for the Recovered Increment $\Delta_1 \widehat{B}_n^{(M)}$ (Recovered)

We want to test at level 0.05

$$H_0 : \rho(1) = 0 \quad \text{vs.} \quad H_1 : \rho(1) \neq 0 \quad \text{where } \rho(1) = \text{Corr}(\Delta_1 B_n, \Delta_1 B_{n+1}).$$

To assess the performance of our proposed test statistic based on estimated increments (see Lemma 4.1.1), we compare it with the corresponding statistic based on the true increments; that is, we compare the performance of the statistics:

$$W_{\Delta_1 B}(1) = \sqrt{N} \frac{\widehat{\gamma}_{\Delta_1 B}(1)}{\widehat{\eta}^2} = \sqrt{N} \widehat{\rho}_{\Delta_1 B}(1), \quad W_{\Delta_1 \widehat{B}^{(M)}}(1) = \sqrt{N} \frac{\widehat{\gamma}_{\Delta_1 \widehat{B}^{(M)}}(1)}{\widehat{\eta}^2} = \sqrt{N} \widehat{\rho}_{\Delta_1 \widehat{B}^{(M)}}(1)$$

where $W_{\Delta_1 B}(1)$ is based on the (unobserved) true increments $\Delta_1 B_n$, and $W_{\Delta_1 \widehat{B}^{(M)}}(1)$ is based on the recovered increments $\Delta_1 \widehat{B}_n^{(M)}$. See Corollary 4.1.1 and equation (4.1.1).

Tables 4.1 and 4.2 give the empirical levels $\widehat{\alpha}_{\Delta_1 B}, \widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$ for both tests based on $W_{\Delta_1 B}(1)$ and $W_{\Delta_1 \widehat{B}^{(M)}}(1)$ respectively, over $R = 400$ simulations, with nominal level 0.05. We consider various values of the parameters a and σ .

	$N = 50, M = 100$		$N = 100, M = 100$		$N = 100, M = 300$		$N = 100, M = 500$	
	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$
a=0.9	0.0600	0.0600	0.0375	0.0375	0.0375	0.0375	0.0375	0.0375
a=10	0.0375	0.0400	0.0375	0.0400	0.0475	0.0475	0.0450	0.0450
a=100	0.0325	0.0250	0.0350	0.0475	0.0325	0.0350	0.0500	0.0500
a=1000	0.0375	0.0450	0.0500	0.0450	0.0375	0.0400	0.0400	0.0425

Table 4.1: We fix $\{\sigma = 1, K = 5000, \mu = 0, R = 400\}$

These results are consistent with a nominal level 0.05, since with $R = 400$, the empirical level should fall in the range 0.05 ± 0.021 95% of the time.

	$N = 50, M = 100$		$N = 100, M = 100$		$N = 100, M = 300$		$N = 100, M = 500$	
	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$	$\hat{\alpha}_{\Delta_1 B}$	$\hat{\alpha}_{\Delta_1 \hat{B}^{(M)}}$
$\sigma = 1$	0.0600	0.0600	0.0375	0.0375	0.0375	0.0375	0.0375	0.0375
$\sigma = 10$	0.0625	0.0625	0.0600	0.0600	0.0625	0.0625	0.0625	0.0625
$\sigma = 100$	0.0600	0.0600	0.0425	0.0425	0.0575	0.0575	0.0500	0.0500
$\sigma = 1000$	0.0425	0.0425	0.0525	0.0550	0.0775	0.0775	0.0650	0.0650

Table 4.2: We fix $\{a = 0.9, K = 5000, \mu = 0, R = 400\}$

The test statistics based on the true and recovered increments give us virtually identical empirical levels except for large values of a ($a = 100$ or $a = 1000$). Looking at the formula for the recovered noise (3.3.7), we see that large values of a introduce more volatility and so this result is to be expected.

For a Gaussian driving process, the performance of the test statistics does not seem to be particularly sensitive to the sampling frequency M or the value of the ratio N/M . Also, a sample $N = 50$ seems adequate.

4.2.2 Gamma-driven CAR(1) process

The gamma process G is defined in Section 2.2.1.

Following the same steps as in the case of Brownian motion driven CAR(1) process we simulate the driving process $G(t)$ using the discrete approximation

$$G(t) = \begin{cases} 0 & \text{if } t = 0 \\ \sum_{i=1}^n \Gamma_{\frac{i}{K}} & \text{if } \frac{n-1}{K} < t \leq \frac{n}{K}, \end{cases}$$

where $\Gamma_{\frac{i}{K}} \sim$ i.i.d. $\Gamma(\alpha\frac{1}{K}, \beta) = \Gamma\left(\frac{\mu^2}{\eta^2}\frac{1}{K}, \frac{\eta^2}{\mu}\right)$, $i = 1, 2, \dots, NK$. Figure 4.6 illustrates a simulated sample path $G(t)$.

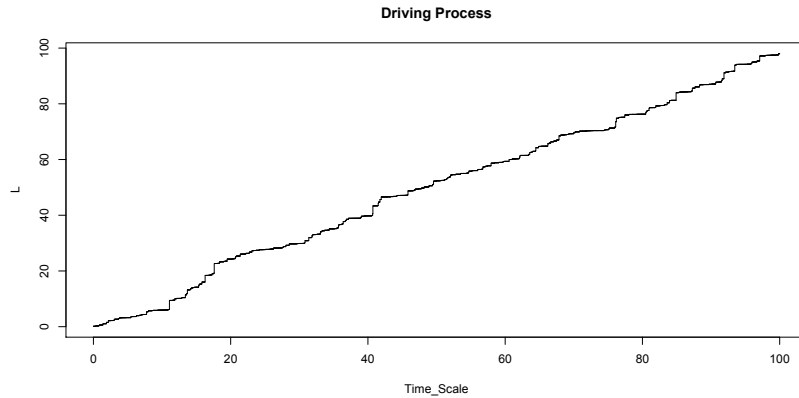


Figure 4.6: Simulated gamma Process for $\{N = 100, K = 5000, \mu = 1, \eta = 1\}$

We also use Euler's scheme to approximate, Y_t , a gamma-driven CAR(1) process, by the discrete equation:

$$Y_t - Y_{t-\frac{1}{K}} = -aY_{t-\frac{1}{K}}\frac{1}{K} + \sigma\Gamma_{\frac{t}{K}},$$

where t is of the form $t = i/K$, $i = 1, \dots, NM$. In this case we begin with any chosen value for Y_0 then use Euler's scheme to approximate the process Y . We drop the first 2000 simulated values in order to achieve stationarity.

Figure 4.7 shows a sample path Y_t with $K = 5000$ and $N = 100$ and Y_{Sampled} which are the values of Y_t observed at the times $\{0, \frac{1}{M}, \frac{2}{M}, \dots, N\}$ with $M = 500$.

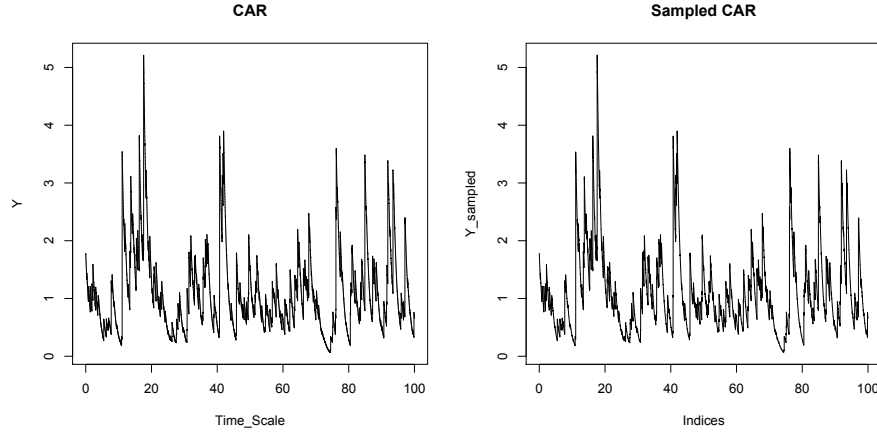


Figure 4.7: Simulated gamma-driven CAR(1) process for $\{N = 100, K = 5000, \mu = 1, \eta = 1, \sigma = 1, a = 1\}$ and Y_{Sampled} for $M = 500$

In Figure 4.8 we display $\Delta_1 G_n = G_n - G_{n-1}$ and $\Delta_1 \widehat{G}_n^{(M)}$ computed by equation (3.3.7), using Y_{Sampled} from Figure 4.7.

Also, the differences $(\Delta_1 G_n - \Delta_1 \widehat{G}_n^{(M)})$ are displayed in Figure 4.9 as well as the sample autocorrelation functions in Figure 4.10. As before, the sample autocorrelations support our theoretical results.

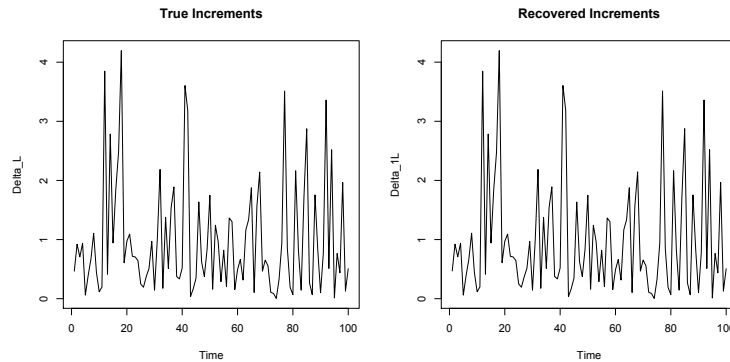


Figure 4.8: True Increments: $\Delta_1 G_n$, Recovered Increments: $\Delta_1 \widehat{G}_n^{(M)}$

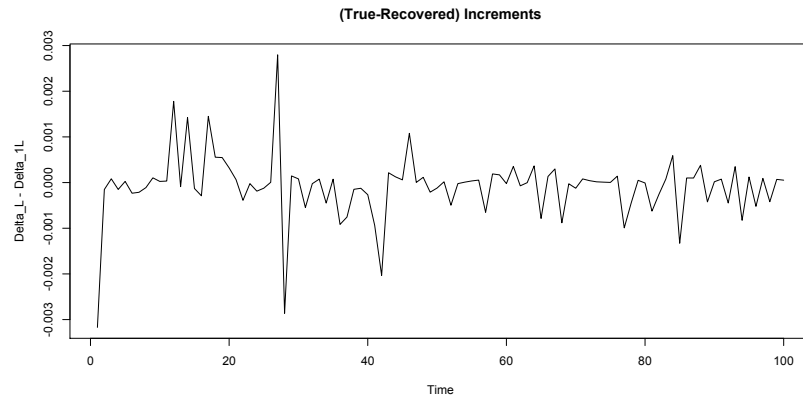


Figure 4.9: (True-Recovered) Increments: $\Delta_1 G_n - \Delta_1 \widehat{G}_n^{(M)}$ ($M = 500$)

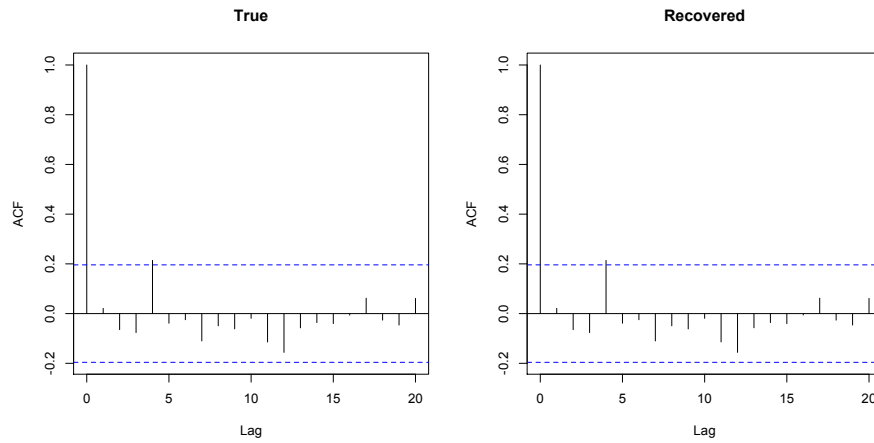


Figure 4.10: The sample autocorrelation function for the True Increments $\Delta_1 G_n$ (True) and for the Recovered Increments $\Delta_1 \widehat{G}_n^{(M)}$ (Recovered).

Tables 4.3 and 4.4 give computed empirical levels $\hat{\alpha}_{\Delta_1 G}, \hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$ for tests based on $W_{\Delta_1 G_n}(1)$ and $W_{\Delta_1 \hat{G}^{(M)}}(1)$ respectively, following the same procedures as before, over $R = 400$ simulations, with nominal level 0.05. We consider various values of the parameters a and σ as well.

	$N = 50, M = 100$		$N = 100, M = 100$		$N = 100, M = 300$		$N = 100, M = 500$	
	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$
$a = 0.9$	0.0200	0.0200	0.0725	0.0725	0.0500	0.0500	0.0375	0.0375
$a = 10$	0.0250	0.0250	0.0550	0.0550	0.0400	0.0425	0.0425	0.0450
$a = 100$	0.0350	0.0250	0.0450	0.0325	0.0500	0.0525	0.0375	0.0350
$a = 1000$	0.0300	0.0175	0.0475	0.0200	0.0300	0.0225	0.0300	0.0300

Table 4.3: We fix $\{\sigma = 1, \mu = 1, \eta = 1, K = 5000, R = 400\}$

	$N = 50, M = 100$		$N = 100, M = 100$		$N = 100, M = 300$		$N = 100, M = 500$	
	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$
$\sigma = 1$	0.0200	0.0200	0.0725	0.0725	0.0500	0.0500	0.0375	0.0375
$\sigma = 10$	0.0425	0.0350	0.0375	0.0425	0.0525	0.0500	0.0500	0.0525
$\sigma = 100$	0.0350	0.0350	0.0250	0.0325	0.0525	0.0475	0.0325	0.0300
$\sigma = 1000$	0.0375	0.0300	0.0400	0.0425	0.0500	0.0525	0.06250	0.0650

Table 4.4: We fix $\{a = 0.9, \mu = 1, \eta = 1, K = 5000, R = 400\}$

We note larger discrepancies between the empirical levels $\hat{\alpha}_{\Delta_1 G}$ and $\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$ for small M than we did for Brownian motion. This is likely due to the larger discrepancies between the true and recovered increments as illustrated in Figure 4.11, for $M = 100$. As a result, in the case of a gamma process driving function, the test statistics are more sensitive to the sampling frequency M and to the sample size N . However, the ratio N/M does not appear to play a significant rule.

In the left hand plot in Figure 4.11, we see the true-recovered increments with $M = 100$. There are four large peaks that obscure the remaining differences. Removing these peaks allow us to illustrate the remaining differences in the right hand plot.

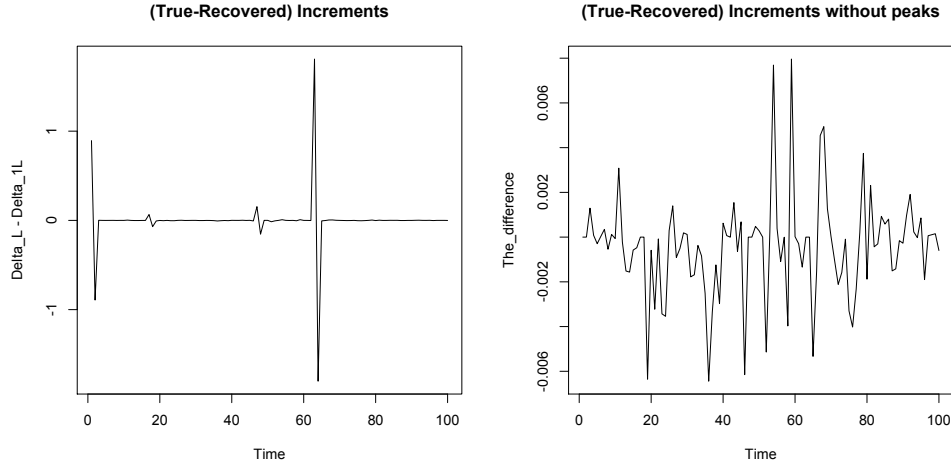


Figure 4.11: (True-Recovered) Increments with and without peaks for small $M = 100$

4.2.3 Poisson-driven CAR(1) process

The Poisson process P is defined in Section 2.2.1.

Following the same steps as in the case of Brownian motion driven CAR(1) process we simulate the driving process $P(t)$ using the discrete approximation

$$P(t) = \begin{cases} 0 & \text{if } t = 0 \\ \sum_{i=1}^n P_{\frac{i}{K}} & \text{if } \frac{n-1}{K} < t \leq \frac{n}{K}, \end{cases}$$

where $P_{\frac{i}{K}} \sim$ i.i.d. $P(\lambda \frac{1}{K})$, $i = 1, 2, \dots, NK$. Figure 4.12 illustrates a simulated sample path of P .

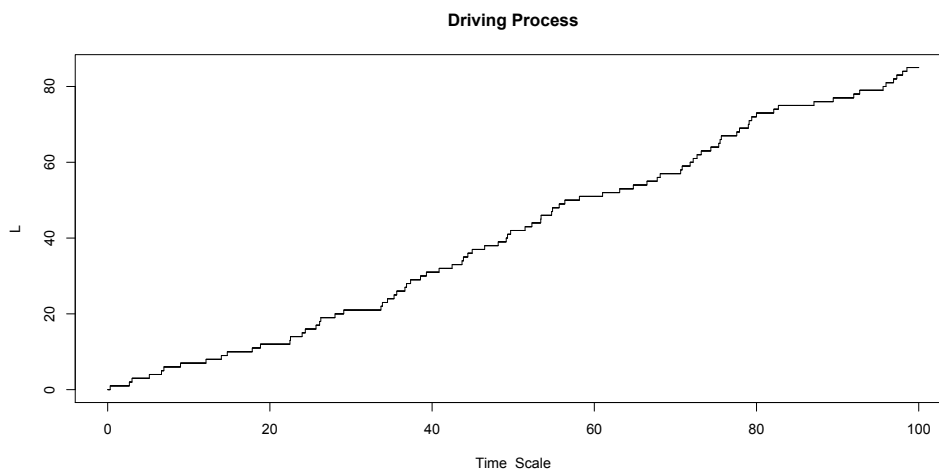


Figure 4.12: Simulated Poisson Process for $\{N = 100, K = 5000, \lambda = 1\}$

We also use Euler's scheme to approximate, Y_t , Poisson-driven CAR(1) process, by the discrete equation:

$$Y_t - Y_{t-\frac{1}{K}} = -aY_{t-\frac{1}{K}} \frac{1}{K} + \sigma P_{\frac{i}{K}},$$

where t is of the form $t = i/K$, $i = 1, \dots, NM$.

In the same way as in the gamma driving CAR(1) process we drop the first 2000 simulations to be closer to stationarity.

Figure 4.13 shows a sample path Y_t with $K = 5000$ and $N = 100$ and Y_{Sampled} which are the values of Y_t observed at the times $\{0, \frac{1}{M}, \frac{2}{M}, \dots, N\}$ with $M = 500$.

In Figure 4.14 we display $\Delta_1 P_n = P_n - P_{n-1}$ and $\Delta_1 \widehat{P}_n^{(M)}$ computed by equation (3.3.7), using Y_{Sampled} from Figure 4.13.

Also, the differences $(\Delta_1 P_n - \Delta_1 \widehat{P}_n^{(M)})$ are displayed in Figure 4.15 as well as the sample autocorrelation functions in Figure 4.16. As before, the sample autocorrelations

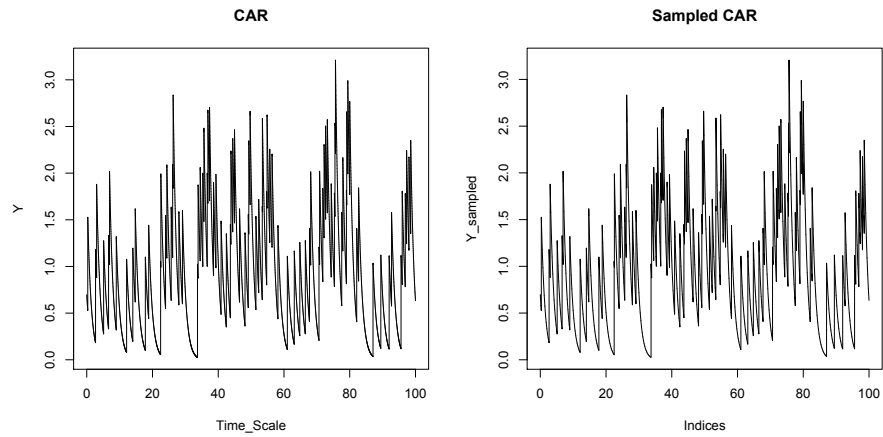


Figure 4.13: Simulated gamma-driven CAR(1) process for $\{N = 100, K = 5000, \lambda = 1, \sigma = 1, a = 1\}$ and Y_{Sampled} for $M = 500$

support our theoretical results.

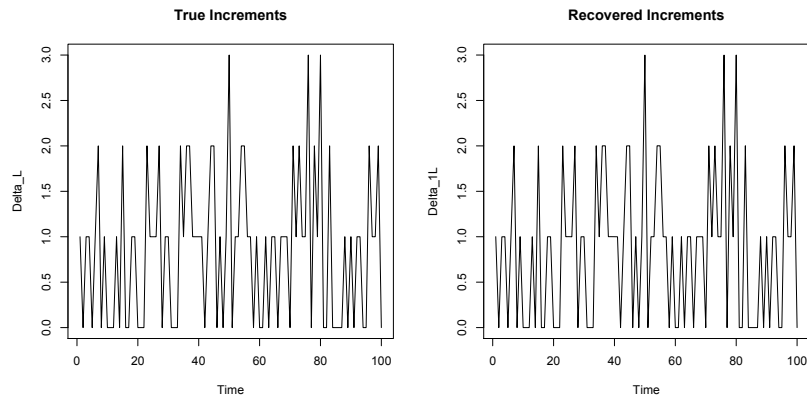


Figure 4.14: True Increments: $\Delta_1 P_n$, Recovered Increments: $\Delta_1 \hat{P}_n^{(M)}$ ($M = 500$)

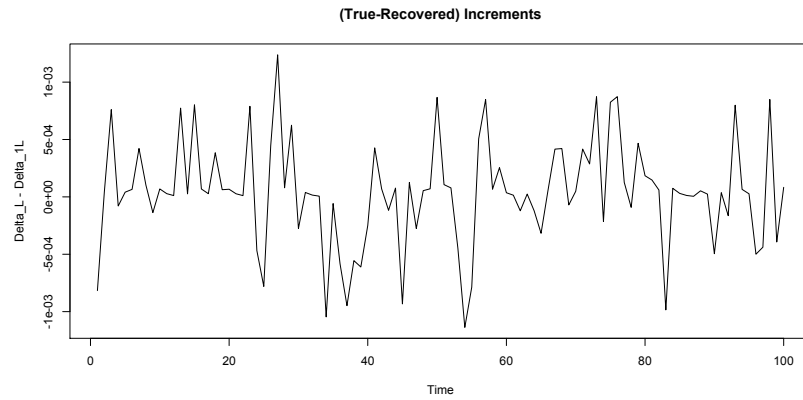


Figure 4.15: (True-Recovered) Increments: $\Delta_1 P_n - \Delta_1 \hat{P}_n^{(M)}$ ($M = 500$)

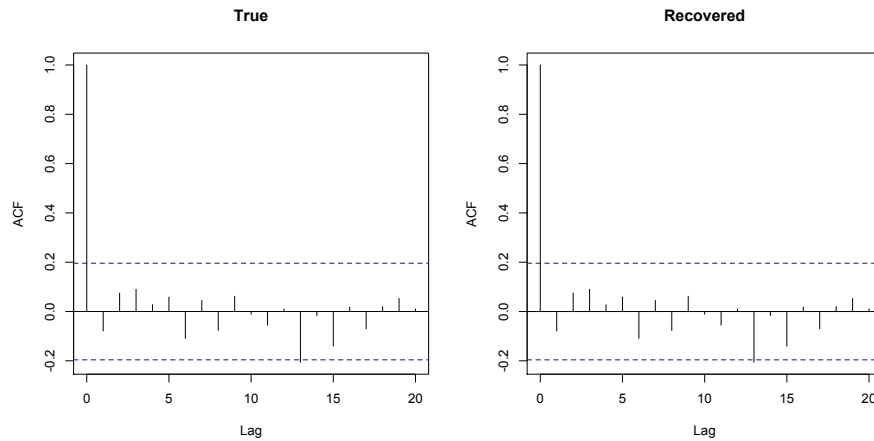


Figure 4.16: The sample autocorrelation function for the True Increments $\Delta_1 P_n$ (True) and for the Recovered Increments $\Delta_1 \hat{P}_n^{(M)}$ (Recovered).

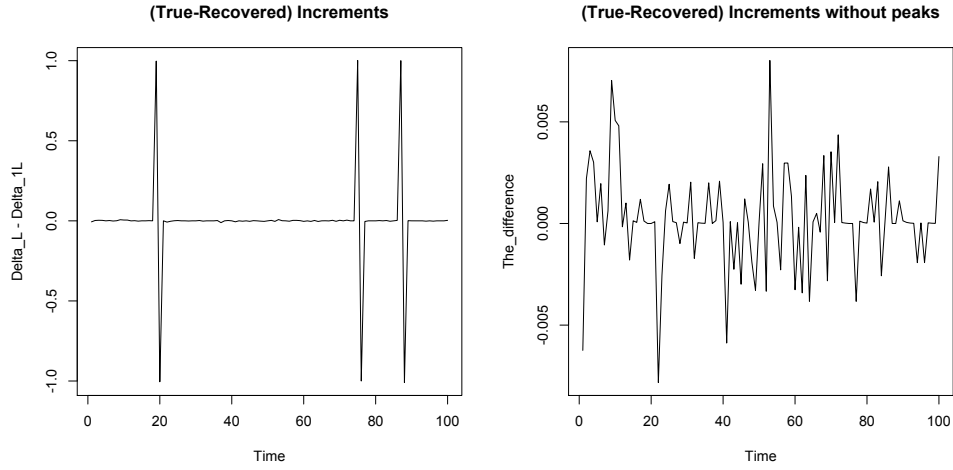


Figure 4.17: (True-Recovered) Increments with and without peaks for small $M = 100$

Tables 4.5 and 4.6 give computed empirical levels $\hat{\alpha}_{\Delta_1 P}$, $\hat{\alpha}_{\Delta_1 \hat{P}(M)}$ for tests based on $W_{\Delta_1 P_n}(1)$ and $W_{\Delta_1 \hat{P}(M)}(1)$ respectively, following the same procedures as before, over $R = 400$ simulations, with nominal level 0.05. We consider various values of the parameters a and σ as well.

	$N = 50, M = 100$		$N = 100, M = 100$		$N = 100, M = 300$		$N = 100, M = 500$	
	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}(M)}$	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}(M)}$	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}(M)}$	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}(M)}$
$a = 0.9$	0.0450	0.0450	0.0450	0.0475	0.0525	0.0525	0.0475	0.0525
$a = 10$	0.0475	0.0575	0.0375	0.0350	0.0575	0.0475	0.0525	0.0550
$a = 100$	0.0575	0.0600	0.0500	0.0325	0.0325	0.0300	0.0550	0.0575
$a = 1000$	0.0375	0.0350	0.0550	0.0350	0.0275	0.0300	0.0550	0.0425

Table 4.5: We fix $\{\sigma = 1, \mu = 1, \eta = 1, K = 5000, R = 400\}$

As in the case of Brownian motion, the performance of the test statistic does not seem to be particularly sensitive to the sampling frequency M or the value of the

	$N = 50, M = 100$		$N = 100, M = 100$		$N = 100, M = 300$		$N = 100, M = 500$	
	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}^{(M)}}$	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}^{(M)}}$	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}^{(M)}}$	$\hat{\alpha}_{\Delta_1 P}$	$\hat{\alpha}_{\Delta_1 \hat{P}^{(M)}}$
$\sigma = 1$	0.0450	0.0450	0.0450	0.0475	0.0525	0.0525	0.0475	0.0525
$\sigma = 10$	0.0500	0.0450	0.0475	0.0500	0.0325	0.0300	0.0375	0.0400
$\sigma = 100$	0.0400	0.0400	0.0325	0.0300	0.0425	0.0450	0.0325	0.0325
$\sigma = 1000$	0.0375	0.0400	0.0525	0.0475	0.0400	0.0500	0.0475	0.0475

Table 4.6: We fix $\{a = 0.9, \mu = 1, \eta = 1, K = 5000, R = 400\}$

ratio N/M . Also, a sample $N = 50$ seems adequate.

4.2.4 An alternative case

To illustrate the behavior of the test statistics $W_{\Delta_1 L}(1)$ and $W_{\Delta_1 \hat{L}^{(M)}}(1)$ under an alternative case, we simulate fractional Brownian motion $B_H(t)$ with Hurst parameter H (see Section 2.3) by using the techniques developed in [18].

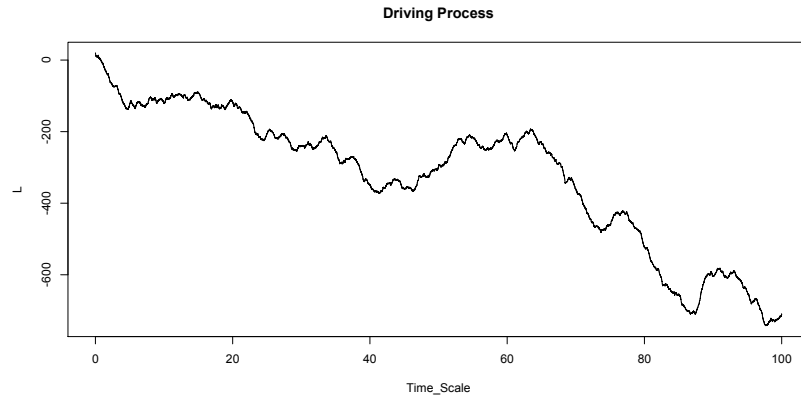


Figure 4.18: Simulated fractional Brownian motion for $N = 100$, $K = 5000$ and $H = 0.8$.

Following the same procedures as before we approximate the process Y_t (B_H -CAR(1)) Y_t and Y_{Sampled} in Figure 4.19, with $H = 0.8$. In Figure 4.20 we consider the same

Y_t and Y_{Sampled} from Figure 4.19 and we computed $\Delta_1 B_{H,n} = B_{H,n} - B_{H,n-1}$ and $\Delta_1 \widehat{B}_{H,n}^{(M)}$ by equation (3.3.7). Also, the differences $(\Delta_1 B_{H,n} - \Delta_1 \widehat{B}_{H,n}^{(M)})$ are illustrated in Figure 4.21.

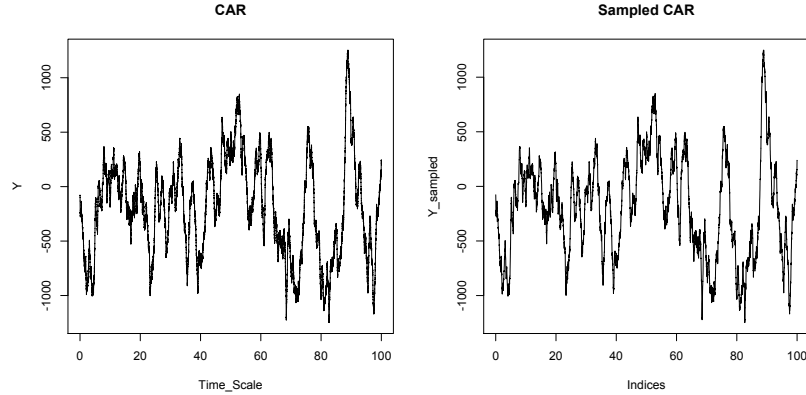


Figure 4.19: Simulated fractional Brownian motion driven CAR(1) process for $N = 100$, $K = 5000$ and Y_{Sampled} for $M = 500$

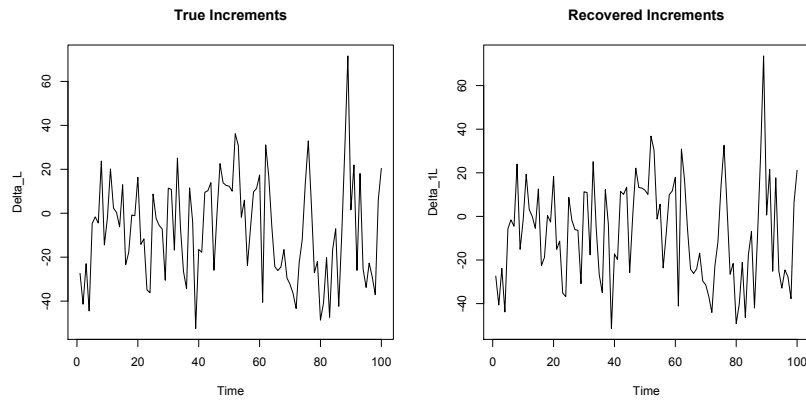


Figure 4.20: True Increments: $\Delta_1 B_{H,n}$, Recovered Increments: $\Delta_1 \widehat{B}_{H,n}^{(M)}$

The sample autocorrelation functions are displayed in Figure 4.22 where we can see that the positive correlation at lag 1 is reflected by both the recovered increments

$(\Delta_1 \widehat{B}_{H,n}^{(M)})$ and the true increments $(\Delta_1 B_{H,n})$.

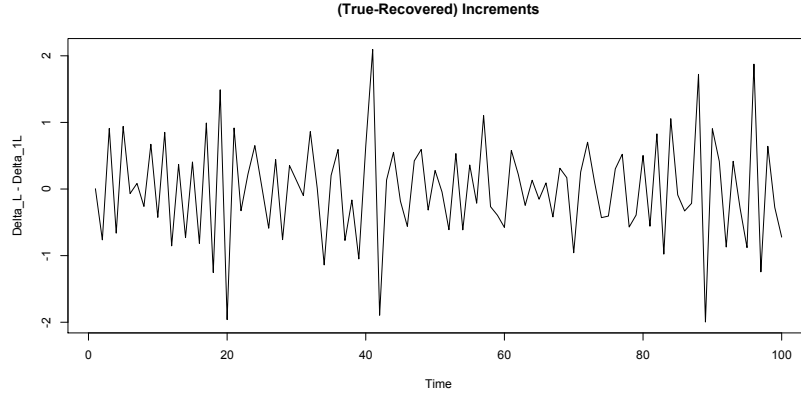


Figure 4.21: (True-Recovered) Increments: $\Delta_1 B_{H,n} - \Delta_1 \widehat{B}_{H,n}^{(M)}$

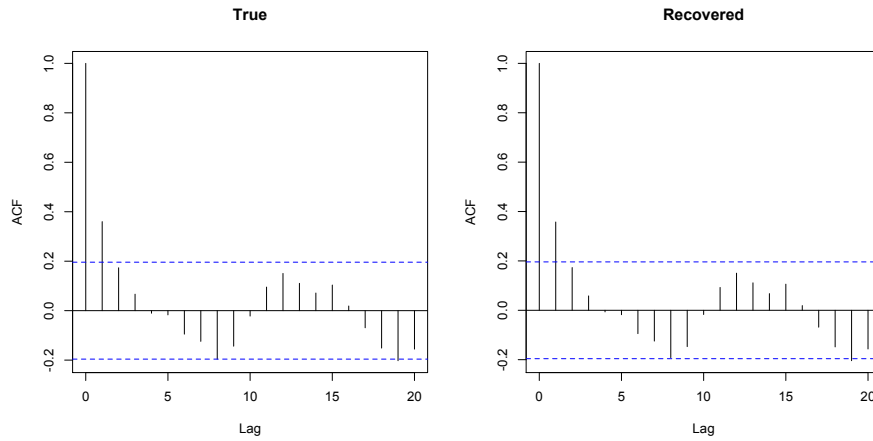


Figure 4.22: The sample autocorrelation function for the True Increments $\Delta_1 B_{H,n}$ (True) and for the Recovered Increment $\Delta_1 \widehat{B}_{H,n}^{(M)}$ (Recovered).

To graph the power functions for both tests at level 0.05 based on $W_{\Delta_1 B_H}(1)$ and $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$, we computed the empirical rejection rate $\widehat{\beta}_{\Delta_1 B_H}, \widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$ for different values of the Hurst parameter H . The power functions are illustrated in Figure 4.23.

H	$\widehat{\beta}_{\Delta_1 B_H}$	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	H	$\widehat{\beta}_{\Delta_1 B_H}$	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$
0.05	0.92	0.96	0.55	0.04	0.04
0.10	0.92	0.90	0.60	0.18	0.16
0.15	0.74	0.78	0.65	0.18	0.16
0.20	0.66	0.68	0.70	0.32	0.32
0.25	0.58	0.52	0.75	0.70	0.70
0.30	0.36	0.44	0.80	0.60	0.60
0.35	0.24	0.30	0.85	0.80	0.80
0.40	0.24	0.26	0.90	0.90	0.90
0.45	0.04	0.04	0.95	0.98	0.98

Table 4.7: $N = 50, K = 1000, R = 50, a = 0.9, \sigma = 1$

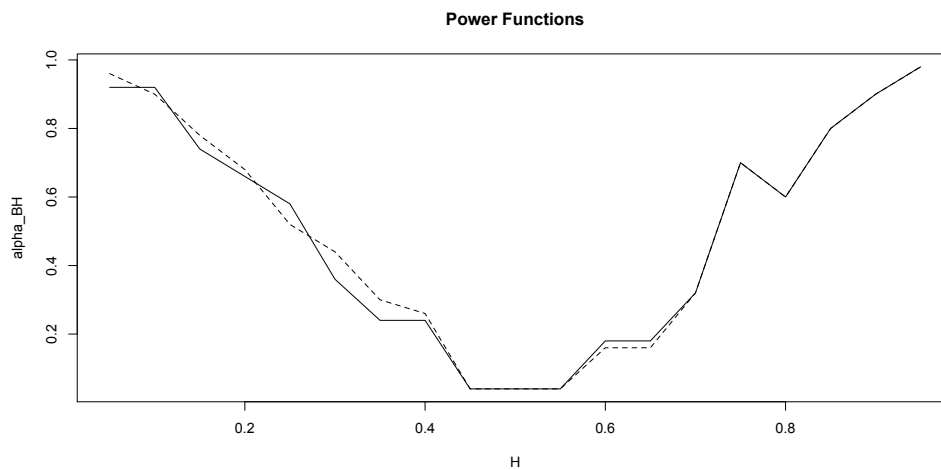


Figure 4.23: The power function for the test based on $W_{\Delta_1 B_H}(1)$ is the solid line and the one based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the dashed line .

Chapter 5

Asymptotics for the sampled process with estimated parameter

In this chapter we study the consistency of the test statistic $W_{\Delta_1 \hat{L}^{(M)}}(1)$ defined in Corollary 4.1.1 if we replace the parameter a by an estimator.

Recall equation (2.2.2):

$$\mathbb{E}[L(t)] = \mu t \quad \text{and} \quad \text{Var}(L(t)) = \eta^2 t, \quad t \geq 0, \quad (5.0.1)$$

and the Definition 2.4.1 of Y , a Lévy-driven CAR(1) process:

$$Y(t) = e^{-at}Y(0) + \sigma \int_0^t e^{-a(t-u)} dL(u), \quad t \geq 0,$$

where $a \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$.

We notice that σ and η are not identifiable, but without loss of generality we can assume that $\sigma = 1$ since if Y is a CAR(1) process driven by L with $\sigma \neq 1$ then Y_t can be also seen as CAR(1) processes driven by the Lévy process $L'(t) = \sigma L(t)$ with $\sigma' = 1$. This discussion is concluded in the following remark.

Remark 5.0.1. Because we test the hypothesis of 0 correlation of the increments, the assumption that $\sigma = 1$ will not affect the test statistic $W_{\Delta_1 \hat{L}^{(M)}}(k)$.

Hence, the only parameter we need to estimate is a .

In what follows, in Section 5.1 we discuss different estimators of a (cf.[44]). We prove the asymptotic normality of the chosen estimator in Theorem 5.1.1.

In Section 5.2 we state the main result, Theorem 5.2.1, that deals with asymptotic normality of the test statistic for zero covariance. This result shows an interesting effect of replacing a with its estimator. Theorem 5.2.1 should be compared to Corollary 4.1.1. The proof is lengthy and technical.

The final section deals with simulation studies. We illustrate Theorem 5.2.1 by showing different power of the test statistics when we use the true value of a and its estimated value. We observe the difference for moderate values of a , as suggested by the Theorem.

All the results in this chapter are new.

5.1 Estimation of a

Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L .

Consider first the classical least squares estimator for a :

$$\tilde{a}_N^{(M)} \equiv \frac{\sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) Y_{\frac{n-1}{M}}}{\frac{1}{M} \sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} \right)^2}.$$

If the driving process L is a Brownian motion, the weak and strong consistency of the estimator $\tilde{a}_N^{(M)}$ is studied in [32], [21], and [29]; the asymptotic normality and asymptotic efficiency of $\tilde{a}_N^{(M)}$ was proven in [37]. Recently Zhang et al. [44] have shown the strong consistency and asymptotic normality for $\tilde{a}_N^{(M)}$ if the driving process

L is a zero mean ($\mu = 0$), second order Lévy process. In this case, it is proven in [44] that

- $\tilde{a}_N^{(M)} \xrightarrow{a.s.} a$ as $N, M \rightarrow \infty$,
- $\sqrt{N} \left(\tilde{a}_N^{(M)} - a \right) \xrightarrow{d} N(0, 2a)$ as $N \rightarrow \infty$ and $N/M^2 \rightarrow 0$.

To perform our test with an estimated value of a in the test statistic $W_{\Delta_1 \hat{L}^{(M)}}(k)$, we need to consider an estimator for a for the general second order driving process L (i.e. $\mu \neq 0$).

The above discussion motivates us to introduce the following estimator of a :

$$\hat{a}_N^{(M)} = \frac{\sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) \left(Y_{\frac{n-1}{M}} - \bar{Y} \right)}{\frac{1}{M} \sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - \bar{Y} \right)^2} \quad (5.1.1)$$

where

$$\bar{Y} = \frac{1}{NM} \sum_{n=1}^{NM} Y_{\frac{n}{M}}.$$

Theorem 5.1.1. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (5.0.1) holds. If $\hat{a}_N^{(M)}$ defined as in (5.1.1), then,*

- (i) $\hat{a}_N^{(M)} \xrightarrow{p} a$ as $N, M \rightarrow \infty$
- (ii) $\sqrt{N} \left(\hat{a}_N^{(M)} - a \right) \xrightarrow{d} N(0, 2a)$ as $N \rightarrow \infty$ and $N/M^2 \rightarrow 0$.

In order to prove Theorem 5.1.1, we start with the following lemma.

Lemma 5.1.2. *Let Y be a CAR(1) process driven by a second-order Lévy process L with parameters a, σ such that (5.0.1) holds. Then the process $Y - \frac{\mu\sigma}{a}$ is a CAR(1) process driven by the centered second-order Lévy process $L(t) - \mu t$ with the same parameters a, σ .*

Proof :

Recall equation (2.4.2):

$$Y(t) = e^{-at}Y(0) + \sigma L(t) - a\sigma \int_0^t e^{-a(t-u)}L(u)du, \quad t \geq 0.$$

Then

$$\begin{aligned} Y(t) - \frac{\mu\sigma}{a} &= \\ &= e^{-at} \left(Y(0) - \frac{\mu\sigma}{a} \right) + e^{-at} \frac{\mu\sigma}{a} - \frac{\mu\sigma}{a} + \sigma L(t) - a\sigma \int_0^t e^{-a(t-u)}L(u)du \\ &= e^{-at} \left(Y(0) - \frac{\mu\sigma}{a} \right) + e^{-at} \frac{\mu\sigma}{a} - \frac{\mu\sigma}{a} + \sigma L(t) - a\sigma \int_0^t e^{-a(t-u)} (L(u) - \mu u) du \\ &\quad - a\mu\sigma e^{-at} \int_0^t u e^{au} du \\ &= e^{-at} \left(Y(0) - \frac{\mu\sigma}{a} \right) + e^{-at} \frac{\mu\sigma}{a} - \frac{\mu\sigma}{a} + \sigma L(t) - a\sigma \int_0^t e^{-a(t-u)} (L(u) - \mu u) du \\ &\quad - a\mu\sigma e^{-at} \left(\frac{te^{at}}{a} - \frac{1}{a^2} (e^{at} - 1) \right) \\ &= e^{-at} \left(Y(0) - \frac{\mu\sigma}{a} \right) + \sigma (L(t) - \mu t) - a\sigma \int_0^t e^{-a(t-u)} (L(u) - \mu u) du \end{aligned}$$

□

Proof of Theorem 5.1.1:

Assuming first that the driving process L is a centered second order Lévy process

($\mu = 0$), we rewrite $\widehat{a}_N^{(M)}$ as

$$\begin{aligned} \widehat{a}_N^{(M)} &= \frac{\sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) Y_{\frac{n-1}{M}} + \bar{Y} \sum_{n=1}^{NM} Y_{\frac{n-1}{M}} - \bar{Y} \sum_{n=1}^{NM} Y_{\frac{n}{M}}}{\frac{1}{M} \sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - \bar{Y} \right)^2} \\ &= \frac{\frac{1}{NM} \left(\sum_{n=1}^{NM} M \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) Y_{\frac{n-1}{M}} + M\bar{Y} \sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) \right)}{\frac{1}{NM} \left(\sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} \right)^2 - 2\bar{Y} \sum_{n=1}^{NM} Y_{\frac{n-1}{M}} + NM (\bar{Y})^2 \right)} \\ &= \frac{\frac{1}{NM} \left(\sum_{n=1}^{NM} M \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) Y_{\frac{n-1}{M}} + M\bar{Y} (Y_0 - Y_N) \right)}{\frac{1}{NM} \left(\sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} \right)^2 - 2\bar{Y} \sum_{n=1}^{NM} Y_{\frac{n-1}{M}} + NM (\bar{Y})^2 \right)} \end{aligned}$$

$$= \frac{\frac{1}{NM} \left(\sum_{n=1}^{NM} M \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) Y_{\frac{n-1}{M}} \right) + \frac{1}{N} \bar{Y} (Y_0 - Y_N)}{\frac{1}{NM} \sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} \right)^2 - 2\bar{Y} \frac{1}{NM} \sum_{n=1}^{NM} Y_{\frac{n-1}{M}} + (\bar{Y})^2}. \quad (5.1.2)$$

Now, $\bar{Y} \xrightarrow{p} 0$ as $N \wedge M \rightarrow \infty$ by Corollary 3.4.1 (i). Therefore, $\frac{\sqrt{N}}{N} \bar{Y} (Y_0 - Y_N) \xrightarrow{p} 0$ as $N \rightarrow \infty$ and $-2\bar{Y} \frac{1}{NM} \sum_{n=1}^{NM} Y_{\frac{n-1}{M}} + (\bar{Y})^2 \xrightarrow{p} 0$ as $N \wedge M \rightarrow \infty$.

This shows that when the driving process L is a centered second order Lévy process, then $\sqrt{N} \left(\hat{a}_N^{(M)} - a \right)$ and $\sqrt{N} \left(\tilde{a}_N^{(M)} - a \right)$ have the same asymptotic behaviour. Hence (i) and (ii) hold for $\hat{a}_N^{(M)}$ as well.

If L is a general Lévy process, then the result follows in light of Lemma 5.1.2 because the estimator $\hat{a}_N^{(M)}$ defined for Y is the same as the estimator $\tilde{a}_N^{(M)}$ defined for $Y - \frac{\mu\sigma}{a}$. Namely,

$$\frac{\sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - Y_{\frac{n}{M}} \right) \left(Y_{\frac{n-1}{M}} - \bar{Y} \right)}{\frac{1}{M} \sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - \bar{Y} \right)^2} = \frac{\sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - \frac{\mu\sigma}{a} - \left(Y_{\frac{n}{M}} - \frac{\mu\sigma}{a} \right) \right) \left(Y_{\frac{n-1}{M}} - \frac{\mu\sigma}{a} - (\bar{Y} - \frac{\mu\sigma}{a}) \right)}{\frac{1}{M} \sum_{n=1}^{NM} \left(Y_{\frac{n-1}{M}} - \frac{\mu\sigma}{a} - (\bar{Y} - \frac{\mu\sigma}{a}) \right)^2}.$$

□

5.2 Test statistics: asymptotic normality

In order to proceed, in analogy to equation (3.3.7), we define the recovered increments using an estimator of a :

$$\begin{aligned} \widehat{\Delta_1 \hat{L}_n^{(M)}} &\equiv \frac{\hat{a}_N^{(M)}}{M} \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}} + \left(1 - \frac{\hat{a}_N^{(M)}}{2M} \right) (Y_n - Y_{n-1}) \\ &= \hat{a}_N^{(M)} \frac{S_n^{(M)}}{M} + \left(1 - \frac{\hat{a}_N^{(M)}}{2M} \right) (Y_n - Y_{n-1}). \end{aligned} \quad (5.2.1)$$

Also, we define $W_{\widehat{\Delta_1 \hat{L}_n^{(M)}}}(1)$ to be the test statistic defined in Corollary 4.1.1 if we replace the parameter a by $\hat{a}_N^{(M)}$ and we define $\widehat{\eta}^2$ to be $\hat{\eta}^2$ defined in Corollary 3.3.10

if we replace a by $\widehat{a}_N^{(M)}$.

Let Y_t be a discretely sampled stochastic process. If Y is a CAR(1) model driven by a process L , we can use the estimated increments to test H_0 that L has uncorrelated increments, which will be true if L is a Lévy process. We reject H_0 for a large absolute value of the statistic $W_{\Delta_1 \widehat{L}^{(M)}}(1)$ which behaves as follows:

Theorem 5.2.1. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (5.0.1) holds. Consider Case (III), then*

$$W_{\Delta_1 \widehat{L}^{(M)}}(1) = \sqrt{N} \frac{\widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(1)}{\widehat{\eta}^2} + \epsilon_N^{(M)}(a) + o_p(1),$$

where

- as $N \rightarrow \infty$, $N/M \rightarrow 0$,

$$\sqrt{N} \frac{\widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(1)}{\widehat{\eta}^2} \xrightarrow{d} N(0, 1),$$

- as $N \rightarrow \infty$, $N/M \rightarrow 0$ and $a \rightarrow \infty$, $\epsilon_N^{(M)}(a) \xrightarrow{p} 0$,
- $o_p(1) \xrightarrow{p} 0$ as $N \rightarrow \infty$, $N/M \rightarrow 0$.

Under H_0 , for large N, M, a and N/M small we have

$$\alpha \approx P\left(|W_{\Delta_1 \widehat{L}^{(M)}}(1)| > z_{\alpha/2}\right).$$

where $z_{\alpha/2}$ is the critical Z -value where $Z \sim N(0, 1)$.

The proof of Theorem 5.2.1 is lengthy and requires several preliminary results.

5.2.1 Preliminary lemmas

Now we will present a sequence of lemmas that will help to prove Theorem 5.2.1. Because of the non-identifiability between σ and η (see Remark 5.0.1), without loss of generality we will assume $\sigma = 1$ in what follows.

Lemma 5.2.2. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (5.0.1) holds. Then*

- (i) $\frac{1}{N} \sum_{n=1}^N a \int_{n-1}^n Y(s) ds (Y_{n+1} - Y_n) \xrightarrow{a.s.} -\frac{\eta^2}{2a} (e^{-a} - 1)^2$ as $N \rightarrow \infty$.
- (ii) $\frac{1}{N} \sum_{n=1}^N a \int_n^{n+1} Y(s) ds (Y_n - Y_{n-1}) \xrightarrow{a.s.} \frac{\eta^2}{2a} (e^{-a} - 1)^2$ as $N \rightarrow \infty$.
- (iii) $\frac{1}{N} \sum_{n=1}^N a \int_{n-1}^n Y(s) ds (Y_n - Y_{n-1}) \xrightarrow{a.s.} 0$ as $N \rightarrow \infty$.
- (iv) $\frac{1}{N} \sum_{n=1}^N a^2 \int_n^{n+1} Y(s) ds \int_{n-1}^n Y(s) ds \xrightarrow{a.s.} \frac{\eta^2}{2a} (e^{-a} - 1)^2 + \mu^2$ as $N \rightarrow \infty$.

Proof : Recall the notation $\Delta_1 L_n = L_n - L_{n-1}$. From equation (3.3.6) we have:

$$\Delta_1 L_n = Y(n) - Y(n-1) + a \int_{n-1}^n Y(s) ds.$$

(i) By ergodicity we have,

$$\frac{1}{N} \sum_{n=1}^N \Delta_1 L_n (Y_{n+1} - Y_n) \xrightarrow{a.s.} \mathbb{E} [\Delta_1 L_n (Y_{n+1} - Y_n)] = -\frac{\eta^2}{a} (e^{-a} - 1)^2,$$

since

$$\begin{aligned} & \mathbb{E} [\Delta_1 L_n (Y_{n+1} - Y_n)] \\ &= \mathbb{E} \left[\left(Y_n - Y_{n-1} + a \int_{n-1}^n Y(s) ds \right) (Y_{n+1} - Y_n) \right] \\ &= \mathbb{E} [(Y_n - Y_{n-1}) (Y_{n+1} - Y_n)] + a \mathbb{E} \left[\int_{n-1}^n Y(s) (Y_{n+1} - Y_n) ds \right] \\ &= \mathbb{E} [(Y_1 - Y_0) (Y_2 - Y_1)] + a \mathbb{E} \left[\int_0^1 Y(s) (Y_2 - Y_1) ds \right] \\ &= \mathbb{E} [Y_1 Y_2 - Y_1 Y_1 - Y_0 Y_2 + Y_0 Y_1] + a \int_0^1 \mathbb{E} [Y(s) Y_2] ds - a \int_0^1 \mathbb{E} [Y(s) Y_1] ds \\ &= \frac{\eta^2}{2a} \left(2e^{-a} - 1 - e^{-2a} + ae^{-2a} \int_0^1 e^{as} ds - ae^{-a} \int_0^1 e^{as} ds \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{\eta^2}{2a} (2e^{-a} - 1 - e^{-2a} + e^{-2a}(e^a - 1) - e^{-a}(e^a - 1)) \\
&= \frac{\eta^2}{2a} (4e^{-a} - 2 - 2e^{-2a}) = \frac{\eta^2}{2a} (-2(e^{-a} - 1)^2).
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{1}{N} \sum_{n=1}^N \Delta_1 L_n (Y_{n+1} - Y_n) &= \frac{1}{N} \sum_{n=1}^N \left(Y_n - Y_{n-1} + a \int_{n-1}^n Y(s) ds \right) (Y_{n+1} - Y_n) \\
&= \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1}) (Y_{n+1} - Y_n) + \frac{1}{N} \sum_{n=1}^N \left(a \int_{n-1}^n Y(s) ds \right) (Y_{n+1} - Y_n).
\end{aligned}$$

By ergodicity we have,

$$\begin{aligned}
\frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1}) (Y_{n+1} - Y_n) &\xrightarrow{a.s.} \mathbb{E}[(Y_n - Y_{n-1}) (Y_{n+1} - Y_n)] \\
&= \frac{\sigma^2 \eta^2}{2a} (\mathbb{E}Y_2 Y_1 - \mathbb{E}Y_2 Y_0 - \mathbb{E}Y_0^2 + \mathbb{E}Y_1 Y_0) \\
&= \frac{\eta^2}{2a} (2e^{-a} - e^{-2a} - 1) = -\frac{\eta^2}{2a} (e^{-a} - 1)^2,
\end{aligned} \tag{5.2.2}$$

and the result follows.

(ii): Similarly to (i),

$$\begin{aligned}
&\frac{1}{N} \sum_{n=1}^N \Delta_1 L_{n+1} (Y_n - Y_{n-1}) \\
&= \frac{1}{N} \sum_{n=1}^N \left(Y_{n+1} - Y_n + a \int_n^{n+1} Y(s) ds \right) (Y_n - Y_{n-1}) \\
&= \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1}) (Y_{n+1} - Y_n) + \frac{1}{N} \sum_{n=1}^N \left(a \int_n^{n+1} Y(s) ds \right) (Y_n - Y_{n-1}).
\end{aligned}$$

By ergodicity we have,

$$\frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1}) (Y_{n+1} - Y_n) \xrightarrow{a.s.} -\frac{\eta^2}{2a} (e^{-a} - 1)^2.$$

Also by ergodicity and independence

$$\frac{1}{N} \sum_{n=1}^N \Delta_1 L_{n+1}(Y_n - Y_{n-1}) \xrightarrow{a.s.} \mathbb{E}[\Delta_1 L_{n+1}(Y_n - Y_{n-1})] = 0$$

The result follows.

(iii) : By ergodicity we have,

$$\frac{1}{N} \sum_{n=1}^N \Delta_1 L_n(Y_{n+1} - Y_n) \xrightarrow{a.s.} \mathbb{E}[\Delta L_n(Y_{n+1} - Y_n)] = \frac{\eta^2}{a} (1 - e^{-a}),$$

since

$$\begin{aligned} & \mathbb{E}[\Delta_1 L_n(Y_n - Y_{n-1})] \\ &= \mathbb{E}\left[\left(Y_n - Y_{n-1} + a \int_{n-1}^n Y(s) ds\right) (Y_n - Y_{n-1})\right] \\ &= \mathbb{E}[(Y_n - Y_{n-1})^2] + a \mathbb{E}\left[\int_{n-1}^n Y(s) (Y_n - Y_{n-1}) ds\right] \\ &= \mathbb{E}[(Y_1 - Y_0)^2] + a \mathbb{E}\left[\int_0^1 Y(s) (Y_1 - Y_0) ds\right] \\ &= \text{Var}(Y_1 - Y_0) + a \int_0^1 \mathbb{E}[Y(s)Y_1] ds - a \int_0^1 \mathbb{E}[Y(s)Y_0] ds \\ &= \frac{\eta^2}{2a} \left(2 - 2e^{-a} + ae^a \int_0^1 e^{-as} ds - a \int_0^1 e^{as} ds\right) \\ &= \frac{\eta^2}{2a} (2 - 2e^{-a} - e^a (e^{-a} - 1) - (e^a - 1)) \\ &= \frac{\eta^2}{2a} (2 - 2e^{-a} - 1 + e^a - e^a + 1) = \frac{\eta^2}{a} (1 - e^{-a}). \end{aligned}$$

Now,

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \Delta_1 L_n(Y_n - Y_{n-1}) &= \frac{1}{N} \sum_{n=1}^N \left(Y_n - Y_{n-1} + a \int_{n-1}^n Y(s) ds\right) (Y_n - Y_{n-1}) \\ &= \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 + \frac{1}{N} \sum_{n=1}^N \left(a \int_{n-1}^n Y(s) ds\right) (Y_n - Y_{n-1}). \end{aligned}$$

By ergodicity we have,

$$\frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 \xrightarrow{a.s.} \text{Var}(Y_n - Y_{n-1}) = \frac{\eta^2}{a} (1 - e^{-a}), \quad (5.2.3)$$

and the result follows.

(iv) : The result follows directly from (i), (ii), and the equation:

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \Delta_1 L_{n+1} \Delta_1 L_n = & \\ & \frac{1}{N} \sum_{n=1}^N a^2 \int_n^{n+1} Y(s) ds \int_{n-1}^n Y(s) ds + \frac{1}{N} \sum_{n=1}^N a \int_{n-1}^n Y(s) ds (Y_{n+1} - Y_n) \\ & + \frac{1}{N} \sum_{n=1}^N a \int_n^{n+1} Y(s) ds (Y_n - Y_{n-1}) + \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1}) (Y_{n+1} - Y_n) \end{aligned}$$

since

$$\frac{1}{N} \sum_{n=1}^N \Delta L_{n+1} \Delta L_n \xrightarrow{a.s.} \mu^2 \text{ as } N \rightarrow \infty.$$

□

Lemma 5.2.3. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (5.0.1) holds. For $n \geq 1$ define*

$$S_n^{(M)} \equiv \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}}. \quad (5.2.4)$$

Then

(i) as $N \wedge M \rightarrow \infty$,

$$\frac{1}{N} \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \xrightarrow{p} 0,$$

(ii) as $N \rightarrow \infty$ and $N/M \rightarrow 0$,

$$\frac{1}{\sqrt{N}} \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \xrightarrow{d} N \left(0, \frac{\eta^2}{a^2} \right).$$

Proof: Using the definition of $S_n^{(M)}$, the result follows from Corollary 3.4.1 (i) and (ii). □

Corollary 5.2.4. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (5.0.1) holds. If $S_n^{(M)}$ as defined in (5.2.4) then,*

- (i) $\frac{1}{N} \sum_{n=1}^N a \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) \xrightarrow{p} -\frac{\eta^2}{2a} (e^{-a} - 1)^2$ as $N \wedge M \rightarrow \infty$.
- (ii) $\frac{1}{N} \sum_{n=1}^N a \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) \xrightarrow{p} \frac{\eta^2}{2a} (e^{-a} - 1)^2$ as $N \wedge M \rightarrow \infty$.
- (iii) $\frac{1}{N} \sum_{n=1}^N a \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) \xrightarrow{p} 0$ as $N \wedge M \rightarrow \infty$.
- (iv) $\frac{1}{N} \sum_{n=1}^N a^2 \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \xrightarrow{p} \frac{\eta^2}{2a} (e^{-a} - 1)^2$ as $N \wedge M \rightarrow \infty$.
- (v) $\frac{a^2}{N} \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M} \right)^2 \xrightarrow{p} \frac{\eta^2}{a} (e^{-a} + a - 1) + \mu^2$ as $N \wedge M \rightarrow \infty$.

Proof:

The proofs of (i), (ii), (iii) and (iv) depend on showing the L_1 closeness of the terms in Corollary 5.2.4 to the corresponding ones in Lemma 5.2.2.

(i) :

$$\frac{1}{N} \sum_{n=1}^N a \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) = \frac{1}{N} \sum_{n=1}^N a \frac{S_n^{(M)}}{M} (Y_{n+1} - Y_n) + \frac{\mu}{N} (Y_{N+1} - Y_1).$$

The last term converges to 0 in probability as $N \rightarrow \infty$.

Now we show the L_1 closeness between the term $\frac{1}{N} \sum_{n=1}^N a \frac{S_n^{(M)}}{M} (Y_{n+1} - Y_n)$ and the corresponding one in Lemma 5.2.2(i). We have

$$\begin{aligned} & \frac{1}{N} \mathbb{E} \left| \sum_{n=1}^N a \frac{S_n^{(M)}}{M} (Y_{n+1} - Y_n) - \sum_{n=1}^N a \int_{n-1}^n Y(s) ds (Y_{n+1} - Y_n) \right| \\ & \leq \frac{a}{N} \sum_{n=1}^N \mathbb{E} \left| \frac{S_n^{(M)}}{M} (Y_{n+1} - Y_n) - \int_{n-1}^n Y(s) ds (Y_{n+1} - Y_n) \right| \end{aligned}$$

$$\begin{aligned}
&\leq \frac{a}{N} \sum_{n=1}^N \left\| \frac{S_n^{(M)}}{M} - \int_{n-1}^n Y(s) ds \right\|_{L_2} \| (Y_{n+1} - Y_n) \|_{L_2} \\
&= a \left\| \frac{S_n^{(M)}}{M} - \int_{n-1}^n Y(s) ds \right\|_{L_2} \| (Y_{n+1} - Y_n) \|_2 \\
&\leq \frac{a\eta}{\sqrt{a}} \left(1 - e^{-\frac{a}{M}}\right)^{\frac{1}{2}} \| (Y_1 - Y_0) \|_{L_2} \quad \text{by Lemma 3.3.6} \\
&\rightarrow 0 \quad \text{as } M \rightarrow \infty, \text{ uniformly in } N.
\end{aligned}$$

Hence the result follows by Lemma 5.2.2 (i).

(ii), (iii) : Similar to (i).

(iv) : We have

$$\begin{aligned}
&\frac{1}{N} \sum_{n=1}^N a^2 \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) = \\
&\frac{1}{N} \sum_{n=1}^N a^2 \frac{S_{n+1}^{(M)}}{M} \frac{S_n^{(M)}}{M} - \frac{a\mu}{N} \sum_{n=1}^N \frac{S_{n+1}^{(M)}}{M} - \frac{a\mu}{N} \sum_{n=1}^N \frac{S_n^{(M)}}{M} + \mu^2.
\end{aligned}$$

By Lemma 5.2.3 (i) we have that the terms

$$\frac{a\mu}{N} \sum_{n=1}^N \frac{S_{n+1}^{(M)}}{M} \xrightarrow{p} \mu^2 \quad \text{and} \quad \frac{a\mu}{N} \sum_{n=1}^N \frac{S_n^{(M)}}{M} \xrightarrow{p} \mu^2 \quad \text{as } N \wedge M \rightarrow \infty.$$

Now for the term first term, similar to (i), we have:

$$\begin{aligned}
&\frac{1}{N} \mathbb{E} \left| \sum_{n=1}^N a^2 \frac{S_{n+1}^{(M)}}{M} \frac{S_n^{(M)}}{M} - \sum_{n=1}^N a^2 \int_n^{n+1} Y(s) ds \int_{n-1}^n Y(s) ds \right| \\
&\leq \frac{a^2}{N} \sum_{n=1}^N \mathbb{E} \left| \frac{S_{n+1}^{(M)}}{M} \frac{S_n^{(M)}}{M} - \int_n^{n+1} Y(s) ds \int_{n-1}^n Y(s) ds \right| \\
&\leq a^2 \left(\left\| \frac{S_n^{(M)}}{M} - \int_{n-1}^n Y(s) ds \right\|_{L_2} \left\| \frac{S_{n+1}^{(M)}}{M} \right\|_{L_2} + \left\| \frac{S_{n+1}^{(M)}}{M} - \int_n^{n+1} Y(s) ds \right\|_{L_2} \left\| \int_{n-1}^n Y(s) ds \right\|_{L_2} \right).
\end{aligned}$$

The upper bound converges to 0, uniformly in N , by using the same argument as in

(i). Combining the terms, the result follows.

(v) : Recall the definition and the consistency of $\widehat{\eta}^2$ from Lemma 3.3.10:

$$\begin{aligned}
\widehat{\eta}^2 &= \frac{1}{N} \sum_{n=1}^N \left(\Delta_1 \widehat{L}_n^{(M)} \right)^2 - \left(\overline{\Delta_1 \widehat{L}^{(M)}} \right)^2 \\
&= \frac{1}{N} \sum_{n=1}^N \left(a \frac{S_n^{(M)}}{M} + \left(1 - \frac{a}{2M} \right) (Y_n - Y_{n-1}) \right)^2 - \left(\overline{\Delta_1 \widehat{L}^{(M)}} \right)^2 \\
&= \frac{1}{N} a^2 \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M} \right)^2 + \left(2 - \frac{a}{M} \right) \frac{1}{N} \sum_{n=1}^N a \frac{S_n^{(M)}}{M} (Y_n - Y_{n-1}) \\
&\quad + \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 - \frac{a}{M} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 + \frac{a^2}{4M^2} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 \\
&\quad - \left(\overline{\Delta_1 \widehat{L}^{(M)}} \right)^2. \tag{5.2.5}
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{1}{N} a^2 \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M} \right)^2 &= \widehat{\eta}^2 - \left(2 - \frac{a}{M} \right) \frac{1}{N} \sum_{n=1}^N a \frac{S_n^{(M)}}{M} (Y_n - Y_{n-1}) - \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 \\
&\quad + \frac{a}{M} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 - \frac{a^2}{4M^2} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 + \left(\overline{\Delta_1 \widehat{L}^{(M)}} \right)^2.
\end{aligned}$$

Now by (iii), as $N \wedge M \rightarrow \infty$, we have

$$\frac{1}{N} \sum_{n=1}^N a \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) \xrightarrow{p} 0$$

which implies that $\frac{1}{N} \sum_{n=1}^N a \frac{S_n^{(M)}}{M} (Y_n - Y_{n-1}) \xrightarrow{p} 0$, since $\frac{1}{N} \sum_{n=1}^N a \frac{\mu}{a} (Y_n - Y_{n-1}) \xrightarrow{p} 0$.

Hence by ergodicity and (iii), as $N \wedge M \rightarrow \infty$ we have:

$$\frac{1}{N} a^2 \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M} \right)^2 \xrightarrow{p} \eta^2 - \text{Var}(Y_1 - Y_0) + \mu^2 = \eta^2 - \frac{\eta^2}{a} + \frac{\eta^2}{a} e^{-a} + \mu^2.$$

□

5.2.2 Asymptotics for Sample Mean and Sample Variance

Lemma 5.2.5. Let $\overline{\Delta_1 \widehat{L}^{(M)}}$ and $\widehat{\eta}^2$ be the sample mean and sample variance, respectively, of $\widehat{\Delta_1 \widehat{L}_n^{(M)}}$, i.e.

$$\overline{\Delta_1 \widehat{L}^{(M)}} \equiv \frac{1}{N} \sum_{n=1}^N \widehat{\Delta_1 \widehat{L}_n^{(M)}} \quad \text{and} \quad \widehat{\eta}^2 \equiv \frac{1}{N} \sum_{n=1}^N \left(\widehat{\Delta_1 \widehat{L}_n^{(M)}} - \overline{\Delta_1 \widehat{L}^{(M)}} \right)^2.$$

Then,

- (i) $\overline{\Delta_1 \widehat{L}^{(M)}} \xrightarrow{p} \mu$ as $N \wedge M \rightarrow \infty$.
- (ii) $\sqrt{N} \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \xrightarrow{d} N(0, \eta^2)$ as $N \rightarrow \infty$ and $N/M \rightarrow 0$.
- (iii) $\widehat{\eta}^2 \xrightarrow{p} \eta^2$ as $N \wedge M \rightarrow \infty$.

Proof of (i), (ii) : Using equation (5.2.1),

$$\begin{aligned} \widehat{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu &= \widehat{a}_N^{(M)} \frac{S_n^{(M)}}{M} - \frac{\widehat{a}_N^{(M)}}{a} \mu + \left(1 - \frac{\widehat{a}_N^{(M)}}{2M} \right) (Y_n - Y_{n-1}) \\ &= \widehat{a}_N^{(M)} \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) + \left(1 - \frac{\widehat{a}_N^{(M)}}{2M} \right) (Y_n - Y_{n-1}). \end{aligned}$$

Using equation (3.3.7) we have:

$$\begin{aligned} &\widehat{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu - \left(\Delta_1 \widehat{L}_n^{(M)} - \mu \right) \\ &= \widehat{a}_N^{(M)} \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) + \left(1 - \frac{\widehat{a}_N^{(M)}}{2M} \right) (Y_n - Y_{n-1}) - a \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \\ &\quad - \left(1 - \frac{a}{2M} \right) (Y_n - Y_{n-1}) = \left(\widehat{a}_N^{(M)} - a \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) + \left(\frac{a - \widehat{a}_N^{(M)}}{2M} \right) (Y_n - Y_{n-1}). \end{aligned}$$

Now,

$$\begin{aligned} &\sum_{n=0}^N \left(\widehat{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu - \left(\Delta_1 \widehat{L}_n^{(M)} - \mu \right) \right) = \\ &= \left(\widehat{a}_N^{(M)} - a \right) \sum_{n=0}^N \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) + \left(\frac{a - \widehat{a}_N^{(M)}}{2M} \right) (Y_N - Y_0). \end{aligned}$$

Part (i) follows from Theorem 5.1.1 (i), Lemma 5.2.3 (i) and Corollary 3.3.10 (i).

Part (ii) follows Theorem 5.1.1 (i), Lemma 5.2.3 (ii) and Corollary 3.3.10 (ii).

(iii): By replacing the parameter a by its estimator $\widehat{a}_N^{(M)}$ in equation (5.2.5) we have,

$$\begin{aligned}\widehat{\eta}^2 &= \frac{1}{N} \left(\widehat{a}_N^{(M)}\right)^2 \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M}\right)^2 + \left(2 - \frac{\widehat{a}_N^{(M)}}{M}\right) \frac{1}{N} \sum_{n=1}^N \widehat{a}_N^{(M)} \frac{S_n^{(M)}}{M} (Y_n - Y_{n-1}) \\ &+ \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 - \frac{\widehat{a}_N^{(M)}}{M} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 \\ &+ \frac{\left(\widehat{a}_N^{(M)}\right)^2}{4M^2} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 - \left(\overline{\Delta_1 \widehat{L}^{(M)}}\right)^2.\end{aligned}$$

By (i) the last term converges in probability to μ^2 as $N \wedge M \rightarrow \infty$.

By ergodicity and Slutsky's Theorem we have as $N \wedge M \rightarrow \infty$,

$$\begin{aligned}-\frac{\widehat{a}_N^{(M)}}{M} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 + \frac{\left(\widehat{a}_N^{(M)}\right)^2}{4M^2} \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 &\xrightarrow{p} 0, \\ \frac{1}{N} \sum_{n=1}^N (Y_n - Y_{n-1})^2 &\xrightarrow{p} \text{Var}(Y_n - Y_{n-1}) = \frac{\eta^2}{a} - \frac{\eta^2}{a} e^{-a}.\end{aligned}$$

Using Theorem 5.1.1 (i), Corollary 5.2.4 (iii), Slutsky's Theorem and an argument similar to that used to prove Corollary 5.2.4 (v), we have as $N \wedge M \rightarrow \infty$,

$$\left(2 - \frac{\widehat{a}_N^{(M)}}{M}\right) \frac{1}{N} \sum_{n=1}^N \widehat{a}_N^{(M)} \frac{S_n^{(M)}}{M} (Y_n - Y_{n-1}) \xrightarrow{p} 0.$$

Using Theorem 5.1.1(i), Corollary 5.2.4(v), and Slutsky's Theorem, we have as $N \wedge M \rightarrow \infty$,

$$\frac{1}{N} \left(\widehat{a}_N^{(M)}\right)^2 \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M}\right)^2 = \frac{\left(\widehat{a}_N^{(M)}\right)^2}{a^2} \frac{1}{N} \sum_{n=1}^N \left(\frac{S_n^{(M)}}{M}\right)^2 \xrightarrow{p} \frac{\eta^2}{a} (e^{-a} + a - 1) + \mu^2.$$

The result follows by putting all the terms together:

$$\widehat{\eta}^2 \xrightarrow{p} \frac{\eta^2}{a} (e^{-a} + a - 1) + \mu^2 + \frac{\eta^2}{a} - \frac{\eta^2}{a} e^{-a} - \mu^2 = \eta^2.$$

□

5.2.3 Asymptotics for Sample Covariances

In this section we consider sample covariances of the estimated increments with estimated parameter a at lag 1, defined by

$$\widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(1) \equiv \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\widehat{\Delta_1 L}_{n+1}^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right) \left(\widehat{\Delta_1 L}_n^{(M)} - \overline{\Delta_1 \widehat{L}^{(M)}} \right),$$

where

$$\overline{\Delta_1 \widehat{L}^{(M)}} = \frac{1}{N} \sum_{n=1}^N \Delta \widehat{L}_n^{(M)}.$$

Theorem 5.2.6. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (5.0.1) holds. Then*

$$\sqrt{N} \widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(1) = \sqrt{N} \gamma_{\Delta_1 \widehat{L}^{(M)}}(1) + \epsilon_N^{(M)}(a, \eta) + o_p(1),$$

where as $N \rightarrow \infty$ and $N/M \rightarrow 0$:

- $\sqrt{N} \widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(1) \xrightarrow{d} N(0, \eta^4)$ (by Theorem 3.3.11(ii)),
- $\epsilon_N^{(M)}(a, \eta) \xrightarrow{d} N\left(0, 2 \frac{\eta^4 (e^{-a} - 1)^4}{a^4}\right)$,
- $o_p(1) \xrightarrow{p} 0$.

Proof: We decompose $\sqrt{N} \widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(1)$ as:

$$\begin{aligned} \sqrt{N} \widehat{\gamma}_{\Delta_1 \widehat{L}^{(M)}}(1) &= \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\left(\widehat{\Delta_1 L}_{n+1}^{(M)} - \frac{\widehat{a}_N^{(M)}}{a} \mu - \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \right) \right. \\ &\quad \left. \left(\widehat{\Delta_1 L}_n^{(M)} - \frac{\widehat{a}_N^{(M)}}{a} \mu - \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \right) \right) \\ &= \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\widehat{\Delta_1 L}_{n+1}^{(M)} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\widehat{\Delta_1 L}_n^{(M)} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \\ &\quad - \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\widehat{\Delta_1 L}_{n+1}^{(M)} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \end{aligned}$$

$$\begin{aligned}
& - \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \\
& + \sqrt{N} \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right)^2 = I_1 - I_2 - I_3 + I_4. \tag{5.2.6}
\end{aligned}$$

As $N \rightarrow \infty$ and $N/M \rightarrow 0$, we have by Lemma 5.2.5 (i), (ii) and Slutsky's Theorem the following asymptotics:

- $I_4 = \sqrt{N} \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \xrightarrow{p} 0$,
- $I_3 = \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \xrightarrow{p} 0$,
- $I_2 = \left(\overline{\Delta_1 \widehat{L}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\overline{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \xrightarrow{p} 0$.

Hence,

$$\sqrt{N} \gamma_{\overline{\Delta_1 \widehat{L}^{(M)}}}(1) = \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\overline{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) + o_p(1),$$

where $o_p(1)$ converges in probability to 0 as $N \rightarrow \infty$ and $N/M \rightarrow 0$. Now, from the proof of Theorem 3.3.11 (i) we have that

$$\sqrt{N} \gamma_{\overline{\Delta_1 \widehat{L}^{(M)}}} = \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\overline{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \mu \right) \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \mu \right) + o_p(1),$$

and

$$\frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\overline{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \mu \right) \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \mu \right) \xrightarrow{d} N(0, \eta^4) \text{ as } N \rightarrow \infty \text{ and } \frac{N}{M} \rightarrow 0.$$

Next we consider the following difference:

$$\frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\left(\overline{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) - \left(\overline{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \mu \right) \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \mu \right) \right).$$

We have,

$$\left(\overline{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\overline{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right)$$

$$\begin{aligned}
&= \left(\widehat{a}_N^{(M)} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) + \left(1 - \frac{\widehat{a}_N^{(M)}}{2M} \right) (Y_{n+1} - Y_n) \right) \left(\widehat{a}_N^{(M)} \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \right. \\
&\quad \left. + \left(1 - \frac{\widehat{a}_N^{(M)}}{2M} \right) (Y_n - Y_{n-1}) \right) \\
&= \left(\widehat{a}_N^{(M)} \right)^2 \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) + \widehat{a}_N^{(M)} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) \\
&\quad - \frac{\left(\widehat{a}_N^{(M)} \right)^2}{2M} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) \\
&\quad + \widehat{a}_N^{(M)} \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) - \frac{\left(\widehat{a}_N^{(M)} \right)^2}{2M} \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) \\
&\quad + \left(1 - \frac{\widehat{a}_N^{(M)}}{2M} \right)^2 (Y_{n+1} - Y_n) (Y_n - Y_{n-1}).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\left(\widehat{\Delta}_1 \widehat{L}_{n+1}^{(M)} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\widehat{\Delta}_1 \widehat{L}_n^{(M)} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) - \left(\Delta_1 \widehat{L}_{n+1}^{(M)} - \mu \right) \left(\Delta_1 \widehat{L}_n^{(M)} - \mu \right) \right) \\
&= \sqrt{N} \left(\left(\widehat{a}_N^{(M)} \right)^2 - a^2 \right) \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \\
&\quad + \sqrt{N} \left(\widehat{a}_N^{(M)} - a \right) \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) \\
&\quad + \sqrt{N} \left(\widehat{a}_N^{(M)} - a \right) \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) \\
&\quad + \sqrt{N} \frac{a^2 - \left(\widehat{a}_N^{(M)} \right)^2}{2M} \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) \\
&\quad + \sqrt{N} \frac{a^2 - \left(\widehat{a}_N^{(M)} \right)^2}{2M} \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) \\
&\quad + \sqrt{N} \frac{a - \widehat{a}_N^{(M)}}{M} \frac{1}{N-1} \sum_{n=1}^{N-1} (Y_{n+1} - Y_n) (Y_n - Y_{n-1}) \\
&\quad + \sqrt{N} \frac{\left(\widehat{a}_N^{(M)} \right)^2 - a^2}{4M^2} \frac{1}{N-1} \sum_{n=1}^{N-1} (Y_{n+1} - Y_n) (Y_n - Y_{n-1}).
\end{aligned}$$

Simplifying,

$$\begin{aligned}
& \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\left(\widehat{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\widehat{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) - \left(\Delta_1 \widehat{L}_{n+1}^{(M)} - \mu \right) \left(\Delta_1 \widehat{L}_n^{(M)} - \mu \right) \right) \\
&= \sqrt{N} \left(\left(\widehat{a}_N^{(M)} \right)^2 - a^2 \right) \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \\
&\quad + \left(\widehat{a}_N^{(M)} - a \right) \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) + \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) \right) \\
&\quad + \frac{a^2 - \left(\widehat{a}_N^{(M)} \right)^2}{2M} \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) + \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) \right) \\
&\quad + \sqrt{N} \frac{a - \widehat{a}_N^{(M)}}{M} \frac{1}{N-1} \sum_{n=1}^{N-1} (Y_{n+1} - Y_n) (Y_n - Y_{n-1}) \\
&\quad + \sqrt{N} \frac{\left(\widehat{a}_N^{(M)} \right)^2 - a^2}{4M^2} \frac{1}{N-1} \sum_{n=1}^{N-1} (Y_{n+1} - Y_n) (Y_n - Y_{n-1}). \tag{5.2.7}
\end{aligned}$$

By Corollary 5.2.4 (i),(ii), as $N \wedge M \rightarrow \infty$ and $N/M \rightarrow 0$,

$$\frac{1}{N-1} \sum_{n=1}^{N-1} \left(\left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) (Y_n - Y_{n-1}) + \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) (Y_{n+1} - Y_n) \right) \xrightarrow{p} 0.$$

Hence, using Lemma 5.1.1 (ii), Corollary 5.2.4 (i),(ii), ergodicity, and Slutsky's Theorem, all terms of equation (5.2.7) converge to 0 in probability as $N \rightarrow \infty$ and $N/M \rightarrow 0$ with the exception of the first term.

For the first term in (5.2.7),

$$\begin{aligned}
& \sqrt{N} \left(\left(\widehat{a}_N^{(M)} \right)^2 - a^2 \right) \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) \\
&= \sqrt{N} \left(\widehat{a}_N^{(M)} - a \right) \left(\widehat{a}_N^{(M)} + a \right) \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\left(\frac{S_{n+1}^{(M)}}{M} - \frac{\mu}{a} \right) \left(\frac{S_n^{(M)}}{M} - \frac{\mu}{a} \right) - \frac{\eta^2}{2a^3} (e^{-a} - 1)^2 \right) \\
&\quad + \sqrt{N} \left(\widehat{a}_N^{(M)} - a \right) \left(\widehat{a}_N^{(M)} + a \right) \frac{\eta^2}{2a^3} (e^{-a} - 1)^2. \tag{5.2.8}
\end{aligned}$$

Using Theorem 5.1.1 (ii), Corollary 5.2.4 (iv), and Slutsky's Theorem, the first term of equation (5.2.8) converges to 0 in probability as $N \rightarrow \infty$ and $N/M \rightarrow 0$.

For the second term, using Theorem 5.1.1 (ii), as $N \rightarrow \infty$ and $N/M \rightarrow 0$ we have,

$$\epsilon_N^{(M)}(a, \eta) \equiv \sqrt{N} \left(\widehat{a}_N^{(M)} - a \right) \left(\widehat{a}_N^{(M)} + a \right) \frac{\eta^2}{2a^3} (e^{-a} - 1)^2 \xrightarrow{d} N \left(0, 2 \frac{\eta^4 (e^{-a} - 1)^4}{a^4} \right).$$

Summarizing, we have shown that

$$\begin{aligned} \sqrt{N} \widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(1) &= \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\widehat{\Delta_1 \widehat{L}_{n+1}^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) \left(\widehat{\Delta_1 \widehat{L}_n^{(M)}} - \frac{\widehat{a}_N^{(M)}}{a} \mu \right) + o_p(1) \\ &= \frac{\sqrt{N}}{N-1} \sum_{n=1}^{N-1} \left(\Delta_1 \widehat{L}_{n+1}^{(M)} - \mu \right) \left(\Delta_1 \widehat{L}_n^{(M)} - \mu \right) + \epsilon_N^{(M)}(a, \eta) + o_p(1) \\ &= \sqrt{N} \widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}} + \epsilon_N^{(M)}(a, \eta) + o_p(1). \end{aligned}$$

□

5.2.4 Proof of Theorem 5.2.1

By Theorem 5.2.6, Lemma 5.2.5 (iii), and Slutsky's Theorem we have that

$$\sqrt{N} \frac{\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(1)}{\widehat{\eta}^2} = \sqrt{N} \frac{\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(1)}{\widehat{\eta}^2} + \frac{\epsilon_N^{(M)}(a, \eta)}{\widehat{\eta}^2} + \frac{o_p(1)}{\widehat{\eta}^2}.$$

By Lemma 4.1.1,

$$\sqrt{N} \frac{\widehat{\gamma_{\Delta_1 \widehat{L}^{(M)}}}(1)}{\widehat{\eta}^2} \xrightarrow{d} N(0, 1) \text{ as } N \rightarrow \infty \text{ and } \frac{N}{M} \rightarrow 0.$$

By Slutsky's Theorem

$$\epsilon_N^{(M)}(a) = \frac{\epsilon_N^{(M)}(a, \eta)}{\widehat{\eta}^2} \xrightarrow{p} 0$$

as $N \rightarrow \infty$, $N/M \rightarrow 0$, $a \rightarrow \infty$, and $\frac{o_p(1)}{\widehat{\eta}^2} \xrightarrow{p} 0$ as $N \rightarrow \infty$, $N/M \rightarrow 0$. □

5.3 Simulation Studies

5.3.1 Brownian motion driven CAR(1) process

Table 5.1 gives computed empirical levels $\widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$, $\widehat{\alpha}_{\Delta_1 \widehat{\widehat{B}^{(M)}}}$ for tests based on $W_{\Delta_1 \widehat{B}^{(M)}}(1)$ (a known) and $W_{\Delta_1 \widehat{\widehat{B}^{(M)}}}(1)$ (a estimated) respectively, following the same procedures as before, over $R = 400$ simulations, with nominal level 0.05. We consider various values of the parameter a .

a	$\widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 \widehat{\widehat{B}^{(M)}}}$	a	$\widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 \widehat{\widehat{B}^{(M)}}}$	a	$\widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 \widehat{\widehat{B}^{(M)}}}$
0.1	0.0475	0.0375	7.0	0.0600	0.0625	14.0	0.0625	0.0625
0.5	0.0475	0.0075	7.5	0.0475	0.0475	14.5	0.0525	0.0500
1.0	0.0575	0.0200	8.0	0.0475	0.0525	15.0	0.0575	0.0600
1.5	0.0400	0.0225	8.5	0.0175	0.0175	15.5	0.0475	0.0475
2.0	0.0475	0.0400	9.0	0.0375	0.0400	16.0	0.0450	0.0475
2.5	0.0575	0.0475	9.5	0.0450	0.0400	20.0	0.0250	0.0225
3.0	0.0650	0.0500	10.0	0.0425	0.0475	30.0	0.0600	0.0575
3.5	0.0450	0.0600	10.5	0.0350	0.0325	40.0	0.0225	0.0225
4.0	0.0525	0.0425	11.0	0.0500	0.0525	50.0	0.0525	0.0525
4.5	0.0350	0.0275	11.5	0.0450	0.0475	60.0	0.0475	0.0475
5.0	0.0375	0.0450	12.0	0.0650	0.0675	70.0	0.0375	0.0350
5.5	0.0375	0.0300	12.5	0.0375	0.0400	80.0	0.0475	0.0475
6.0	0.0325	0.0350	13.0	0.0725	0.0725	90.0	0.0500	0.0475
6.5	0.0300	0.0250	13.5	0.0600	0.0600	100	0.0475	0.0475

Table 5.1: $N = 100, M = 500, R = 400, \sigma = 1$

5.3.2 Gamma-driven CAR(1) process

Table 5.2 gives computed empirical levels $\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$, $\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$ for tests based on $W_{\Delta_1 \widehat{G}^{(M)}}(1)$ and $W_{\Delta_1 \widehat{G}^{(M)}}(1)$ respectively, following the same procedures as before, over $R = 400$ simulations, with nominal level 0.05. We consider various values of the parameter a .

a	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$	a	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$	a	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$
0.1	0.0425	0.0275	7.0	0.025	0.020	14.0	0.0550	0.0550
0.5	0.0400	0.0075	7.5	0.0300	0.0325	14.5	0.0300	0.0300
1.0	0.0600	0.0225	8.0	0.0300	0.0300	15.0	0.0375	0.0375
1.5	0.0425	0.0250	8.5	0.0200	0.0275	15.5	0.0475	0.0500
2.0	0.0425	0.0150	9.0	0.0375	0.0375	16.0	0.0300	0.0350
2.5	0.0325	0.0200	9.5	0.0400	0.0350	20.0	0.0425	0.0450
3.0	0.0475	0.0475	10.0	0.0325	0.0350	30.0	0.0425	0.0425
3.5	0.0600	0.0325	10.5	0.0475	0.0475	40.0	0.0500	0.0500
4.0	0.0250	0.0225	11.0	0.0325	0.0325	50.0	0.0300	0.0300
4.5	0.0300	0.0225	11.5	0.0350	0.0450	60.0	0.0350	0.0350
5.0	0.0325	0.0300	12.0	0.0375	0.0350	70.0	0.0425	0.0450
5.5	0.0375	0.0350	12.5	0.0325	0.0275	80.0	0.0475	0.0475
6.0	0.0650	0.0650	13.0	0.0475	0.0500	90.0	0.0400	0.0400
6.5	0.0425	0.0400	13.5	0.0325	0.0300	100	0.0400	0.0400

Table 5.2: $N = 100, M = 500, R = 400, \sigma = 1$

The empirical levels of the test statistics using an estimator of a are close to those using the true value of a for $a \geq 3$. The variability for $a < 3$ is to be expected, because $\epsilon_N^{(M)}(a) \sim N\left(0, 2\frac{(e^{-a}-1)^4}{a^4}\right)$ has large variance for small a .

5.3.3 An alternative case

To illustrate the behavior of the test statistics $W_{\Delta_1 \widehat{L}^{(M)}}(1)$ and $W_{\Delta_1 \widehat{L}^{(M)}}(1)$ under an alternative case we follow the same procedures as in Section 4.2.4.

To graph the power functions for both tests at level 0.05 based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ and $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$, we computed the empirical rejection rate $\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}, \widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$ for different values of the Hurst parameter H with $a = 5$. The power functions are illustrated in Figure 5.1.

H	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	H	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$
0.05	1.00	0.74	0.55	0.10	0.10
0.1	1.00	0.76	0.60	0.31	0.12
0.15	0.99	0.69	0.65	0.47	0.48
0.2	0.94	0.46	0.7	0.77	0.85
0.25	0.89	0.29	0.75	0.95	0.98
0.3	0.66	0.13	0.8	0.96	0.98
0.35	0.28	0.10	0.85	1.00	0.99
0.4	0.24	0.05	0.9	1.00	0.95
0.45	0.04	0.04	0.95	1.00	0.97

Table 5.3: $N = 100, M = 200, R = 50, a = 5, \sigma = 1$

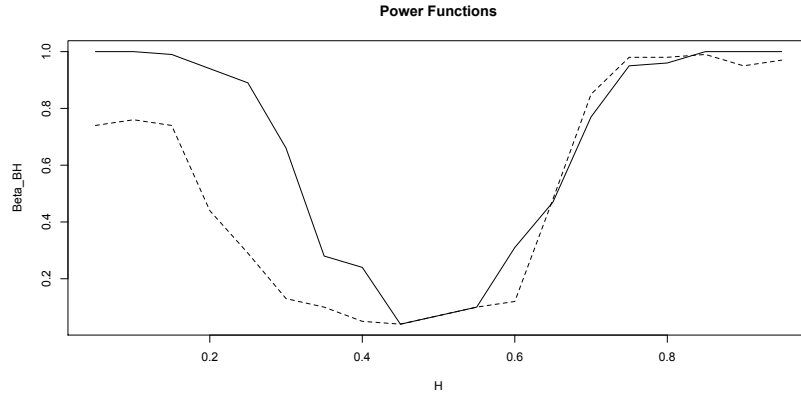


Figure 5.1: The power function for the test based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the solid line and the one based on $W_{\widehat{\Delta_1 B}_H^{(M)}}(1)$ is the dashed line. Fix $a = 5$.

Table 5.4 gives computed empirical rejection rate $\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$, $\widehat{\beta}_{\widehat{\Delta_1 B}_H^{(M)}}$ for the tests based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ and $W_{\widehat{\Delta_1 B}_H^{(M)}}(1)$ respectively for different values of the parameter a and fixed Hurst parameter $H = 0.8$. The power function is illustrated in Figure 5.2.

The power is good for small and large values of a , but the test may lack power for moderate values of a .

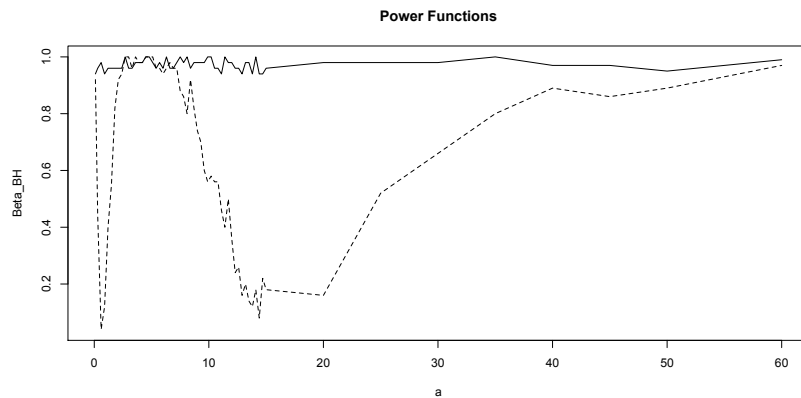


Figure 5.2: The power function for the test based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the solid line and the one based on $W_{\widehat{\Delta_1 B}_H^{(M)}}(1)$ is the dashed line. Fix $H = 0.8$.

a	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	a	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	a	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$
0.1	0.94	0.92	6.3	1.00	0.96	12.6	0.96	0.26
0.3	0.96	0.44	6.6	0.96	0.98	12.9	0.94	0.16
0.6	0.98	0.04	6.9	0.96	0.96	13.2	0.98	0.20
0.9	0.94	0.12	7.2	0.98	0.96	13.5	0.98	0.14
1.2	0.96	0.40	7.5	1.00	0.88	13.8	0.94	0.12
1.5	0.96	0.66	7.8	0.98	0.86	14.1	1.00	0.18
1.8	0.96	0.82	8.1	1.00	0.80	14.4	0.94	0.08
2.1	0.96	0.92	8.4	0.96	0.92	14.7	0.94	0.22
2.4	0.96	0.94	8.7	0.98	0.82	15.0	0.96	0.18
2.7	1.00	1.00	9.0	0.98	0.74	20.0	0.98	0.16
3.0	0.96	1.00	9.3	0.98	0.70	25.0	0.98	0.52
3.3	0.96	0.96	9.6	0.98	0.60	30.0	0.98	0.66
3.6	0.98	1.00	9.9	1.00	0.56	35.0	1.00	0.80
3.9	0.98	0.98	10.2	1.00	0.58	40.0	0.97	0.89
4.2	0.98	0.98	10.5	0.96	0.56	45.0	0.97	0.86
4.5	1.00	1.00	10.8	0.96	0.56	50.0	0.95	0.89
4.8	1.00	1.00	11.1	0.94	0.46	60.0	0.99	0.97
5.1	0.98	1.00	11.4	1.00	0.40	70.0	0.98	0.95
5.4	0.96	0.96	11.7	0.98	0.50	80.0	0.99	0.97
5.7	0.98	0.96	12.0	0.98	0.36	90.0	1.00	0.98
6.0	0.96	0.94	12.3	0.96	0.24	100	0.99	0.97

Table 5.4: $N = 100, M = 200, R = 50, H = 0.8, \sigma = 1$

Table 5.5 gives computed empirical rejection rate $\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$, $\widehat{\beta}_{\widehat{\Delta_1 B}_H^{(M)}}$ for the tests based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ and $W_{\widehat{\Delta_1 B}_H^{(M)}}(1)$ respectively for different values of the parameter a and fixed Hurst parameter $H = 0.65$. The power function is illustrated in Figure 5.3. Again we observe that the power is good for small and large values of a , but the test lacks power for moderate values of a .

a	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\widehat{\Delta_1 B}_H^{(M)}}$	a	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\widehat{\Delta_1 B}_H^{(M)}}$	a	$\widehat{\beta}_{\Delta_1 \widehat{B}_H^{(M)}}$	$\widehat{\beta}_{\widehat{\Delta_1 B}_H^{(M)}}$
0.1	0.42	0.26	7.0	0.40	0.17	14.0	0.48	0.13
0.5	0.48	0.07	7.5	0.57	0.10	14.5	0.48	0.11
1.0	0.48	0.44	8.0	0.55	0.12	15.0	0.57	0.20
1.5	0.51	0.78	8.5	0.51	0.09	15.5	0.47	0.09
2.0	0.47	0.83	9.0	0.50	0.05	16.0	0.56	0.09
2.5	0.50	0.90	9.5	0.55	0.07	20	0.48	0.17
3.0	0.48	0.84	10.0	0.50	0.10	30.0	0.47	0.30
3.5	0.53	0.81	10.5	0.46	0.08	40.0	0.51	0.37
4.0	0.48	0.72	11.0	0.44	0.03	50.0	0.42	0.39
4.5	0.53	0.52	11.5	0.54	0.07	60.0	0.49	0.42
5.0	0.51	0.60	12.0	0.47	0.08	70.0	0.53	0.50
5.5	0.49	0.46	12.5	0.48	0.07	80.0	0.42	0.38
6.0	0.42	0.41	13.0	0.48	0.06	90.0	0.38	0.38
6.5	0.46	0.34	13.5	0.38	0.08	100	0.53	0.49

Table 5.5: $N = 100, M = 200, R = 50, H = 0.65, \sigma = 1$

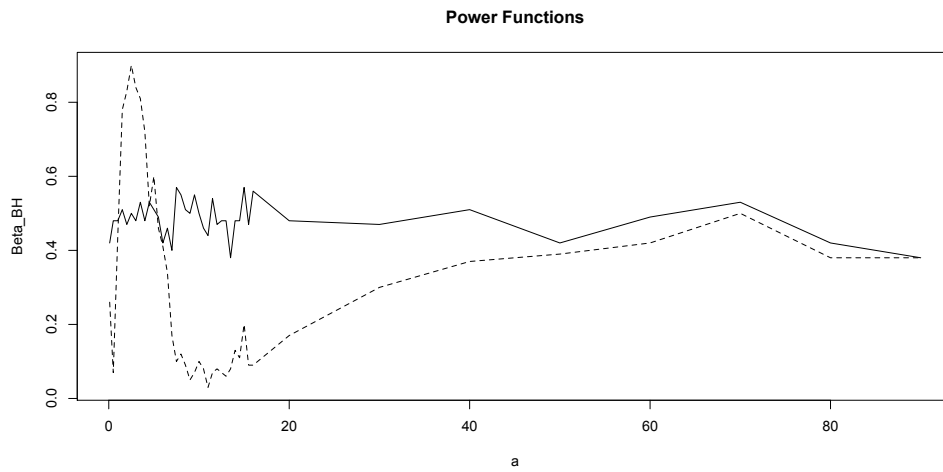


Figure 5.3: The power function for the test based on $W_{\Delta_1 \widehat{B}_H^{(M)}}(1)$ is the solid line and the one based on $W_{\widehat{B}_H^{(M)}}(1)$ is the dashed line. Fix $H = 0.65$.

Chapter 6

Goodness-of-Fit Test

Given a CAR(1) model, if the hypothesis of independent increments of the driving process has not been rejected, the next step is to identify the underlying Lévy process. Therefore, in this chapter we consider empirical processes and goodness-of-fit tests based on estimated increments of the recovered driving process. The goal is to develop a test of goodness of fit for the unobserved Lévy driving process of the CAR(1) model. We will test the composite hypothesis $H_0 : L(1) \sim F$, where $F(\cdot) = F(\cdot; \mu, \eta^2) \in \mathcal{A}$, where \mathcal{A} is a class of distributions completely determined by μ and η^2 (for example, this will be the case if L is a Brownian motion or a gamma process).

We proceed as follows. In the following section, we start with the empirical process based on true increments of the Lévy process. The asymptotic theory is classical. Then we proceed with complications: we have to estimate parameters μ and η which induces an additional error ([22], [15]). Furthermore, we do not observe the driving process. It turns out that all these issues can be handled by a beautiful bootstrap-type trick from [16]. Our main result is Theorem 6.1.1. The proof is long and technical (Section 6.2). We conclude with numerical studies (6.3).

6.1 Introduction and Main Results

We now describe our setup in detail. We assume that Y is a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (2.2.2) holds:

$$\mathbb{E}[L(t)] = \mu t \quad \text{and} \quad \text{Var}(L(t)) = \eta^2 t, \quad t \geq 0. \quad (6.1.1)$$

Recall equations (3.3.6) and (3.3.7) for the true and estimated increments:

$$\Delta_1 L_n = \frac{a}{\sigma} \int_{n-1}^n Y(s) ds + \frac{1}{\sigma} (Y(n) - Y(n-1)), \quad n = 1, \dots, N, \quad (6.1.2)$$

$$\Delta_1 \widehat{L}_n^{(M)} = \frac{a}{M\sigma} \sum_{i=(n-1)M+1}^{nM} Y_{\frac{i}{M}} + \left(\frac{1}{\sigma} - \frac{a}{2M\sigma} \right) (Y_n - Y_{n-1}), \quad n = 1, \dots, N. \quad (6.1.3)$$

By Lemma 3.3.5 (i) we have that $\forall n \in \mathbb{N}$,

$$\left\| \Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n \right\|_{L_2} \leq \eta \sqrt{a} \left(1 - e^{-\frac{a}{M}} \right)^{\frac{1}{2}} + \frac{\eta \sqrt{a}}{2M} (1 - e^{-a})^{\frac{1}{2}}. \quad (6.1.4)$$

The sequence $\{\Delta_1 L_n, n = 1, \dots, N\}$ is an *i.i.d.* sequence with distribution function F . We assume that the distribution of F belongs to a class \mathcal{A} determined by μ and η^2 . For example, $\mathcal{A} = \{N(\mu, \eta^2), \mu \in \mathbb{R}, \eta > 0\}$ or $\mathcal{A} = \{\Gamma(\alpha, \beta), \alpha = \frac{\mu^2}{\eta^2}, \beta = \frac{\eta^2}{\mu}, \alpha > 0, \beta > 0\}$. To emphasize this dependence on μ and η , we write $F = F(\cdot; \mu, \eta^2)$.

We introduce the following notation:

$$\widehat{F}_{\Delta_1 L_N}(t) \equiv \frac{1}{N} \sum_{n=1}^N \mathbf{I}\{\Delta_1 L_n \leq t\} \quad (6.1.5)$$

denotes the empirical distribution function of $\{\Delta_1 L_n, n = 1, \dots, N\}$. By the empirical central limit theorem we have

$$\sqrt{N} \left(\widehat{F}_{\Delta_1 L_N}(\cdot) - F(\cdot; \mu, \eta^2) \right) \xrightarrow{\mathcal{D}} B_F(\cdot) \quad \text{as } N \rightarrow \infty,$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution in the Skorohod topology on $D(\mathbb{R})$, $F = F(\cdot; \mu, \eta^2)$ is the cdf of $\Delta_1 L_1$ and $B_F(\cdot)$ is a Gaussian process with zero mean and covariance given by:

$$\text{Cov}(B_F(s), B_F(t)) = E[B_F(s)B_F(t)] = F(s) \wedge F(t) - F(s)F(t).$$

However, even if L is observable, since the hypothesis is composite and μ and η^2 are unspecified, we will estimate μ and η^2 by the sample mean $\overline{\Delta_1 L}$ and the sample variance $\widehat{\eta}_L^2$ of $\Delta_1 L_n$ respectively:

$$\overline{\Delta_1 L} \equiv \frac{1}{N} \sum_{n=1}^N \Delta_1 L_n \quad \text{and} \quad \widehat{\eta}_L^2 \equiv \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \overline{\Delta_1 L})^2. \quad (6.1.6)$$

Unfortunately, the process

$$\sqrt{N} \left(\widehat{F}_{\Delta_1 L_N}(\cdot) - F(\cdot; \overline{\Delta_1 L}, \widehat{\eta}_L^2) \right)$$

does not lead to distribution free test (see [16], [22] and [15]). In order to resolve this problem we use a technique of Burke and Gombay [16] that uses bootstrapped estimators of μ and η^2 .

To be more precise, let $\{w_n, n = 1, \dots, N\}$ be *i.i.d.* multinomial $(1, \frac{1}{N}, \dots, \frac{1}{N})$, on $\{1, 2, \dots, N\}$, i.e. $P(w_n = i) = \frac{1}{N}, n = 1, \dots, N$, and the sequence w_n is independent of both $\Delta_1 L_n$ and $\Delta_1 \widehat{L}_n^{(M)}$ for all $n \geq 1$.

Define $\{\Delta_1 L_n^*, \Delta_1 \widehat{L}_n^{*(M)}, n = 1, \dots, N\}$ as

$$\Delta_1 L_n^* \equiv \sum_{i=1}^N \mathbf{I}\{w_n = i\} \Delta_1 L_i \quad \text{and} \quad \Delta_1 \widehat{L}_n^{*(M)} \equiv \sum_{i=1}^N \mathbf{I}\{w_n = i\} \Delta_1 \widehat{L}_i^{(M)}. \quad (6.1.7)$$

Define the sample mean and variance based on the bootstrapped true increments and bootstrapped estimated increments as,

$$\overline{\Delta_1 L^*} \equiv \frac{1}{N} \sum_{n=1}^N \Delta_1 L_n^* \quad \text{and} \quad \widehat{\eta}_L^{*2} \equiv \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n^* - \overline{\Delta_1 L^*})^2 \quad (6.1.8)$$

and

$$\overline{\Delta_1 \widehat{L}^{*(M)}} \equiv \frac{1}{N} \sum_{n=1}^N \Delta_1 \widehat{L}_n^{*(M)} \quad \text{and} \quad \widehat{\eta}^{*2} \equiv \frac{1}{N} \sum_{n=1}^N \left(\Delta_1 \widehat{L}_n^{*(M)} - \overline{\Delta_1 \widehat{L}^{*(M)}} \right)^2. \quad (6.1.9)$$

We will show in Proposition 6.2.1 that the conditions of Theorem 3.5 of [16] are satisfied, and so

$$\sqrt{N} \left(\widehat{F}_{\Delta_1 L_N}(\cdot) - F \left(\cdot; \overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2} \right) \right) \xrightarrow{d} B_F(\cdot) \quad \text{as } N \rightarrow \infty. \quad (6.1.10)$$

However, a further complication is that we do not observe the true increments, $\Delta_1 L_n$, which must be replaced by the estimated increments, $\Delta_1 \widehat{L}_n^{(M)}$. Therefore we will extend (6.1.10) to the weak convergence in $D(\mathbb{R})$ of:

$$\sqrt{N} \left(\widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(\cdot) - F \left(\cdot; \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right),$$

where

$$\widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(t) = \frac{1}{N} \sum_{n=1}^N \mathbf{I} \left\{ \Delta_1 \widehat{L}_n^{(M)} \leq t \right\}.$$

Our main result is Theorem 6.1.1. From this, we will be able to develop distribution-free tests of the composite hypothesis, $H_0 : L(1) \sim F \in \mathcal{A}$.

Theorem 6.1.1. *Let Y be a strictly stationary CAR(1) process driven by a second-order Lévy process L such that (6.1.1) holds. We assume that $F(\cdot; \mu, \eta^2)$, the cdf of $\Delta_1 L_1$, belongs to a class \mathcal{A} of distributions that are characterized by μ, η^2 , and for all $F(\cdot; \mu, \eta^2) \in \mathcal{A}$:*

- $F(t; \mu, \eta^2)$ is continuous with bounded derivative in t ;
- The vector

$$\nabla_{\mu, \eta^2} F(t; \mu, \eta^2) = \left(\frac{\partial F}{\partial \mu}, \frac{\partial F}{\partial \eta^2} \right)^T \quad (6.1.11)$$

is uniformly continuous in t and $(\mu, \eta^2) \in \Lambda$, where Λ is the closure of a given neighbourhood of the true value of (μ, η^2) ;

- The sample mean $(\overline{\Delta_1 L})$ and the sample variance $(\widehat{\eta}_L^2)$ are MLE's of μ and η^2 that satisfy condition (3.2) in [16];
- F has a finite fourth moment.

Then,

$$\sqrt{N} \left(\widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(\cdot) - F \left(\cdot ; \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right) \xrightarrow{d} B_F(\cdot) \quad \text{as } \frac{N^2}{M} \rightarrow 0 \quad \text{and } N \rightarrow \infty,$$

where $B_F(t)$ is a Gaussian process with zero mean and covariance given by:

$$\text{Cov}(B_F(s), B_F(t)) = E[B_F(s)B_F(t)] = F(s) \wedge F(t) - F(s)F(t).$$

6.2 Proof of Theorem 6.1.1

The proof appears as a series of propositions. First we prove (6.1.10).

Proposition 6.2.1. *Under the assumptions of Theorem 6.1.1, we have*

$$\sqrt{N} \left(\widehat{F}_{\Delta_1 L_N}(\cdot) - F \left(\cdot ; \overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2} \right) \right) \xrightarrow{\mathcal{D}} B_F(t) \quad \text{as } N \rightarrow \infty,$$

where

$$\widehat{F}_{\Delta_1 L_N}(t) = \frac{1}{N} \sum_{n=1}^N \mathbf{I}\{\Delta_1 L_n \leq t\},$$

and $B_F(t)$ is a Gaussian process with zero mean and covariance given by:

$$\text{Cov}(B_F(s), B_F(t)) = E[B_F(s)B_F(t)] = F(s) \wedge F(t) - F(s)F(t).$$

Proof: In what follows, we show that the vector of estimators $\left(\overline{\Delta_1 L}, \widehat{\eta}_L^2 \right)^T$ satisfies conditions $(A_1) - (A_6)$ of Burke & Gombay [16]:

(A_1) : We verify

$$\sqrt{N} \left(\left(\begin{array}{c} \overline{\Delta_1 L} \\ \widehat{\eta}_L^2 \end{array} \right) - \left(\begin{array}{c} \mu \\ \eta^2 \end{array} \right) \right)$$

$$\begin{aligned}
&= \sqrt{N} \begin{pmatrix} \overline{\Delta_1 L} - \mu \\ \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \mu)^2 - \eta^2 \end{pmatrix} + \sqrt{N} \begin{pmatrix} 0 \\ \hat{\eta}_L^2 - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \mu)^2 \end{pmatrix} \\
&= \frac{1}{\sqrt{N}} \sum_{n=1}^N \begin{pmatrix} \Delta_1 L_n - \mu \\ (\Delta_1 L_n - \mu)^2 - \eta^2 \end{pmatrix} + \begin{pmatrix} 0 \\ \sqrt{N} \left(\hat{\eta}_L^2 - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \mu)^2 \right) \end{pmatrix},
\end{aligned}$$

and

$$\begin{aligned}
\sqrt{N} \left(\hat{\eta}_L^2 - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \mu)^2 \right) &= \frac{1}{\sqrt{N}} \sum_{n=1}^N \left((\Delta_1 L_n - \overline{\Delta_1 L})^2 - (\Delta_1 L_n - \mu)^2 \right) \\
&= \frac{1}{\sqrt{N}} \sum_{n=1}^N \left(-2\Delta_1 L_n \overline{\Delta_1 L} + (\overline{\Delta_1 L})^2 + 2\mu\Delta_1 L_n - \mu^2 \right) \\
&= \frac{1}{\sqrt{N}} \sum_{n=1}^N \left(2\Delta_1 L_n (\mu - \overline{\Delta_1 L}) - (\mu^2 - (\overline{\Delta_1 L})^2) \right) \\
&= \sqrt{N} (\mu - \overline{\Delta_1 L}) \frac{1}{N} \sum_{n=1}^N (2\Delta_1 L_n - (\mu + \overline{\Delta_1 L})) \\
&= \sqrt{N} (\mu - \overline{\Delta_1 L}) (\overline{\Delta_1 L} - \mu) \xrightarrow{p} 0 \text{ as } N \rightarrow \infty.
\end{aligned} \tag{6.2.1}$$

Hence by letting

$$l(x, \mu, \eta^2) = \begin{bmatrix} x - \mu \\ (x - \mu)^2 - \eta^2 \end{bmatrix},$$

condition (A₁) of [16] is satisfied.

$$(A_2) : \mathbb{E} \begin{bmatrix} \Delta_1 L_n - \mu \\ (\Delta_1 L_n - \mu)^2 - \eta^2 \end{bmatrix} = 0.$$

(A₃) :

$$\begin{aligned}
\Sigma &= \mathbb{E} \left[\begin{pmatrix} \Delta_1 L_n - \mu \\ (\Delta_1 L_n - \mu)^2 - \eta^2 \end{pmatrix}^T \begin{pmatrix} \Delta_1 L_n - \mu \\ (\Delta_1 L_n - \mu)^2 - \eta^2 \end{pmatrix} \right] \\
&= \begin{pmatrix} \text{Var}(\Delta_1 L_n - \mu) & \text{Cov}(\Delta_1 L_n - \mu, (\Delta_1 L_n - \mu)^2 - \eta^2) \\ \text{Cov}(\Delta_1 L_n - \mu, (\Delta_1 L_n - \mu)^2 - \eta^2) & \text{Var}((\Delta_1 L_n - \mu)^2 - \eta^2) \end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} \eta^2 & \mathbb{E} [\Delta_1 L_n - \mu]^3 \\ \mathbb{E} [\Delta_1 L_n - \mu]^3 & \mathbb{E} [(L_1 - \mu)^4] - \eta^4 \end{pmatrix}$$

which is a finite non-negative definite matrix if $\mathbb{E} [L_1^4] < \infty$.

(A₄): The vector $\nabla_{\mu, \eta^2} F(t; \mu, \eta^2) = \left(\frac{\partial F}{\partial \mu}, \frac{\partial F}{\partial \eta^2} \right)^T$ is uniformly continuous in t and $(\mu, \eta^2) \in \Lambda$ by assumption.

(A₅): Each component of the vector function

$$l(x, \mu, \eta^2) = \begin{bmatrix} x - \mu \\ (x - \mu)^2 - \eta^2 \end{bmatrix}$$

is of bounded variation on each finite interval.

We assume that the sample mean $(\overline{\Delta_1 L})$ and the sample variance $(\widehat{\eta_L^2})$ satisfy condition (3.2) in [16]; i.e.,

$$l(x, \mu, \eta^2) = \nabla_{\mu, \eta^2} \log f(x, \mu, \eta^2) \cdot I^{-1}(\mu, \eta^2),$$

where f is the density function of F and $I^{-1}(\mu, \eta^2)$ is the inverse of the Fisher information matrix:

$$I(\mu, \eta^2) = \mathbb{E} [\nabla_{\mu, \eta^2} \log f(\Delta L_1, \mu, \eta^2)]' \cdot [\nabla_{\mu, \eta^2} \log f(\Delta L_1, \mu, \eta^2)]$$

Brownian motion and Poisson process are examples of Lévy processes that satisfy this condition; gamma process, however, does not.

(A₆):

Remark 6.2.2. *There is a small typo on condition (A₆) of [16] and the correct version is:*

$$\sqrt{N} \left(\begin{pmatrix} \overline{\Delta_1 L^*} \\ \widehat{\eta_L^{*2}} \end{pmatrix} - \begin{pmatrix} \overline{\Delta_1 L} \\ \widehat{\eta_L^2} \end{pmatrix} \right) = \frac{1}{\sqrt{N}} \sum_{n=1}^N (l(\Delta_1 L_n^*, \mu, \eta^2) - l(\Delta_1 L_n, \mu, \eta^2)) + \epsilon_N,$$

where $\epsilon_N \xrightarrow{p} 0$ as $N \rightarrow \infty$.

We have,

$$\begin{aligned}
& \sqrt{N} \left(\left(\begin{array}{c} \overline{\Delta_1 L^*} \\ \widehat{\eta}_L^{*2} \end{array} \right) - \left(\begin{array}{c} \overline{\Delta_1 L} \\ \widehat{\eta}_L^2 \end{array} \right) \right) \\
&= \sqrt{N} \left(\left(\begin{array}{c} \overline{\Delta_1 L^*} \\ \widehat{\eta}_L^{*2} \end{array} \right) - \begin{pmatrix} \mu \\ \eta^2 \end{pmatrix} - \left(\left(\begin{array}{c} \overline{\Delta_1 L} \\ \widehat{\eta}_L^2 \end{array} \right) \right) - \begin{pmatrix} \mu \\ \eta^2 \end{pmatrix} \right) \\
&= \frac{1}{\sqrt{N}} \sum_{n=1}^N \begin{pmatrix} \Delta_1 L_n^* - \mu \\ (\Delta_1 L_n^* - \mu)^2 - \eta^2 \end{pmatrix} + \begin{pmatrix} 0 \\ \sqrt{N} \left(\widehat{\eta}_L^{*2} - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n^* - \mu)^2 \right) \end{pmatrix} \\
&\quad - \frac{1}{\sqrt{N}} \sum_{n=1}^N \begin{pmatrix} \Delta_1 L_n - \mu \\ (\Delta_1 L_n - \mu)^2 - \eta^2 \end{pmatrix} - \begin{pmatrix} 0 \\ \sqrt{N} \left(\widehat{\eta}_L^2 - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \mu)^2 \right) \end{pmatrix} \\
&= \frac{1}{\sqrt{N}} \sum_{n=1}^N \begin{pmatrix} \Delta_1 L_n^* - \mu \\ (\Delta_1 L_n^* - \mu)^2 - \eta^2 \end{pmatrix} - \frac{1}{\sqrt{N}} \sum_{n=1}^N \begin{pmatrix} \Delta_1 L_n - \mu \\ (\Delta_1 L_n - \mu)^2 - \eta^2 \end{pmatrix} \\
&\quad + \begin{pmatrix} 0 \\ \sqrt{N} \left(\widehat{\eta}_L^{*2} - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n^* - \mu)^2 \right) \end{pmatrix} - \begin{pmatrix} 0 \\ \sqrt{N} \left(\widehat{\eta}_L^2 - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \mu)^2 \right) \end{pmatrix}.
\end{aligned}$$

We have shown in (A₁) that

$$\sqrt{N} \left(\widehat{\eta}_L^2 - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n - \mu)^2 \right) \xrightarrow{p} 0 \text{ as } N \rightarrow \infty,$$

and similarly to equation (6.2.1) we have

$$\sqrt{N} \left(\widehat{\eta}_L^{*2} - \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n^* - \mu)^2 \right) = \sqrt{N} (\mu - \overline{\Delta_1 L^*}) (\overline{\Delta_1 L^*} - \mu).$$

Recall equation (6.1.7),

$$\overline{\Delta_1 L^*} = \frac{1}{N} \sum_{n=1}^N \Delta_1 L_n^* = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^N \mathbf{I}\{w_n = i\} \Delta_1 L_i.$$

Note that: $\mathbb{E} [\overline{\Delta_1 L^*}] = \mu$. Now, noting that

$$\mathbf{I}(w_n = i) \mathbf{I}(w_n = j) = 0, \quad i \neq j,$$

we have

$$\begin{aligned}
\|\overline{\Delta_1 L^\star} - \mu\|_{L_2}^2 &= \mathbb{E} [\overline{\Delta_1 L^\star} - \mu]^2 = \mathbb{E} [\overline{\Delta_1 L^\star}]^2 - \mu^2 \\
&= \frac{1}{N^2} \mathbb{E} \left[\sum_{n=1}^N \sum_{i=1}^N \mathbf{I}\{w_n = i\} \Delta_1 L_i \right]^2 - \mu^2 \\
&= \frac{1}{N^2} \mathbb{E} \left[\sum_{m=1}^N \sum_{j=1}^N \sum_{n=1}^N \sum_{i=1}^N \mathbf{I}\{w_n = i\} \mathbf{I}\{w_m = j\} \Delta_1 L_i \Delta_1 L_j \right] - \mu^2 \\
&= \frac{1}{N^2} \mathbb{E} \left[\sum_{n=m=1}^N \sum_{j=1}^N \sum_{i=1}^N \mathbf{I}\{w_n = i\} \mathbf{I}\{w_n = j\} \Delta_1 L_i \Delta_1 L_j \right] \\
&\quad + \frac{1}{N^2} \mathbb{E} \left[\sum_{n \neq m}^N \sum_{j=1}^N \sum_{i=1}^N \mathbf{I}\{w_n = i\} \mathbf{I}\{w_m = j\} \Delta_1 L_i \Delta_1 L_j \right] - \mu^2 \\
&= \frac{N}{N^2} \sum_{n=1}^N P(w_n = i) \mathbb{E} [\Delta_1 L_1]^2 \\
&\quad + \frac{1}{N^2} \sum_{n \neq m}^N \sum_{i=j=1}^N P(w_n = i) P(w_m = i) \mathbb{E} [\Delta_1 L_i]^2 \\
&\quad + \frac{1}{N^2} \mathbb{E} \left[\sum_{n \neq m}^N \sum_{i \neq j}^N \mathbf{I}(w_n = i) \mathbf{I}(w_m = j) \Delta_1 L_i \Delta_1 L_j \right] - \mu^2 \\
&= \frac{1}{N} \mathbb{E} [\Delta_1 L_1]^2 \\
&\quad + \frac{1}{N^4} (N^2 - N) N \mathbb{E} [\Delta_1 L_i]^2 \\
&\quad + \frac{1}{N^2} (N^2 - N)^2 \frac{1}{N^2} \mu^2 - \mu^2 \rightarrow 0 \text{ as } N \rightarrow \infty,
\end{aligned}$$

hence, $\overline{\Delta_1 L^\star} \xrightarrow{p} \mu$ as $N \rightarrow \infty$. Furthermore, it is standard that almost surely, the conditional distribution of $\sqrt{N} (\overline{\Delta_1 L^\star} - \mu)$ converges weakly to a $N(0, \eta^2)$ distribution. Therefore, for any subsequence (N') we can take a further subsequence (N'') so that with probability 1, as $N'' \rightarrow \infty$, $(\overline{\Delta_1 L^\star} - \mu) \rightarrow 0$ and $\sqrt{N} (\overline{\Delta_1 L^\star} - \mu) \xrightarrow{d} N(0, \eta^2)$, and so almost surely $\sqrt{N''} (\overline{\Delta_1 L^\star} - \mu) (\overline{\Delta_1 L^\star} - \mu) \xrightarrow{p} 0$ as $N'' \rightarrow \infty$.

For any $k > 0$ and for almost all ω ,

$$P\left(\sqrt{N''} (\overline{\Delta_1 L^\star} - \mu) (\overline{\Delta_1 L^\star} - \mu) > k \mid L(\cdot, \omega)\right) \rightarrow 0 \text{ as } N'' \rightarrow \infty.$$

Therefore, by dominated convergence,

$$P\left(\sqrt{N''}(\overline{\Delta_1 L^*} - \mu)(\overline{\Delta_1 L^*} - \mu) > k\right) \rightarrow 0 \text{ as } N'' \rightarrow \infty,$$

for all $k > 0$, and so $\sqrt{N''}(\overline{\Delta_1 L^*} - \mu)(\overline{\Delta_1 L^*} - \mu) \xrightarrow{p} 0$.

Now suppose that

$$\sqrt{N}(\overline{\Delta_1 L^*} - \mu)(\overline{\Delta_1 L^*} - \mu) \not\xrightarrow{p} 0.$$

Then there exist $\epsilon > 0$, $\delta > 0$ and a subsequence (N') such that

$$P\left(\sqrt{N'}(\overline{\Delta_1 L^*} - \mu)(\overline{\Delta_1 L^*} - \mu) > \epsilon\right) > \delta \text{ for all } N'.$$

But by the argument above there exists a subsequence (N'') such that

$$P\left(\sqrt{N''}(\overline{\Delta_1 L^*} - \mu)(\overline{\Delta_1 L^*} - \mu) > \epsilon\right) \rightarrow 0 \text{ as } N'' \rightarrow \infty,$$

which is a contradiction. Therefore,

$$\sqrt{N}(\overline{\Delta_1 L^*} - \mu)(\overline{\Delta_1 L^*} - \mu) \xrightarrow{p} 0 \text{ as } N \rightarrow \infty,$$

hence (A_6) is satisfied. Proposition 6.2.1 follows by Theorem 3.5 of [16]. \square

It remains to show that $\Delta_1 L_n$ can be replaced by $\Delta_1 \widehat{L}_n^{(M)}$.

Proposition 6.2.3. *Let $\overline{\Delta_1 L^*}$, $\widehat{\eta}_L^{*2}$, $\overline{\Delta_1 \widehat{L}^{*(M)}}$, and $\widehat{\eta}^{*2}$ be defined as in equations (6.1.8) and (6.1.9). Under the assumptions of Theorem 6.1.1 we have*

$$\sup_t \left| \sqrt{N} \left(F\left(t; \overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2}\right) - F\left(t; \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2}\right) \right) \right| \xrightarrow{p} 0 \text{ as } \sqrt{\frac{N}{M}} \rightarrow 0 \text{ and } N \rightarrow \infty$$

To prove Proposition 6.2.3 we need two lemmas that show the closeness between the estimators of μ , η^2 defined by the bootstrapped true increments and bootstrapped estimated increments.

Lemma 6.2.4. *Let $\Delta_1 L_n^*$ and $\Delta_1 \widehat{L}_n^{*(M)}$ be defined as in equation (6.1.7). Then*

$$\left\| \Delta_1 L_n^* - \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2} = O\left(\frac{1}{\sqrt{M}}\right) \text{ uniformly in } n \in \mathbb{N}.$$

Proof of Lemma 6.2.4: We have

$$\begin{aligned}
\left\| \Delta_1 L_n^* - \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2}^2 &= \left\| \sum_{i=1}^N \mathbf{I}\{w_n = i\} \left(\Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right) \right\|_{L_2}^2 \\
&= \mathbb{E} \left[\sum_{i=1}^N \mathbf{I}\{w_n = i\} \left(\Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right) \right]^2 \\
&= \mathbb{E} \left[\sum_{j=1}^N \sum_{i=1}^N \mathbf{I}\{w_n = i\} \mathbf{I}\{w_n = j\} \left(\Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right) \left(\Delta_1 L_j - \Delta_1 \widehat{L}_j^{(M)} \right) \right] \\
&= \mathbb{E} \left[\sum_{i \neq j}^N \mathbf{I}\{w_n = i\} \mathbf{I}\{w_n = j\} \left(\Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right) \left(\Delta_1 L_j - \Delta_1 \widehat{L}_j^{(M)} \right) \right. \\
&\quad \left. + \sum_{i=j=1}^N \mathbf{I}\{w_n = i\} \left(\Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right)^2 \right] \\
&= \mathbb{E} \left[\sum_{i=1}^N \mathbf{I}\{w_n = i\} \left(\Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right)^2 \right]
\end{aligned}$$

since for all n , $\mathbf{I}\{w_n = i\} \mathbf{I}\{w_n = j\} = 0$ if $i \neq j$.

Hence, by independence and Lemma 3.3.5

$$\begin{aligned}
\left\| \Delta_1 L_n^* - \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2}^2 &= \sum_{i=1}^N P(w_n = i) \left\| \Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right\|_{L_2}^2 \\
&= \left\| \Delta_1 L_i - \Delta_1 \widehat{L}_i^{(M)} \right\|_{L_2}^2 = O\left(\frac{1}{M}\right).
\end{aligned}$$

□

Lemma 6.2.5. Let $\overline{\Delta_1 L^*}$, $\widehat{\eta}_L^{*2}$, $\overline{\Delta_1 \widehat{L}^{*(M)}}$, and $\widehat{\eta}^{*2}$ be defined as in equations (6.1.8) and (6.1.9). Then

- (i) $\sqrt{N} \left\| \overline{\Delta_1 L^*} - \overline{\Delta_1 \widehat{L}^{*(M)}} \right\|_{L_2} = O\left(\sqrt{\frac{N}{M}}\right)$,
- (ii) $\sqrt{N} \left\| \widehat{\eta}_L^{*2} - \widehat{\eta}^{*2} \right\|_{L_1} = O\left(\sqrt{\frac{N}{M}}\right)$.

Proof of Lemma 6.2.5:

$$\begin{aligned} (i) : \sqrt{N} \left\| \overline{\Delta_1 L^*} - \overline{\Delta_1 \widehat{L}^{*(M)}} \right\|_{L_2} &\leq \frac{1}{\sqrt{N}} \sum_{n=1}^N \left\| \Delta_1 L_n^* - \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2} \\ &= \sqrt{N} \left\| \Delta_1 L_n^* - \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2} = O \left(\sqrt{\frac{N}{M}} \right) \end{aligned}$$

by Lemma 6.2.4.

$$\begin{aligned} (ii) : \sqrt{N} \left\| \widehat{\eta}_L^{*2} - \widehat{\eta}^{*2} \right\|_{L_1} &= \sqrt{N} \left\| \frac{1}{N} \sum_{n=1}^N (\Delta_1 L_n^* - \overline{\Delta_1 L^*})^2 - \frac{1}{N} \sum_{n=1}^N (\Delta_1 \widehat{L}_n^{*(M)} - \overline{\Delta_1 \widehat{L}^{*(M)}})^2 \right\|_{L_1} \\ &\leq \sqrt{N} \left\| \Delta_1 L_n^* + \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2} \left\| \Delta_1 L_n^* - \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2} \\ &\quad + \sqrt{N} \left\| \overline{\Delta_1 L^*} + \overline{\Delta_1 \widehat{L}^{*(M)}} \right\|_{L_2} \left\| \overline{\Delta_1 L^*} - \overline{\Delta_1 \widehat{L}^{*(M)}} \right\|_{L_2} \\ &\leq 2\sqrt{N} \left\| \Delta_1 L_n^* + \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2} \left\| \Delta_1 L_n^* - \Delta_1 \widehat{L}_n^{*(M)} \right\|_{L_2} \\ &\leq O \left(\sqrt{\frac{N}{M}} \right) \end{aligned}$$

by Lemma 6.2.4. □

Proof of Proposition 6.2.3:

By Taylor's theorem we have

$$\begin{aligned} &\sqrt{N} \left(F \left(t; \overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2} \right) - F \left(t; \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right) \\ &= \sqrt{N} \left(\left(\overline{\Delta_1 L^*} - \overline{\Delta_1 \widehat{L}^{*(M)}} \right), \left(\widehat{\eta}_L^{*2} - \widehat{\eta}^{*2} \right) \right) \cdot \begin{pmatrix} \frac{\partial F(t; \theta_1, \theta_2)}{\partial \mu} \\ \frac{\partial F(t; \theta_1, \theta_2)}{\partial \eta^2} \end{pmatrix} \end{aligned}$$

for some θ_1, θ_2 such that: $|\theta_1 - \overline{\Delta_1 L^*}| \leq |\overline{\Delta_1 L^*} - \overline{\Delta_1 \widehat{L}^{*(M)}}|$ and $|\theta_2 - \widehat{\eta}_L^{*2}| \leq |\widehat{\eta}_L^{*2} - \widehat{\eta}^{*2}|$.

By Lemma 6.2.5 we have $N \rightarrow \infty$ and $N/M \rightarrow 0$,

$$\sqrt{N} \left(\overline{\Delta_1 L^*} - \overline{\Delta_1 \widehat{L}^{*(M)}} \right) \xrightarrow{p} 0 \quad \text{and} \quad \sqrt{N} \left(\widehat{\eta}_L^{*2} - \widehat{\eta}^{*2} \right) \xrightarrow{p} 0 \quad \text{as} \quad \sqrt{\frac{N}{M}} \rightarrow 0.$$

Since the probability that $\left(\overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2} \right)$ and $\left(\overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right)$ lie within \wedge converges to 1,

$$\sup_t \left| \sqrt{N} \left(F \left(t; \overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2} \right) - F \left(t; \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right) \right| \xrightarrow{p} 0$$

as $N \rightarrow \infty$ and $N/M \rightarrow 0$ by (A_4) . \square

The next result shows that the empirical distribution function based on the estimated increments is close to the empirical distribution function based on the true increments.

Proposition 6.2.6. *Under the assumptions of Theorem 6.1.1 we have*

$$\sup_t \sqrt{N} \left| \widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(t) - \widehat{F}_{\Delta_1 L_N}(t) \right| \xrightarrow{p} 0 \quad \text{as } \frac{N^2}{M} \rightarrow 0 \quad \text{and } N \rightarrow \infty.$$

Proof: Let $\Delta_n^{(M)} \equiv \Delta_1 L_n - \Delta_1 \widehat{L}_n^{(M)}$. Then

$$\begin{aligned} & \sup_t \sqrt{N} \left| \widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(t) - \widehat{F}_{\Delta_1 L_N}(t) \right| \\ &= \sup_t \sqrt{N} \left| \frac{1}{N} \sum_{n=1}^N \mathbf{I}\{\Delta_1 \widehat{L}_n^{(M)} \leq t\} - \frac{1}{N} \sum_{n=1}^N \mathbf{I}\{\Delta_1 L_n \leq t\} \right| \\ &\leq \sup_t \frac{1}{\sqrt{N}} \sum_{n=1}^N \left| \mathbf{I}\{\Delta_1 \widehat{L}_n^{(M)} - \Delta_1 L_n + \Delta_1 L_n \leq t\} - \mathbf{I}\{\Delta_1 L_n \leq t\} \right| \\ &= \sup_t \frac{1}{\sqrt{N}} \sum_{n=1}^N \left| \mathbf{I}\{\Delta_1 L_n \leq t + \Delta_n^{(M)}\} - \mathbf{I}\{\Delta_1 L_n \leq t\} \right| \\ &\leq \sup_t \frac{1}{\sqrt{N}} \sum_{n=1}^N \mathbf{I}\{t - |\Delta_n^{(M)}| \leq \Delta_1 L_n \leq t + |\Delta_n^{(M)}|\} \\ &= \sup_t \frac{1}{\sqrt{N}} \sum_{n=1}^N \mathbf{I}\{-|\Delta_n^{(M)}| \leq \Delta_1 L_n - t \leq |\Delta_n^{(M)}|\} \\ &= \sup_t \frac{1}{\sqrt{N}} \sum_{n=1}^N \mathbf{I}\{|\Delta_n^{(M)}| \geq |\Delta_1 L_n - t|\} \\ &\leq \sup_t \frac{1}{\sqrt{N}} \sum_{n=1}^N \mathbf{I}\{\delta_N^{(M)} \geq |\Delta_1 L_n - t|\} \end{aligned}$$

where $\delta_N^{(M)} \equiv \sup_{n=1, \dots, N} |\Delta_n^{(M)}|$.

Now, for given $\epsilon > 0$ and (a_N) any sequence that converges to 0, we calculate:

$$\begin{aligned}
& P\left(\sup_t \sqrt{N} \left| \widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(t) - \widehat{F}_{\Delta_1 L_N}(t) \right| > \epsilon\right) \\
& \leq P\left(\sup_t \frac{1}{\sqrt{N}} \sum_{n=1}^N \mathbf{I}\left\{\delta_N^{(M)} \geq |\Delta_1 L_n - t|\right\} > \epsilon\right) \\
& = P\left(\sup_t \sum_{n=1}^N \mathbf{I}\left\{\delta_N^{(M)} \geq |\Delta_1 L_n - t|\right\} > \sqrt{N}\epsilon\right) \\
& = P\left(\sup_t \left[\sum_{n=1}^N \mathbf{I}\left\{\delta_N^{(M)} \geq |\Delta_1 L_n - t|, |\Delta_1 L_n - t| > a_N\right\} \right. \right. \\
& \quad \left. \left. + \sum_{n=1}^N \mathbf{I}\left\{\delta_N^{(M)} \geq |\Delta_1 L_n - t|, |\Delta_1 L_n - t| \leq a_N\right\} \right] > \sqrt{N}\epsilon\right) \\
& \leq P\left(\sup_t \sum_{n=1}^N \mathbf{I}\left\{\delta_N^{(M)} \geq |\Delta_1 L_n - t|, |\Delta_1 L_n - t| > a_N\right\} > \sqrt{N}\frac{\epsilon}{2}\right) \\
& \quad + P\left(\sup_t \sum_{n=1}^N \mathbf{I}\left\{\delta_N^{(M)} \geq |\Delta_1 L_n - t|, |\Delta_1 L_n - t| \leq a_N\right\} > \sqrt{N}\frac{\epsilon}{2}\right) \\
& \leq P\left(N\mathbf{I}\left\{\delta_N^{(M)} > a_N\right\} > \sqrt{N}\frac{\epsilon}{2}\right) + P\left(\sup_t \sum_{n=1}^N \mathbf{I}\{|\Delta_1 L_n - t| \leq a_N\} > \sqrt{N}\frac{\epsilon}{2}\right) \\
& \leq P\left(\delta_N^{(M)} > a_N\right) + P\left(\sup_t \sqrt{N} \frac{1}{N} \sum_{n=1}^N \mathbf{I}\{t - a_N \leq \Delta_1 L_n \leq t + a_N\} > \frac{\epsilon}{2}\right).
\end{aligned}$$

Now, by Lemma 3.3.5 (i) (see (6.1.4)),

$$\begin{aligned}
P\left(\delta_N^{(M)} > a_N\right) & = P\left(\bigcup_{n=1}^N (\Delta_n^{(M)} > a_N)\right) \leq \frac{\sum_{n=1}^N \mathbb{E} \left| \Delta_n^{(M)} \right|^2}{a_N^2} \leq \frac{N \left\| \Delta_n^{(M)} \right\|_{L_2}^2}{a_N^2} \\
& = O\left(\frac{N}{a_N^2 M}\right)
\end{aligned} \tag{6.2.2}$$

and

$$\begin{aligned}
& P\left(\sup_t \sqrt{N} \frac{1}{N} \sum_{n=1}^N \mathbf{I}\{t - a_N \leq \Delta_1 L_n \leq t + a_N\} > \frac{\epsilon}{2}\right) \\
& = P\left(\sup_t \sqrt{N} \left(\widehat{F}_{\Delta_1 L_N}(t + a_N) - \widehat{F}_{\Delta_1 L_N}(t - a_N) \right) > \frac{\epsilon}{2}\right)
\end{aligned}$$

$$\begin{aligned} &\leq P\left(\sup_t \sqrt{N} |W_{\Delta_1 L_N}(t + a_N) - W_{\Delta_1 L_N}(t - a_N)| > \frac{\epsilon}{4}\right) \\ &\quad + \mathbf{I}\left\{\sup_t \sqrt{N} (F(t + a_N) - F(t - a_N)) > \frac{\epsilon}{4}\right\} \end{aligned}$$

where

$$W_{\Delta_1 L_N}(t) = \frac{1}{N} \sum_{n=1}^N (\mathbf{I}\{\Delta_1 L_n \leq t\} - F(t)).$$

We have

$$P\left(\sup_t \sqrt{N} (W_{\Delta_1 L_N}(t + a_N) - W_{\Delta_1 L_N}(t - a_N)) > \frac{\epsilon}{4}\right) \leq P\left(\omega\left(\sqrt{N}W_{\Delta_1 L_N}, 2a_N\right) > \frac{\epsilon}{4}\right) \rightarrow 0,$$

where

$$\omega(x, \delta) = \sup_{|t-s| < \delta} |x(t) - x(s)|, \quad 0 < \delta \leq 1$$

is the modulus of continuity of the process x . By the argument on pages 198 – 199 of Billingsley [6], (22.13) holds since F is continuous, that is for every positive ϵ and η there exist δ , $0 < \delta < 1$, and n_0 such that for all $n \geq n_0$

$$P\left(\omega\left(\sqrt{N}W_{\Delta_1 L_N}, \delta\right) > \epsilon\right) \leq \eta.$$

Now let (a_N) be any positive sequence that decreasing to 0. Choose N_1 , such that $N_1 \geq N_0$ and $a_N \leq \delta$ for all $N \geq N_1$. Then for all $N \geq N_1$, $\omega\left(\sqrt{N}W_{\Delta_1 L_N}, a_N\right) \leq \omega\left(\sqrt{N}W_{\Delta_1 L_N}, \delta\right)$, and so

$$P\left(\omega\left(\sqrt{N}W_{\Delta_1 L_N}, a_N\right) > \epsilon\right) \leq P\left(\omega\left(\sqrt{N}W_{\Delta_1 L_N}, \delta\right) > \epsilon\right) \leq \eta.$$

Hence $\omega\left(\sqrt{N}W_{\Delta_1 L_N}, 2a_N\right) \xrightarrow{p} 0$ as $a_N \rightarrow 0$.

Since F is continuous with bounded derivative then by the Mean Value Theorem,

$$\sup_t \sqrt{N} (F(t + a_N) - F(t - a_N)) \leq 2 \sup_t \sqrt{N} F'(t) a_N \leq 2K \sqrt{N} a_N. \quad (6.2.3)$$

Hence, if $\frac{N^2}{M} \rightarrow 0$, the result follows from (6.2.2) and (6.2.3) by choosing $a_N = \frac{\epsilon}{\sqrt{N}}$, since ϵ is arbitrary. \square

We can now complete the proof of Theorem 6.1.1:

$$\begin{aligned} \sqrt{N} \left(\widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(t) - F \left(t, \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right) &= \sqrt{N} \left(\widehat{F}_{\Delta_1 L_N}(t) - F \left(t, \overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2} \right) \right) \\ &+ \sqrt{N} \left(F \left(t, \overline{\Delta_1 L^*}, \widehat{\eta}_L^{*2} \right) - F \left(t, \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right) + \sqrt{N} \left(\widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(t) - \widehat{F}_{\Delta_1 L_N}(t) \right). \end{aligned}$$

The result follows by Propositions 6.2.1, 6.2.3 and 6.2.6. \square

6.3 Test Statistics

Under the assumptions of Theorem 6.1.1, one can construct a distribution-free test statistic for the composite hypothesis $H_0 : F \in \mathcal{A}$ where \mathcal{A} is a class of distributions determined by μ and η^2 .

We have shown in Theorem 6.1.1 that:

$$\sqrt{N} \left(\widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(\cdot) - F \left(\cdot ; \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right) \xrightarrow{d} B_F(\cdot) \text{ as } \frac{N^{2+\delta}}{M} \rightarrow 0 \text{ and } N \rightarrow \infty.$$

where $B_F(t)$ is a Gaussian process with zero mean and covariance given by:

$$\text{Cov}(B_F(s), B_F(t)) = E[B_F(s)B_F(t)] = F(s) \wedge F(t) - F(s)F(t).$$

We note that Brownian motion satisfies the hypotheses of Theorem 6.1.1. If $\mu > \eta$, the gamma process does as well.

6.3.1 Kolmogorov-Smirnov Test

Define the Kolmogorov-Smirnov (K-S) test statistics for a test of $H_0 : L(1) \sim F(\cdot; \mu, \eta^2)$ based on the estimated increments $\Delta_1 \widehat{L}_N^{(M)}$ and the bootstrapped estimators $\overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2}$ as

$$KS_{\Delta_1 \widehat{L}_N^{(M)}} \equiv \sup_t \sqrt{N} \left| \widehat{F}_{\Delta_1 \widehat{L}_N^{(M)}}(t) - F \left(t; \overline{\Delta_1 \widehat{L}^{*(M)}}, \widehat{\eta}^{*2} \right) \right|.$$

By Theorem 6.1.1 we have that

$$KS_{\Delta_1 \widehat{L}_N^{(M)}} \xrightarrow{\mathcal{D}} \sup_t |B_F(t)| \stackrel{d}{=} \sup_{0 \leq u \leq 1} |B(u)| \quad \text{as } \frac{N^{2+\delta}}{M} \rightarrow 0 \quad \text{and } N \rightarrow \infty,$$

where $B(\cdot)$ is the standard Brownian bridge on $[0, 1]$. This is also the limiting distribution as $N \rightarrow \infty$ of

$$KS_{\Delta_1 L_N} = \sup_t \sqrt{N} \left| \widehat{F}_{\Delta_1 L_N}(\cdot) - F(\cdot; \mu, \eta^2) \right|,$$

the Kolmogorov-Smirnov test statistic based on the true unobserved increments $\Delta_1 L_n$, $n = 1, \dots, N$ and specified values for μ and η^2 .

First, we consider the Brownian motion driven CAR(1) process. Table 6.1 compares the empirical levels $\widehat{\alpha}_{\Delta_1 B}$ and $\widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$ for tests based on $KS_{\Delta_1 B_N}$ and $KS_{\Delta_1 \widehat{B}_N^{(M)}}$, respectively, with true parameters $\mu = 1$, $\eta = 1$, over $R = 400$ simulations, with nominal level 0.05. We consider various values of the parameter a .

	$N = 50, M = 2500$		$N = 100, M = 500$	
	$\widehat{\alpha}_{\Delta_1 B}$	$\widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 B}$	$\widehat{\alpha}_{\Delta_1 \widehat{B}^{(M)}}$
a=0.1	0.0450	0.0575	0.0400	0.0575
a=0.9	0.0550	0.0525	0.0350	0.0475
a=10	0.0450	0.0400	0.0375	0.0725
a=100	0.0725	0.0300	0.0450	0.0400
a=1000	0.0500	0.0750	0.0500	0.0475

Table 6.1: We fix $\{\sigma = 1, \mu = 0, \eta = 1, R = 400\}$

The test performs well in comparison to the usual K-S test and does not seem overly sensitive to the relation between N and M .

Next, we consider gamma driven CAR(1) processes with $\mu = 10$ and $\eta^2 = 1$. Similar to Table 6.1, in Table 6.2 we computed empirical levels $\widehat{\alpha}_{\Delta_1 G}, \widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$ for Kolmogorov-Smirnov tests based on the true increments and parameters and the estimated increments and bootstrapped estimators respectively, over $R = 400$ simulations, with nominal level 0.05. We consider various values of the parameter a .

	$N = 100, M = 500$		$N = 50, M = 2500$	
	$\widehat{\alpha}_{\Delta_1 G}$	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$	$\widehat{\alpha}_{\Delta_1 G}$	$\widehat{\alpha}_{\Delta_1 \widehat{G}^{(M)}}$
$a = 0.1$	0.0375	0.0600	0.0375	0.0475
$a = 0.9$	0.0400	0.0425	0.0550	0.0425
$a = 10$	0.0450	0.0450	0.0375	0.0475
$a = 100$	0.0450	0.0250	0.0450	0.0400
$a = 1000$	0.0425	0.0625	0.0425	0.0250

Table 6.2: We fix $\{\sigma = 1, \mu = 10, \eta = 1, R = 400\}$

Similar to the Brownian motion case, the test performs well in comparison to the usual K-S test and does not seem overly sensitive to the relation between N and M .

Although, in the case of the gamma driven CAR(1) process, the estimators $(\overline{\Delta_1 L})$ and $(\widehat{\eta}_L^2)$ are not based on maximum likelihood methods, and the Burke and Gombay technique has not been yet justified in this case, our simulations give surprisingly accurate empirical levels. This needs further investigation.

In Table 6.3, we illustrate the behavior of the test statistic for different values of μ , with $\eta = 1$ and $a = 0.9$. The test behaves well for all values of $\mu > 1$. However, the differentiability assumption (6.1.1) is violated for $\mu \leq 1$, and that is reflected in the empirical rejection rates for $\mu = 0.5$ and 1.0.

μ	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	μ	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$	a	$\hat{\alpha}_{\Delta_1 G}$	$\hat{\alpha}_{\Delta_1 \hat{G}^{(M)}}$
0.5	0.0500	0.3400	3.5	0.0350	0.0525	6.5	0.0600	0.0675
1.0	0.0425	0.1025	4.0	0.0550	0.0475	7.0	0.0400	0.0625
1.5	0.0425	0.0625	4.5	0.0300	0.0500	7.5	0.0475	0.0575
2.0	0.0425	0.0550	5.0	0.0375	0.0500	8.0	0.0425	0.0400
2.5	0.0675	0.0575	5.5	0.0300	0.0425	8.5	0.0450	0.0425
3.0	0.0275	0.0450	6.0	0.0400	0.0500	9.0	0.0475	0.0275

Table 6.3: $N = 100, M = 500, R = 400, a = 0.9, \eta = 1, \sigma = 1$

Next we consider power by computing empirical rejection rate $\hat{\beta}_{\Delta_1 G}$ and $\hat{\beta}_{\Delta_1 \hat{G}^{(M)}}$ for the tests based on $KS_{\Delta_1 G_n}$ and $KS_{\Delta_1 \hat{G}_n^{(M)}}$, respectively, for different values of the parameter a . First, we simulate the gamma driven CAR(1) process and test the claim that the driving process is Brownian motion. Empirical rejection rates are given in Table 6.4 and the power function is illustrated in Figure 6.1.

a	$\hat{\beta}_{\Delta_1 G}$	$\hat{\beta}_{\Delta_1 \hat{G}^{(M)}}$	a	$\hat{\beta}_{\Delta_1 G}$	$\hat{\beta}_{\Delta_1 \hat{G}^{(M)}}$	a	$\hat{\beta}_{\Delta_1 G}$	$\hat{\beta}_{\Delta_1 \hat{G}^{(M)}}$
0.5	1.00	0.94	3.5	1.00	0.88	6.5	1.00	0.93
1.0	1.00	0.85	4.0	1.00	0.87	7.0	1.00	0.94
1.5	1.00	0.93	4.5	1.00	0.92	7.5	1.00	0.86
2.0	1.00	0.94	5.0	1.00	0.91	8.0	1.00	0.96
2.5	1.00	0.96	5.5	1.00	0.88	8.5	1.00	0.90
3.0	1.00	0.92	6.0	1.00	0.91	9.0	1.00	0.87

Table 6.4: $N = 100, M = 500, R = 100, \mu = 1, \eta = 1, \sigma = 1$

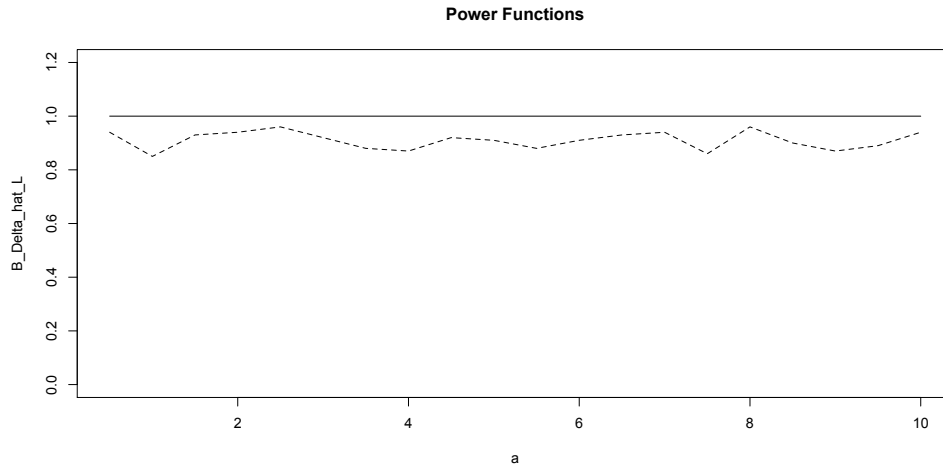


Figure 6.1: The power functions for the test based on $KS_{\Delta_1 G_n}$ is the solid line and the one based on $KS_{\Delta_1 \hat{G}_n^{(M)}}$ is the dashed line .

Next, we simulate the Brownian motion driven CAR(1) process and test the claim that the driving process is a gamma process. Empirical rejection rates are given in Table 6.5. In this case, we have a 100% rejection rate, reflecting how well we approximate Brownian increments.

a	$\hat{\beta}_{\Delta_1 B}$	$\hat{\beta}_{\Delta_1 \hat{B}^{(M)}}$	a	$\hat{\beta}_{\Delta_1 B}$	$\hat{\beta}_{\Delta_1 \hat{B}^{(M)}}$	a	$\hat{\beta}_{\Delta_1 B}$	$\hat{\beta}_{\Delta_1 \hat{B}^{(M)}}$
0.5	1.00	1.00	3.5	1.00	1.00	6.5	1.00	1.00
1.0	1.00	1.00	4.0	1.00	1.00	7.0	1.00	1.00
1.5	1.00	1.00	4.5	1.00	1.00	7.5	1.00	1.00
2.0	1.00	1.00	5.0	1.00	1.00	8.0	1.00	1.00
2.5	1.00	1.00	5.5	1.00	1.00	8.5	1.00	1.00
3.0	1.00	1.00	6.0	1.00	1.00	9.0	1.00	1.00

Table 6.5: $N = 100, M = 500, R = 100, \mu = 10, \eta = 1, \sigma = 1$

Comment: As observed above, the Burke and Gombay technique has not been justified in the case in which the driving process is gamma. Nonetheless, Tables 6.2, 6.3 and 6.4 illustrate that remarkably accurate empirical levels are achieved, as well as good power. Indeed, the proof given in Section 6.2 demonstrates that Theorem 6.2 of [16] is valid for the gamma process, and although the KS test statistic is no longer nonparametric our simulations indicate that the error introduced in the variance is likely small (see Figure 6.2). This needs to be investigated further.

Moreover, we conjecture that an alternative approach proposed in [39] is valid for a wide range of Lévy driving processes, including both Brownian motion and the gamma process. In this case, a parametric bootstrap technique is used. Assuming that one observes L , a random sample X_1^*, \dots, X_N^* is simulated from the distribution $F(\cdot, \overline{\Delta_1 L}, \hat{\eta}_L^2)$. Next, let $\overline{X^*}$ and $\hat{\eta}_{X^*}^2$ denote the bootstrapped estimators of μ and η^2 . Then if \hat{F}_{X^*} is the empirical distribution based on the bootstrapped sample, both

$$\sqrt{N} \left(\hat{F}_{\Delta_1 L}(\cdot) - F(\cdot, \overline{\Delta_1 L}, \hat{\eta}_L^2) \right)$$

and

$$\sqrt{N} \left(\hat{F}_{X^*}(\cdot) - F(\cdot, \overline{X^*}, \hat{\eta}_{X^*}^2) \right)$$

have the same limiting distribution. We conjecture that in light of Propositions 6.2.3 and 6.2.6, the same would be true using recovered values of L . In this case, repeated bootstrap samples would allow us to calculate an appropriate critical value for our test statistic. However, we note that although this technique would be more widely applicable than that of Burke and Gombay, it is computationally far more intensive.

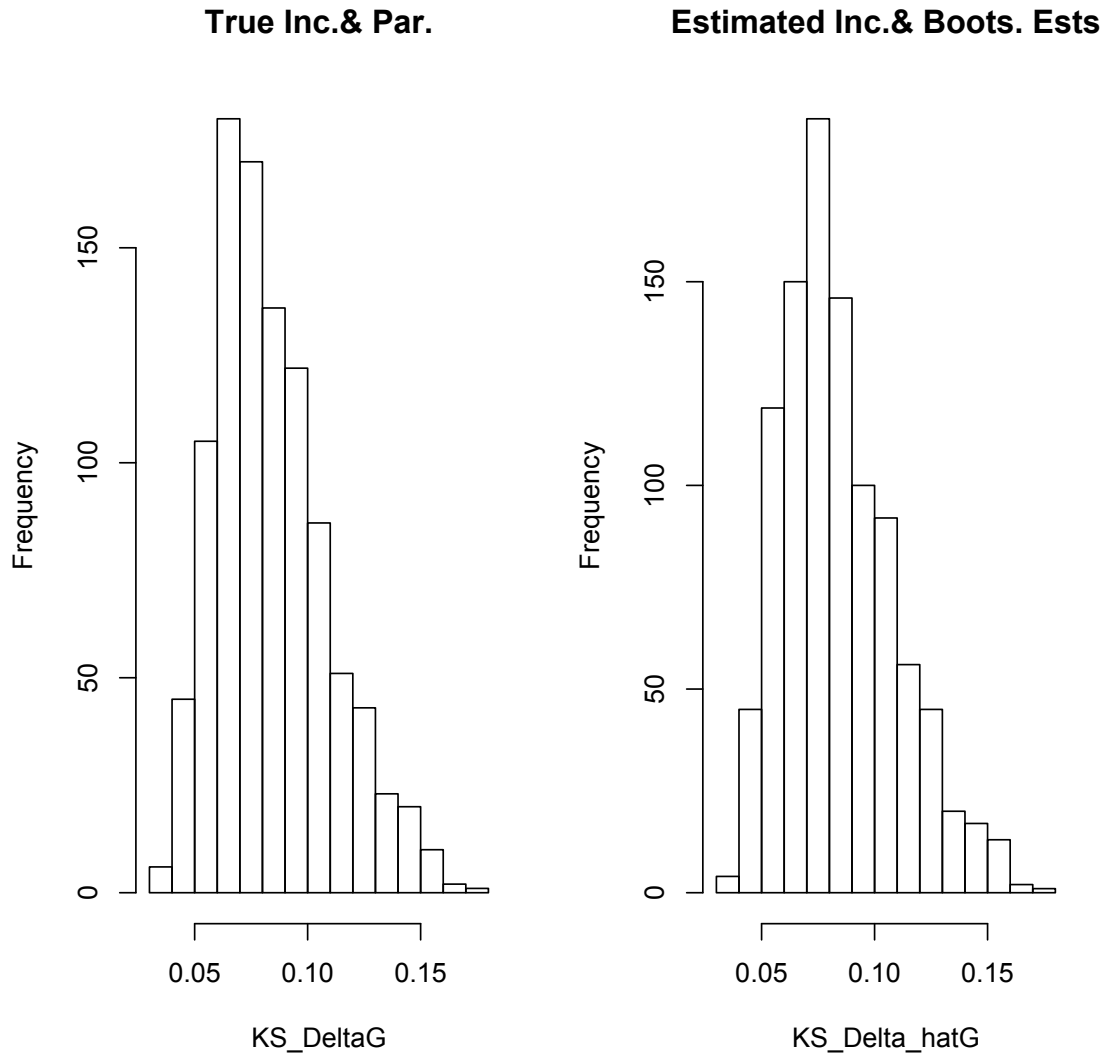


Figure 6.2: Histograms of the Kolmogorov-Smirnov test values based on the true gamma process increments and parameters and the one that based on the estimated gamma process increments and bootstrapped estimators respectively. $R = 1000$, $N = 100$, $M = 500$, $\eta = 1$ and $\mu = 10$.

Chapter 7

Brief Summary and Extensions

In summary, if the CAR(1) process is observed at discrete times, we have constructed test statistics to test, initially, the Lévy assumption of uncorrelated increments of the driving process. Then, if the Lévy assumption is not rejected, we have created a more precise test statistic to examine which candidate of the Lévy family could be the driving process.

7.1 Future Work and Extensions

We proved the asymptotic normality of the test statistics for uncorrelatedness of the increments of the driving process. It would be interesting to consider other tests for independence of the increments.

We suspect that techniques from [25] and [2] can be used to improve the rate condition $N^2/M \rightarrow 0$ in Theorem 6.1.1.

We will verify that the technique of [39] is valid using recovered values of L , and compare its performance to that of Burke and Gombay.

Another interesting topic for immediate extension is to further investigate the behavior of the test statistic defined in Chapter 6 if we replace the parameter a by the estimator defined in Chapter 5.

An extension to the Lévy-driven CARMA(2,1) process is an important problem. This model was employed by Todorov and Tauchen [41] and Todorov [40] to represent stochastic volatility in the Deutsche Mark/U.S Dollar daily exchange rate.

Subsequently, a natural question is whether our results are extendable to general CARMA(p, q) or at least CARMA($p, 1$) models. A close inspection of our proofs show that they rely on the inversion formula and second order properties of Y . Hence, our results should be extendable, but this will require detailed analysis.

Last but not least, in [24], Garcia et al. used an α -stable Levy process as a driving process for the CARMA(2,1) process to model spot prices from the Singapore New Electricity Market. Accordingly, it would be desirable to extend some of our results to CARMA(p, q) models driven by a stable Lévy process. Clearly our techniques will no longer be appropriate since they rely on an L_2 approximation of the noise L by the process Y , and so a different approach will be needed.

Bibliography

- [1] Ibrahim Abdelrazeq, B. Gail Ivanoff, and Rafa Kulik. Model verification for Lévy-driven Ornstein-Uhlenbeck processes. *Electronic Journal of Statistics*, 8(1):1029–1062, 2014.
- [2] Philippe Barbe, Christian Genest, Kilani Ghoudi, and Bruno Rémillard. On Kendall’s process. *J. Multivariate Anal.*, 58(2):197–229, 1996.
- [3] Ole E. Barndorff-Nielsen, Elisa Nicolato, and Neil Shephard. Some recent developments in stochastic volatility modelling. *Quant. Finance*, 2(1):11–23, 2002. Special issue on volatility modelling.
- [4] Ole E. Barndorff-Nielsen and Neil Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 63(2):167–241, 2001.
- [5] Fred Espen Benth, Martin Groth, and Rodwell Kufakunesu. Valuing volatility and variance swaps for a non-Gaussian Ornstein-Uhlenbeck stochastic volatility model. *Appl. Math. Finance*, 14(4):347–363, 2007.
- [6] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley & Sons Inc., New York, 1968.
- [7] Patrick Billingsley. *Probability and Measure*. Wiley Series in Probability and

- Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.
- [8] Fischer Black and Myron S Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–54, May-June 1973.
- [9] P. J. Brockwell. Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.*, 53(1):113–124, 2001. Nonlinear non-Gaussian models and related filtering methods (Tokyo, 2000).
- [10] Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1991.
- [11] Peter J. Brockwell, Richard A. Davis, and Yu Yang. Estimation for nonnegative Lévy-driven Ornstein-Uhlenbeck processes. *J. Appl. Probab.*, 44(4):977–989, 2007.
- [12] Peter J. Brockwell, Richard A. Davis, and Yu Yang. Estimation for non-negative Lévy-driven CARMA processes. *J. Bus. Econom. Statist.*, 29(2):250–259, 2011.
- [13] Peter J. Brockwell and Tina Marquardt. Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statist. Sinica*, 15(2):477–494, 2005.
- [14] Peter J. Brockwell and Eckhard Schlemm. Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations. *J. Multivariate Anal.*, 115:217–251, 2013.
- [15] M. D. Burke, M. Csörgő, S. Csörgő, and P. Révész. Approximations of the empirical process when parameters are estimated. *Ann. Probab.*, 7(5):790–810, 1979.

- [16] Murray D. Burke and Edit Gombay. On goodness-of-fit and the bootstrap. *Statist. Probab. Lett.*, 6(5):287–293, 1988.
- [17] Jessica Cariboni and Wim Schoutens. Jumps in intensity models: investigating the performance of Ornstein-Uhlenbeck processes in credit risk modeling. *Metrika*, 69(2-3):173–198, 2009.
- [18] Jean-Francois Coeurjolly. *dvfBm: Discrete variations of a fractional Brownian motion*, 2009. R package version 1.0.
- [19] S. Cohen and A. Lindner. A central limit theorem for the sample autocorrelations of a lévy driven continuous time moving average process. *Journal of Statistical Planning and Inference*, 143:1295–1306, 2013.
- [20] J. L. Doob. The elementary Gaussian processes. *Ann. Math. Statistics*, 15:229–282, 1944.
- [21] A. Ja. Dorogovcev. The consistency of an estimate of a parameter of a stochastic differential equation. *Theory Prob. Math. Statist.*, 10:73–82, 1976.
- [22] J. Durbin. *Distribution theory for tests based on the sample distribution function*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1973. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 9.
- [23] Vincenzo Ferrazzano and Florian Fuchs. Noise recovery for Lévy-driven CARMA processes and high-frequency behaviour of approximating Riemann sums. *Electron. J. Stat.*, 7:533–561, electronic only, 2013.
- [24] Isabel Garcia, Claudia Klueppelberg, and Gernot Mueller. Estimation of stable carma models with an application to electricity spot prices. *Statistical Modelling*, 11(5):447–470, 2011.

- [25] Kilani Ghoudi and Bruno Remillard. Empirical processes based on pseudo-observations. In *Asymptotic methods in probability and statistics (Ottawa, ON, 1997)*, pages 171–197. North-Holland, Amsterdam, 1998.
- [26] M. Grigoriu. *Applied non-Gaussian processes*. PTR Prentice. Hall, 1995.
- [27] H. Jeffreys and B. S. Jeffreys. *Methods of Mathematical Physics*. Cambridge University Press, England, third edition, 1988.
- [28] G. Jongbloed, F. H. van der Meulen, and A. W. van der Vaart. Nonparametric inference for Lévy-driven ornstein-uhlenbeck processes. *Bernoulli*, 11(5):759–791, 2005.
- [29] R. A. Kasonga. The consistency of a nonlinear least squares estimator from diffusion processes. *Stochastic Process. Appl.*, 30(2):263–275, 1988.
- [30] O Kella and W Stadje. On hitting times for compound Poisson dams with exponential jumps and linear release rate. *Journal of Applied Probability*, 38(3):781–786, SEP 2001.
- [31] H Larralde. A first passage time distribution for a discrete version of the Ornstein-Uhlenbeck process. *Journal of Physics A-Mathematical and General*, 37(12):3759–3767, MAR 26 2004.
- [32] A. Le Breton. On continuous and discrete sampling for parameter estimation in diffusion type processes. *Mathematical Programming Study*, 5:124–144, 1976.
- [33] Boris Leblanc and Olivier Scaillet. Path dependent options on yields in the affine term structure model. *Finance Stoch.*, 2(4):349–367, 1998.
- [34] G Louchard, C Kenyon, and R Schott. Data structures’ maxima. *Siam Journal on Computing*, 26(4):1006–1042, AUG 1997.

- [35] Tina Marquardt and Robert Stelzer. Multivariate CARMA processes. *Stochastic Process. Appl.*, 117(1):96–120, 2007.
- [36] Robert C. Merton. Theory of Rational Option Pricing. *Bell Journal of Economics*, 4(1):141–183, Spring 1973.
- [37] B. L. S. Prakasa Rao. On the exponential rate of convergence of the least squares estimator in the nonlinear regression model with Gaussian errors. *Statist. Probab. Lett.*, 2(3):139–142, 1984.
- [38] Ken-iti Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.
- [39] Winfried Stute, Wenceslao Gonzáles Manteiga, and Manuel Presedo Quindimil. Bootstrap based goodness-of-fit tests. *Metrika*, 40(3-4):243–256, 1993.
- [40] Viktor Todorov. Econometric analysis of jump-driven stochastic volatility models. *J. Econometrics*, 160(1):12–21, 2011.
- [41] Viktor Todorov and George Tauchen. Simulation methods for Lévy-driven continuous-time autoregressive moving average (CARMA) stochastic volatility models. *J. Bus. Econom. Statist.*, 24(4):455–469, 2006.
- [42] Pham Dinh Tuan. Estimation of parameters of a continuous time Gaussian stationary process with rational spectral density. *Biometrika*, 64(2):385–399, 1977.
- [43] G. E. Uhlenbeck and L. S. Ornstein. On the theory of brownian motion. *Phys. Rev.*, 36:823–841, 1930.

-
- [44] Shibin Zhang, Zhengyan Lin, and Xinsheng Zhang. A least squares estimator for Lévy-driven moving averages based on discrete time observations. *Communications in Statistics - Theory and Methods*, (to appear).