



Estimador Local de Whittle para Procesos de Volatilidad Estocástica Localmente Estacionarios de Corta y Larga Memoria

por

Cristian Vásquez Escobar

Tesis presentada al Departamento de Estadística de la
Pontificia Universidad Católica de Chile para optar al
grado de Doctor en Estadística

Enero, 2016

Director de Tesis: **Dr. Wilfredo Palma**

©Copyright por **Cristian Vásquez Escobar**, 2016

FACULTAD DE MATEMÁTICAS
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

INFORME DE APROBACIÓN
TESIS DE DOCTORADO

Se informa a la Facultad de Matemáticas que la Tesis de Doctorado presentada por el candidato

Cristian Vásquez Escobar

ha sido aprobada por la Comisión de Evaluación de la Tesis como requisito para optar al grado de Doctor en Estadística, en el examen de Defensa de Tesis rendido el día 7 de Enero de 2016.

Director de Tesis

Wilfredo Palma

Comisión de Evaluación de la Tesis

Miguel de Carvalho

Miguel Uribe

Natalia Bahamonde

Ricardo Olea

FACULTAD DE MATEMÁTICAS
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Fecha: 7 de Enero 2016

Autor :**Cristian Vásquez Escobar**
Título :**Estimador Local de Whittle para Procesos de Volatilidad
Estocásticos Localmente Estacionarios de Corta y
Larga Memoria**
Departamento :**Estadística**
Grado :**Doctor**
Convocación :**Enero 2016**

Se le concede permiso para hacer circular y copiar, con propósitos no comerciales,
el título ante dicho para los requerimientos de individuos y/o instituciones.

Firma del Autor

EL AUTOR SE RESERVA LOS DERECHOS DE OTRAS PUBLICACIONES, Y NI LA TESIS NI EXTRACTOS
EXTENSOS DE ELLA, PUEDEN SER IMPRESOS O REPRODUCIDOS SIN EL PERMISO ESCRITO DEL AUTOR.
EL AUTOR ATESTIGUA QUE EL PERMISO SE HA OBTENIDO PARA EL USO DE CUALQUIER MATERIAL
COPYRIGHTED QUE APAREZCA EN ESTA TESIS (CON EXCEPCIÓN DE LOS BREVES EXTRACTOS QUE
REQUIEREN SOLAMENTE EL RECONOCIMIENTO APROPIADO EN LA ESCRITURA DEL ESTUDIANTE) Y
QUE TODO USO ESTÉ RECONOCIDO CLARAMENTE.

AGRADECIMIENTOS

Primero que todo quiero agradecer a mi profesor guía Wilfredo Palma, por el tiempo dedicado, la confianza entregada y el apoyo para poder realizar este proyecto de tesis. Quisiera agradecer a las profesoras Anita Araneda y Claudia Matus, por las cartas de recomendación para la postulación al programa de Doctorado en Estadística y a la Beca de Conicyt. También quiero dar las gracias a Ricardo Olea, por el tiempo y el material de apoyo entregado.

Quiero aprovechar esta instancia para agradecer a mis grandes camaradas, Carlos Piutrin, Ignacia Vicuña y Felipe Barrientos, porque ellos fueron un gran apoyo en lo personal para poder finalizar mi doctorado y muchas ocasiones aportaron acertadas ideas en mi trabajo. Durante mis años en la PUC hice grandes amigos a los que quiero agradecer, Danilo Garrido, Mauricio Tejo, Alvaro Ferrada, José Quinlan, Jaime Arrué, Rodrigo Rubio y Enrique Guerra. La mención honrosa se la llevan mis super secre Luisa Rivas y Valeria Leiva, ellas siempre me apoyaron.

También deseo agradecer al Departamento de Estadística de la Pontificia Universidad Católica de Chile, por aceptarme dos veces en sus programas de Postgrado, primero en el Magister en Estadística y luego en el Doctorado.

Finalmente quiero dar las gracias a mi madre Mercedes Escobar, mi padre Luis Vásquez, mi hermano Felipe, mi tía preferida Ana Vásquez, mi abuelita Sylvia Ibáñez que esta en el cielo y a mi gran amor Luz Tisandié, por todo el cariño y apoyo entregado estos años.

Mi permanencia en el plan de doctorado fue posible gracias a la Beca de Doctorado
CONICYT No. 21110386.

Índice

Lista de Figuras	VII
Lista de Tablas	IX
1. Preliminares	1
1.1. Introducción	1
1.2. Retornos	4
1.3. Hechos Estilizados Sobre los Retornos	6
1.4. Volatilidad	7
1.5. Modelos de Volatilidad Estocástica	8
1.6. Distribuciones especiales	12
2. Procesos de Volatilidad Estocástica Localmente Estacionarios de Corta Memoria	14
2.1. Procesos Localmente Estacionarios de Corta Memoria	15
2.2. Modelo de Volatilidad Estocástica de Corta Memoria	20
2.2.1. Propiedades del proceso de retornos localmente estacionario $\{r_{t,T}\}$	22
2.3. Estimación de modelos paramétricos a procesos de volatilidad localmente estacionarios	23
2.3.1. Selección de Modelos	27

2.4. Experimento de Monte Carlo	29
3. Procesos de Volatilidad Estocástica Localmente Estacionarios de Larga Memoria	41
3.1. Procesos Localmente Estacionarios de Larga Memoria	42
3.2. Modelos de Volatilidad Estocástica de Larga Memoria	46
3.2.1. Propiedades del proceso de retornos localmente estacionario $\{r_{t,T}\}$	48
3.2.2. Estimación de modelos paramétricos a procesos de volatilidad localmente estacionarios de larga memoria	50
3.2.3. Selección de Modelos	53
3.3. Experimento de Monte Carlo	54
4. Ilustración	68
4.1. Índice Bursátil S&P500	68
5. Conclusiones y Trabajos Futuros	84
A. Apéndice	88
B. Apéndice	92
B.1. Proposiciones	92
B.2. Demostración de Teoremas	111
C. Apéndice	116
C.1. Proposiciones	116
C.2. Demostración de Teoremas	147

Referencias

Índice de figuras

2.1. Densidad espectral del proceso LS-AR donde $\beta_0 = 0.4$, $\beta_1 = 0.4$ y $\sigma = 1$. . .	19
2.2. Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARMA con $T = 512$	36
2.3. Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARMA con $T = 1024$	38
3.1. Densidad espectral del proceso LS-FN donde $\beta_0 = 0.3$, $\beta_1 = 0.6$ y $\beta_2 = 0.5$. 45	
3.2. Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARFIMA con $T = 512$	61
3.3. Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARFIMA con $T = 1024$	63
4.1. (a) Índice S&P500 considerando valor de cierre ajustado (b) Retornos logarítmicos de la serie (a).	69
4.2. (a) Función de autocorrelación muestral (b) Función de autocorrelación parcial muestral.	70
4.3. Diagnóstico del modelo.	72
4.4. Gráfico del p valor del test de “Ljung Box”.	73
4.5. (a) Cuadrado de los residuos (b) Función de autocorrelación muestral. . .	74

4.6. Función de autocorrelación (a) intervalo del 26/11/1951 al 27/05/1963 (b) intervalo del 11/05/1964 al 25/09/1978 (c) intervalo del 10/09/1979 al 26/12/1995.	75
4.7. Gráfico de la serie transformada.	76
4.8. Función de autocorrelación (a) intervalo del 26/11/1951 al 27/05/1963 (b) intervalo del 11/05/1964 al 25/09/1978 (c) intervalo del 10/09/1979 al 26/12/1995.	77
4.9. Estimación de los parámetros por bloque.	78
4.10. Gráfico de la evolución de los parámetros estimados.	80
4.11. Serie $\log z_{i,T}^2$ y predicción a un paso.	81
4.12. (a) Gráfico de los Residuos (b) Función de autocorrelación muestral (c) Función de autocorrelación parcial muestral (d) Gráfico de los valores p del test de “Ljung Box”.	82

Índice de tablas

2.1. Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 128$ y salto $S = 20$. . .	31
2.2. Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 128$ y salto $S = 30$. . .	31
2.3. Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 128$ y salto $S = 40$. . .	32
2.4. Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 256$ y salto $S = 20$. . .	33
2.5. Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 256$ y salto $S = 30$. . .	33
2.6. Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 256$ y salto $S = 40$. . .	34
2.7. Estimación de Whittle: tamaño muestral $T = 512$, tamaño de ventana $N = 120$ y salto $S = 30$	37
2.8. Estimación de Whittle: tamaño muestral $T = 1024$, tamaño de ventana $N =$ 210 y salto $S = 45$	39
3.1. Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 128$ y salto $S = 20$. . .	56
3.2. Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 128$ y salto $S = 30$. . .	56
3.3. Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 128$ y salto $S = 40$. . .	57
3.4. Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 256$ y salto $S = 20$. . .	58
3.5. Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 256$ y salto $S = 30$. . .	58
3.6. Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 256$ y salto $S = 40$. . .	59
3.7. Estimación de Whittle: tamaño muestral $T = 512$, tamaño de ventana $N = 130$ y salto $S = 40$	62

3.8. Estimación de Whittle: tamaño muestral $T = 1024$, tamaño de ventana $N =$ 230 y salto $S = 60$	64
3.9. Estimación de Whittle: tamaño muestral $T = 1024$, tamaño de ventana $N =$ 230 y salto $S = 60$ con $\sigma_\epsilon = \pi/\sqrt{2} \approx 2.221$	66
4.1. Test de Ljung Box sobre la serie de retornos	71
4.2. Parámetros estimados	71
4.3. Parámetros estimados	79

RESUMEN

En la presente tesis, se realiza una extensión de los procesos de volatilidad estocástica introducidos por Taylor (1980) y Bollerslev (1986), utilizados para modelar retornos financieros. En este trabajo se considera que la secuencia latente no observada que genera la volatilidad, es un procesos localmente estacionario, de corta y larga memoria.

Se introducen dos nuevas clase de procesos, se desarrollan ejemplos ilustrativos, y se estudian algunas propiedades de momentos. Se presenta un mecanismo de estimación para los parámetros asociados a esta clase de modelos, donde se propone utilizar el estimador por bloques de Whittle. Se presentan propiedades asintóticas de los estimadores para muestras de tamaño grande, como la consistencia y el teorema del límite central.

Para investigar el desempeño en muestras finitas, se realiza un experimento de Monte Carlo en diferentes escenarios y se realiza una aplicación de la metodología.

ABSTRACT

In this thesis we develop an extension of the class of stochastic volatility processes, which are widely used for modeling financial returns. In this work, we consider that the latent sequence generating the volatility model corresponds to a locally stationary process with either short or long memory. A definition of this class of processes is presented and some fundamental properties are studied. A spectral-based block Whittle likelihood technique is proposed to estimate the model parameters. Asymptotic properties of these estimators are established, including consistency and a central limit

theorem. Furthermore, the finite sample performances of the proposed estimates are studied by means of Monte Carlo simulations.

Capítulo 1

Preliminares

1.1. Introducción

Existe una gran variedad de modelos estadísticos para describir el comportamiento de los retornos de una serie de tiempo financiera. En la literatura se pueden distinguir dos grandes familias; los modelos de volatilidad determinísticos, que corresponde a la familia ARCH-GARCH introducidos por Engle (1982) y Bollerslev (1986), donde los retornos pasados de la serie determinan de manera casi segura la volatilidad del proceso y los modelos de volatilidad estocástica (también conocidos como modelos de volatilidad exponencial) introducidos por Taylor (1986), donde la volatilidad del proceso de retorno de una serie está caracterizada de manera exponencial por una serie latente no observada, la cual es independiente de las innovaciones que generan los retornos. Los modelos estadísticos mencionados anteriormente suponen que la volatilidad misma es un proceso estacionario.

Desde la década de los 90, se han desarrollado varios y extensos trabajos para tratar la falta de estacionariedad de las series temporales, donde se destaca el desarrollo de los procesos localmente estacionarios introducidos por Dahlhaus (1996, 1997). A partir de éstos, se desarrolló una gran línea de investigación, donde sobresalen los es-

tudios realizados por Dahlhaus (2000), Sachs and von Sachs and MacGibbon (2000), Dahlhaus and Polonik (2006, 2009), Beran (2009), Palma and Olea (2010) y Roueff and von Sachs (2011), entre otros. Todos ellos, asumen que la función de transferencia varía a través del tiempo lo suficientemente suave como para ser aproximada en una vecindad por la función de transferencia de procesos estacionarios, de corta y larga memoria.

Los modelos estadísticos comúnmente utilizados para describir la volatilidad de una serie de retornos financieros, suponen una estructura estacionaria. Sin embargo, existen dos grandes motivos para realizar una modificación a ello. El primero, se debe a que las series financieras son muy extensas (por la cantidad de información recolectada), lo que dificulta utilizar el supuesto de estacionariedad. Segundo, algunos investigadores sugieren que la misma volatilidad puede ser generada por la falta de estacionariedad. Considerando lo anterior, se ha desarrollado una nueva línea de investigación, donde se incorpora una estructura localmente estacionaria en la volatilidad de los retornos, es ahí donde se destaca un trabajo importante, el artículo Dahlhaus et al. (2006), donde se desarrollan los modelos ARCH localmente estacionarios (considerado como una extensión del modelo ARCH estacionario comúnmente utilizado). En ese trabajo se deriva un estimador no paramétrico y algunas propiedades asintóticas. Existen otros trabajos importantes relacionados donde se pueden destacar Fryzlewicz (2005), Fryzlewicz et al. (2006) y Fryzlewicz and Subba Rao (2011) donde se estudian los procesos ARCH y ARCH(∞) localmente estacionario (time varying). Para más detalles, el artículo de Van Bellegem (2012) presenta una visión general de los procesos ARCH localmente estacionarios.

Los trabajos mencionados, corresponde a extensiones de los modelos de volatili-

dad determinísticos (familia ARCH-GARCH), los cuales son considerados como herramientas muy útiles para realizar ajustes a series de retornos financieros. Sin embargo, un defecto de estos modelos (ARCH localmente estacionarios), son las restricciones paramétricas impuestas para garantizar la existencia de momentos y de las propiedades asintóticas.

Como alternativa, se propone es utilizar modelos de volatilidad estocástica, los cuales relajan varios supuestos para garantizar la existencia del proceso y momentos y para realizar inferencias. En esta tesis se presenta una extensión de los modelos de volatilidad estocástica, donde el proceso latente (no observado) que genera la volatilidad corresponde a una proceso localmente estacionario, el cual puede ser de corta o larga memoria (modelos ARMA y ARFIMA de variación en el tiempo). Para la estimación de los parámetros, se propone utilizar la función de verosimilitud por bloques de Whittle. Bajo este mecanismo, se demuestra que el estimador es asintóticamente consistente y con distribución normal (para procesos de corta y larga memoria). Este trabajo proporciona herramientas estadísticas para modelar y hacer inferencia sobre una clase de procesos no estacionarios en la volatilidad. Una gran ventaja computacional de utilizar la función de verosimilitud por bloques de Whittle, es la velocidad de los cálculos para estimación, dado que solo requiere del periodograma, el cual se puede determinar por medio de la transformada rápida de Fourier.

La tesis se encuentra organizada de la siguiente manera. En el presente capítulo, se entregan preliminares sobre los retornos de una serie financiera, se explica el concepto de volatilidad y el estado del arte sobre técnicas estadísticas para ajustar series financieras desde una perspectiva estacionaria. En el capítulo 2, se introducen los modelos de localmente estacionarios de corta memoria utilizando la definición y

supuestos de Dahlhaus (1997) con algunos ejemplos. Además se presentan los modelos de volatilidad estocástica localmente estacionarios de corta memoria, algunas propiedades de momento como la función de covarianza y algunos detalles sobre la densidad espectral de variación en el tiempo. Se propone un método de estimación basado en la extensión local del estimador de Whittle, y se establece la consistencia y la normalidad asintótica de estos estimadores. Adicionalmente se presenta un estudio de simulación de Montecarlo para ilustrar el desempeño de las estimaciones en muestras finitas. En el capítulo 3 se realiza un idem del capítulo 2 pero para procesos de larga memoria Gaussianos, como estos procesos son más complejo y dado que no se cumplen los supuestos presentados por Dahlhaus (1997), se utilizan los establecidos por Palma and Olea (2010). El capítulo 4 presenta una ilustración con un caso real. Para finalizar el trabajo se presentan las conclusiones en el capítulo 5. Las demostraciones las puede encontrar en el apéndice. En el apéndice A usted encuentran lemas generales para ambos procesos. El apéndice B contiene las demostraciones para la versión corta memoria y en el apéndice C para larga memoria.

1.2. Retornos

Uno de los objetivos en Finanzas es poder medir el riesgo de un portafolio de activos financieros. La manera usual de cuantificar ello, es mediante las variaciones que se producen por los precios de los activos en diferentes períodos de tiempo. Utilizando la nomenclatura estándar, denotaremos a $\{P_t\}$ como el precio de un activo en el instante t . La variación en el precio entre el instante $t - 1$ y t relativo al instante $t - 1$ se conoce como *retorno simple* y esta definido por

$$R_t = \frac{\Delta P_t}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (1.1)$$

Usualmente R_t se expresa en %, dado que representa una variación porcentual sobre el precio relativo a un período de tiempo (día, mes, o año). Una definición frecuentemente utilizada es el *retorno continuamente compuesto* o *log-retorno* que corresponde a

$$r_t = \log \frac{P_t}{P_{t-1}} = \log(1 + R_t) = \log P_t - \log P_{t-1}. \quad (1.2)$$

Esta definición será la utilizada durante la tesis, y la denotaremos directamente como retorno. Es útil trabajar con los retornos, ya que por definición son libres de escala y tienen propiedades estadísticas interesantes. El retorno es una de las principales medidas para cuantificar el riesgo de un activo o de un portafolio, ahí radica la importancia de proponer modelos adecuados y que se adapten a los hechos.

Note que para valores pequeños de u , $\log(1 + u) \approx u$, de esta forma se puede verificar que los *retornos simples* R_t y los *log-retornos* r_t serán valores muy similares.

Una vez establecida la definición de retorno, se puede construir el retorno a k períodos, el cual está dado por

$$r_t[k] = \log \frac{P_t}{P_{t-k}} = \log \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}} = \sum_{j=0}^{k-1} r_{t-j} \quad (1.3)$$

Un breve ejemplo, suponga que se quiere medir el retorno en un período de una semana, donde transcurrieron 5 días hábiles, entonces el *log retorno* esta dado por

$$r_t[5] = r_t + r_{t-1} + \cdots + r_{t-4},$$

para todo t .

1.3. Hechos Estilizados Sobre los Retornos

Los retornos financieros presentan algunas características bien particulares, atípicas a otras series de tiempo. Por ejemplo, si uno observa retornos diarios de cualquier activo, rara vez se presenta tendencia o estacionalidad, estas características se conocen en finanzas como hechos estilizados. Los principales hechos estilizados relativos a los retornos financieros pueden ser resumidos como sigue:

1. Los retornos son usualmente no correlacionados.
2. Los cuadrados de los retornos son autocorrelacionados, por lo general de decaimiento lento.
3. Las series de retornos presentan cluster de volatilidad a lo largo del tiempo (entendiendo por volatilidad a la varianza de los retornos).
4. La distribución incondicional de los retornos presenta colas más pesadas que la distribución normal.
5. La distribución incondicional de los retornos suele ser asimétrica. Si por el contrario, la evidencia empírica muestra una distribución aproximadamente simétrica, igual es leptocúrtica.
6. Algunas series de retornos son no lineales.

Por lo tanto un modelo apropiado para describir una serie de retorno debe considerar estos hechos.

1.4. Volatilidad

El objetivo de esta tesis, es modelar la volatilidad o varianza condicional de un retorno (bajo un enfoque localmente estacionario). Aunque no sea medida directamente, la volatilidad se manifiesta de diferentes maneras en una serie financiera.

Al momento de tratar la volatilidad, podemos utilizar tres enfoques propuestos en la literatura:

- (i) Una forma es comparar el precio de mercado de un activo con el precio modelado. Se obtiene una volatilidad implícita, que usualmente esta basada en la fórmula de Black-Scholes (si son opciones). El problema es que supone volatilidad constante.
- (ii) Una manera usual es modelar directamente la volatilidad de la serie de retornos utilizando familias ya estudiadas, como los modelos ARCH y modelos de volatilidad estocástica. Hoy en día, se dispone de una nueva familia que son los modelos ARCH(∞) localmente estacionarios, propuestos por Dahlhaus et al. (2006) (también conocidos como time varying ARCH).
- (iii) Una alternativa es modelar la volatilidad por medio de un promedio móvil de una función de los últimos k retornos. Este mecanismo es conocido como volatilidad histórica, usualmente se considera un promedio móvil de los retornos cuadrados o de los retornos absolutos para todo t , como

$$v_t = \left[\frac{1}{k} \sum_{j=0}^{k-1} |r_{t-j}|^p \right]^{\frac{1}{p}},$$

donde $p > 0$. Los casos usualmente utilizados son $p = 2$ y $p = 1$.

Las tres metodologías presentadas pueden entregar valores muy distintos. Independiente del enfoque utilizado, la volatilidad es una medida de variabilidad de activos y normalmente es muy difícil prever las variaciones de los precios.

La notación estándar utilizada en la literatura es la siguiente. Sea r_t una serie de retornos, defina

$$\mu_t = \mathbb{E}(r_t | \mathcal{F}_{t-1}), \quad (1.4)$$

$$\sigma_t = \mathbb{E}((r_t - \mu_t)^2 | \mathcal{F}_{t-1}), \quad (1.5)$$

como la media y la varianza condicional de r_t , dada la información pasada al instante t , $\mathcal{F}_{t-1} = \{r_{t-1}, r_{t-2}, \dots\}$.

Un modelo estándar para la volatilidad es de la siguiente forma

$$r_t = \mu_t + \sigma_t u_t, \quad (1.6)$$

donde u_t tiene con media cero, varianza unitaria y típicamente *i.i.d.* con distribución F . Toda la estructura se encuentra en el parámetro σ_t .

1.5. Modelos de Volatilidad Estocástica

En la literatura, existen diferentes familias de modelos estadísticos para ajustar la volatilidad, los más utilizados son la familia ARCH. Estos suponen que la varianza condicional es una función de los retornos pasados. Otra clase son los modelos de

volatilidad estocástica, propuestos por Taylor (1980, 1986), los cuales se basan en el hecho de que la volatilidad misma en el presente es generada por un proceso latente no observado y que depende de ella misma a través de un proceso estacionario, que es no correlacionado con los retornos pasados.

Definición 1.1. La serie de retornos $\{r_t\}$ sigue un modelo de volatilidad estocástica, si el proceso es definido por las siguientes ecuaciones

$$r_t = \sigma_t u_t, \quad (1.7)$$

$$\sigma_t = \sigma \exp(v_t/2), \quad (1.8)$$

donde $\{u_t\}$ es independiente e idénticamente distribuido con media cero y varianza uno, y $\{v_t\}$ es un proceso estacionario independiente de $\{u_t\}$.

Cuando $\{v_t\}$ es especificado por un proceso de corta memoria ARMA(p, q), entonces lo llamaremos *modelo de volatilidad estocástica de corta memoria*. Por otra parte, si $\{v_t\}$ es especificado por un proceso de larga memoria ARFIMA(p, d, q), lo denotaremos como *modelo de volatilidad estocástica de larga memoria*.

Note que este proceso se puede escribir como

$$\log(r_t^2) = \log(\sigma_t^2) + \log(u_t^2), \quad (1.9)$$

$$\log(\sigma_t^2) = \log(\sigma^2) + v_t. \quad (1.10)$$

Sea $Y_t = \log(r_t^2)$, $\mu = \log(\sigma^2) + \mathbb{E}[\log(u_t^2)]$ y $\epsilon_t = \log(u_t^2) - \mathbb{E}[\log(u_t^2)]$, entonces

$$Y_t = \mu + v_t + \epsilon_t. \quad (1.11)$$

Así, la transformación resultante $\{Y_t\}$ será una serie estacionaria más un ruido aditivo (additive noise).

Dada la ortogonalidad de $\{v_t\}$ y $\{\epsilon_t\}$, la función de autocovarianza de $\{Y_t\}$ está dada por

$$\gamma_Y(k) = \gamma_v(k) + \sigma_\epsilon^2 \delta_0(k), \quad (1.12)$$

donde $\delta_0(k) = 1$ para $k = 0$ y $\delta_0(k) = 0$ en otro caso, σ_ϵ^2 es la varianza de ϵ_t y $\gamma_v(k)$ corresponde a la función de autocovarianza del proceso $\{v_t\}$. Por otro lado, la densidad espectral de $\{Y_t\}$, esta dada por

$$f_Y(\lambda) = f_v(\lambda) + \frac{\sigma_\epsilon^2}{2\pi}, \quad (1.13)$$

donde $f_v(\lambda)$ corresponde a la densidad espectral de $\{v_t\}$. Note que el vector de parámetros θ que caracteriza al proceso, esta especificado en la densidad espectral $f_Y(\lambda) = f_\theta(\lambda)$.

En la literatura, se pueden encontrar diferentes mecanismos para estimar los parámetros del modelos de volatilidad estocástica. Esto va a depender de que tipo de proceso está especificado para $\{v_t\}$. La mayoría de los mecanismos trabajan con la transformación $\{Y_t\}$. Cuando $\{v_t\}$ es definido mediante un proceso de corta memoria, hay autores que proponen aproximar el modelo con sistema de espacio estado Gaussiano (QML), como los trabajos de Ruiz (1994) y Harvey et al. (1994), también existe

otros autores como Kim et al. (1998), que trabajan directamente con la función de verosimilitud, pero tiene un costo computacional muy alto. Cuando $\{v_t\}$ es especificado como un proceso ARFIMA(p, d, q), usualmente se proponen estimadores del tipo log regresión (conocidos como GPH) como por ejemplo Deo and Hurvich (2001), también, hay varios autores que trabajan directamente con el estimador de verosimilitud de Whittle (QML) como Breidt et al. (1998), Hurvich and Ray (2003) y Zaffaroni (2009).

A modo de ilustración, se presentará el estimador de verosimilitud espectral de Whittle para el modelo de volatilidad estocástica, ya que es una herramienta útil independiente de la especificación de $\{v_t\}$, y además, porque los resultados de esta tesis son una extensión de los resultados asintóticos de este estimador en la versión localmente estacionaria (que se presentan en los capítulos 2 y 3).

Estimación: Suponga que se observa una trayectoria de n retornos $\{r_1, r_2, \dots, r_n\}$, el vector de parámetros θ puede ser estimado minimizando la verosimilitud espectral de Whittle $\mathcal{L}(\theta)$ definida como

$$\mathcal{L}_n(\theta) = \frac{2\pi}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} \left[\log f_\theta(\lambda_j) + \frac{I(\lambda_j)}{f_\theta(\lambda_j)} \right], \quad (1.14)$$

donde $\lambda_j \in \{2\pi j/n; j = 1, 2, \dots, \lfloor n/2 \rfloor\}$ son las frecuencias de Fourier, $I(\lambda_j)$ es el periodograma de $\{Y_t\}$ (transformación de $\{r_t\}$) y $f_\theta(\lambda_j)$ corresponde a la densidad espectral del proceso $\{Y_t = \log(r_t^2)\}$ definida en (1.13).

Sea $\hat{\theta}_n$ el valor que minimiza $\mathcal{L}_n(\theta)$ sobre el espacio paramétrico Θ . Este estimador satisface los siguientes teoremas bajo ciertas condiciones de regularidad (revisar detalles en citas bibliográficas).

Teorema 1.1. *Breidt et al. (1998).* Asuma que el vector de parámetro θ es un elemento de un espacio paramétrico compacto Θ y asuma que $f_{\theta_1} = f_{\theta_2}$ implica que $\theta_1 = \theta_2$. Sea θ_0 el verdadero valor del parámetro, entonces, $\hat{\theta}_n \rightarrow \theta_0$ casi seguramente cuando $n \rightarrow \infty$.

Teorema 1.2. *Zaffaroni (2009).* Bajo ciertas condiciones de regularidad, el estimador $\hat{\theta}_n$ satisface el teorema de límite central, es decir

$$\sqrt{n} (\hat{\theta}_n - \theta_0) \xrightarrow{D} N_k(0, M^{-1}VM^{-1}),$$

donde

$$\begin{aligned} M(\theta) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\partial \ln f_{\theta}(\lambda)}{\partial \theta} \right] \left[\frac{\partial \ln f_{\theta}(\lambda)}{\partial \theta} \right]' d\lambda, \\ V(\theta) &= \frac{1}{\pi} \int_{-\pi}^{\pi} \left[\frac{\partial \ln f_{\theta}(\lambda)}{\partial \theta} \right] \left[\frac{\partial \ln f_{\theta}(\lambda)}{\partial \theta} \right]' d\lambda, \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{f_{\theta}(\lambda)} \left[\frac{\partial \ln f_{\theta}(\lambda)}{\partial \theta} \right] \frac{1}{f_{\theta}(\mu)} \left[\frac{\partial \ln f_{\theta}(\mu)}{\partial \theta} \right]' g_4(-\lambda, \mu, -\mu) d\lambda d\mu, \end{aligned}$$

g_4 corresponde al cumulante de 4to orden de $\{Y_t\}$ y k representa el número de parámetros involucrados en la densidad espectral $f_Y(\lambda)$, las matrices $M = M(\theta_0)$ y $V = V(\theta_0)$.

1.6. Distribuciones especiales

Usualmente para $\{u_t\}$ se propone una distribución normal estándar o t -Student. Nelson (1991) propone una distribución de errores generalizada ("generalized error distribution", GED). Se dice que la variable aleatoria X sigue una GED, con media cero y varianza uno, si su densidad esta dada por

$$f(x) = \frac{\nu \exp \left[-\left(\frac{1}{2}\right)|x/\lambda|^\nu \right]}{\lambda 2^{(\nu+1)/\nu} \Gamma(1/\nu)},$$

donde $x \in \mathbb{R}$ y

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

El parámetro $\nu > 0$ indica el comportamiento de las colas de la distribución. Si $\nu = 2$ obtenemos la distribución normal estándar, si $\nu = 1$ obtenemos la distribución doble exponencial

$$f(x) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|}.$$

Cuando $\nu < 2$ la distribución tiene colas más pesadas que la normal y cuando $\nu > 2$ tiene colas más livianas que la normal.

Fernández and Steel (1998) proponen para $\{u_t\}$ el uso de una distribución t -Student asimétrica, cuya densidad esta dada por

$$\begin{aligned} f(x|\iota, \nu) &= \frac{2}{\iota + 1/\iota} \left[g(\iota(sx + m)|\nu) I_{(-\infty, 0)}(x + m/s) \right] \\ &\quad + \frac{2}{\iota + 1/\iota} \left[g((sx + m)/\iota|\nu) I_{(0, \infty)}(x + m/s) \right], \end{aligned}$$

donde $g(\cdot|\nu)$ denota una distribución t -Student con ν grados de libertad,

$$\begin{aligned} m &= \frac{\Gamma((\nu + 1)/2) \sqrt{\nu - 2}}{\sqrt{\pi} \Gamma(\nu/2)} (\iota - 1/\iota), \\ s &= \sqrt{(\iota^2 + 1/\iota^2 - 1) - m^2} \end{aligned}$$

siendo ι el parámetro de asimetría, con $\iota = 1$ la distribución t -Student tradicional.

Capítulo 2

Procesos de Volatilidad Estocástica Localmente Estacionarios de Corta Memoria

En este capítulo se introduce una clase de modelos de volatilidad estocástica de variación en el tiempo, para esto, primero se debe revisar la definición de procesos localmente estacionarios realizada por Dahlhaus (1997). Luego se presenta la extensión de los procesos de volatilidad estocástica estacionarios a su versión localmente estacionarios y se estudian propiedades de momento.

En la literatura actual sobre los procesos de volatilidad localmente estacionarios, se encuentran los trabajos de Dahlhaus et al. (2006), Fryzlewicz (2005), Fryzlewicz et al. (2006) y Fryzlewicz and Subba Rao (2011), que están orientados a modelar la volatilidad de manera determinística. El motivo de nuestra propuesta, es que relajamos considerablemente las restricciones sobre el espacio paramétrico de los artículos mencionados.

2.1. Procesos Localmente Estacionarios de Corta Memoria

Definición 2.1. Se considera a una secuencia de un proceso estocástico $X_{t,T}$ ($t = 1, \dots, T$) localmente estacionario con función de transferencia A^0 y tendencia μ , si existe una representación

$$X_{t,T} = \mu(t/T) + \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^0(\lambda) d\xi(\lambda), \quad (2.1)$$

donde

(i) $\xi(\lambda)$ es un proceso estocástico en $[-\pi, \pi]$ con $\overline{\xi(\lambda)} = \xi(-\lambda)$ y

$$\text{cum}\{d\xi(\lambda_1), \dots, d\xi(\lambda_k)\} = \eta \left(\sum_{j=1}^k \lambda_j \right) g_k(\lambda_1, \dots, \lambda_{k-1}) d\lambda_1 \dots d\lambda_k,$$

donde $\text{cum}\{\dots\}$ denota el k -ésimo cumulante, $g_1 = 0$, $g_2(\lambda) = 1$, $g_k(\lambda_1, \dots, \lambda_{k-1}) \leq \text{const}_k$ para todo k y $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ es la extensión de la función delta de Dirac.

(ii) Existe una constante K y una función 2π periódica $A : [0, 1] \times \mathbb{R} \rightarrow \mathbb{C}$ con $A(u, -\lambda) = \overline{A(u, \lambda)}$, tal que

$$\sup_{t,\lambda} \left| A_{t,T}^0(\lambda) - A \left(\frac{t}{T}, \lambda \right) \right| \leq KT^{-1},$$

para todo T . Aquí se asume $A(u, \lambda)$ y $\mu(u)$ continuos en u .

A continuación se presentan algunos ejemplos de procesos localmente estacionarios:

Ejemplo 2.1. (i) Suponga que $\{Y_t\}$ es un proceso estacionario con representación espectral

$$Y_t = \int_{-\pi}^{\pi} \exp(i\lambda t) A(\lambda) d\xi(\lambda),$$

aquí $A(\lambda)$ es la función de transferencia, $\mu : [0, 1] \rightarrow \mathbb{R}$ y $\sigma : [0, 1] \rightarrow \mathbb{R}^+$ ambas continua. Entonces

$$X_{t,T} = \mu(t/T) + \sigma(t/T)Y_t$$

es un proceso localmente estacionario con $A_{t,T}^0(\lambda) = A(t/T, \lambda) = \sigma(t/T)A(\lambda)$.

(ii) Suponga que $\{\varepsilon_t\}$ es una secuencia *i.i.d.* y

$$X_{t,T} = \sum_{j=0}^{\infty} a_j \left(\frac{t}{T} \right) \varepsilon_{t-j},$$

con la restricción $\sum_{j=0}^{\infty} |a_j(u)| < \infty$ para todo $u \in [0, 1]$. Entonces $X_{t,T}$ es localmente estacionario con función de transferencia

$$A_{t,T}^0(\lambda) = A(t/T, \lambda) = \sum_{j=0}^{\infty} a_j(t/T) \exp(-\lambda j).$$

(iii) Considere el siguiente proceso

$$X_{t,T} = \phi \left(\frac{t}{T} \right) X_{t-1,T} + \sigma \left(\frac{t}{T} \right) \varepsilon_t, \quad t \in \{1, \dots, T\}$$

donde $\{\varepsilon_t\}$ una secuencia independiente idénticamente distribuida de media cero y varianza uno. Asumimos que $\sigma(u)$ y $\phi(u)$ son continuos en \mathbb{R} con $\sigma(u) = \sigma(0)$, $\phi(u) = \phi(0)$ para $u < 0$ y $\sigma(u) = \sigma(1)$, $\phi(u) = \phi(1)$ para $u > 1$, las restricciones para los parámetros son $|\phi(u)| < 1$ y $\sigma(u) > 0$ para todo $u \in [0, 1]$, también los parámetros son diferenciables para $u \in (0, 1)$ con derivada acotada.

Dado que la secuencia $\{\varepsilon_t\}$ es *i.i.d.*, el teorema de representación de Cramer garantiza que

$$\varepsilon_t = \int_{-\pi}^{\pi} (2\pi)^{-1/2} \exp(i\lambda t) d\xi(\lambda), \quad \forall t.$$

Aquí $\{\varepsilon_t\}$ satisface la definición 2.1 de manera trivial. Utilizando la expresión anterior, se puede verificar que $X_{t,T}$ satisface la relación

$$X_{t,T} = \mu(t/T) + \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^0(\lambda) d\xi(\lambda),$$

con

$$A_{t,T}^0(\lambda) = \frac{1}{\sqrt{2\pi}} \sum_{\ell=0}^{\infty} \left[\prod_{j=0}^{\ell-1} \phi\left(\frac{t-j}{T}\right) \right] \sigma\left(\frac{t-\ell}{T}\right) \exp(-i\lambda\ell).$$

Una vez que se conoce $A_{t,T}^0(\lambda)$, se puede mostrar que $A(u, \lambda)$ definido por

$$A(u, \lambda) = \frac{\sigma(u)}{\sqrt{2\pi}} \left(1 - \phi(u) \exp(-i\lambda) \right)^{-1},$$

satisface (ii) en la definición 2.1, es decir

$$\sup_{t,\lambda} \left| A_{t,T}^0(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq KT^{-1}.$$

Definición 2.2. (Dahlhaus (1996)). Considere un procesos localmente estacionario $X_{t,T}$ con función de transferencia A^0 , y suponga que existe una función $A(u, \lambda)$ que satisface la condición (ii) de la definición 2.1, entonces la densidad espectral de variación en el tiempo está definida por

$$f(u, \lambda) := |A(u, \lambda)|^2.$$

Esta densidad, corresponde al límite en media cuadrática de la densidad espectral de Wigner-Ville, para ver detalles de la demostración puede consultar el teorema 2.2 de Dahlhaus (1996), y si desea detalles de la densidad espectral de Wigner-Ville puede consultar Martin and Flandrin (1985).

Ejemplo 2.2. Sea $X_{t,T}$ una secuencia autoregresiva con coeficientes de variación en el tiempo, es decir, es solución de la ecuación

$$\sum_{j=0}^p a_j \left(\frac{t}{T} \right) X_{t-j,T} = \sigma \left(\frac{t}{T} \right) \varepsilon_{t-j},$$

donde $\{\varepsilon_t\}$ son *iid* con media cero y varianza 1, y $a_0(u) \equiv 1$. Entonces, si $X_{t,T}$ es un proceso locamente estacionario, por el teorema 2.3 de Dahlhaus (1996) se tiene que la densidad espectral para $X_{t,T}$ será

$$f(u, \lambda) = \frac{\sigma^2(u)}{2\pi} \left| \sum_{j=0}^p a_j(u) \exp(i\lambda j) \right|^{-2}, \quad \text{para todo } (u, \lambda) \in [0, 1] \times [-\pi, \pi]. \quad (2.2)$$

Se define el k -ésimo coeficiente de Fourier de la densidad espectral

$$c(u, k) := \int_{-\pi}^{\pi} f(u, \lambda) \exp(i\lambda k) d\lambda \quad (2.3)$$

como la función de covarianza local de largo k a tiempo u . Por lo anterior, se tiene

$$\begin{aligned} \text{Cov}(X_{[uT],T}, X_{[uT]+k,T}) &= \int_{-\pi}^{\pi} A_{[uT],T}^0(\lambda) A_{[uT]+k,T}^0(-\lambda) \exp(i\lambda k) d\lambda \\ &= c(u, k) + \mathcal{O}(T^{-1}). \end{aligned} \quad (2.4)$$

Ejemplo 2.3. Sea $X_{t,T} = \phi(t/T)X_{t-1,T} + \sigma\varepsilon_t$, donde $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$, y considere que $\phi(u) = \beta_0 + \beta_1 u$, con $|\phi(u)| < 1$ para todo $u \in [0, 1]$. Entonces

$$\text{Cov}(X_{[uT],T}, X_{[uT]+k,T}) = \sigma^2 \frac{(\beta_0 + \beta_1 u)^{|k|}}{1 - (\beta_0 + \beta_1 u)^2} + \mathcal{O}(T^{-1})$$

y la densidad espectral límite es

$$f(u, \lambda) = \frac{\sigma^2}{2\pi} \frac{1}{1 - 2(\beta_0 + \beta_1 u) \cos(\lambda) + (\beta_0 + \beta_1 u)^2}.$$

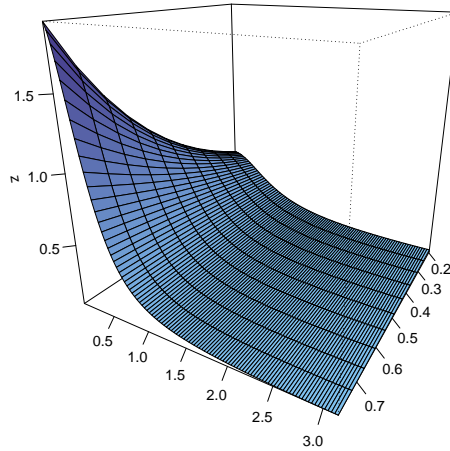


Figura 2.1: Densidad espectral del proceso LS-AR donde $\beta_0 = 0.4$, $\beta_1 = 0.4$ y $\sigma = 1$.

2.2. Modelo de Volatilidad Estocástica de Corta Memoria

Suponga que se dispone de una secuencia de retornos $\{r_t\}$ de largo T , aquí se reescalará el índice de la serie en función del largo de la trayectoria, es decir, se denotará a la secuencia de retornos como $\{r_{t,T}; t \in \{1, \dots, T\}\}$.

Diremos que los retornos $\{r_{t,T}\}$ siguen un modelo de volatilidad estocástica localmente estacionario de corta memoria, si es definido por las siguientes ecuaciones

$$r_{t,T} = \sigma_{t,T} u_t, \quad (2.5)$$

$$\sigma_{t,T} = \sigma \exp(\nu_{t,T}/2), \quad (2.6)$$

donde $\sigma > 0$, $\{u_t\}$ es una secuencia independiente idénticamente distribuida con media cero y varianza uno, y $\{\nu_{t,T}\}$ es un proceso localmente estacionario independiente de $\{u_t\}$ (modelo de dos choques aleatorios).

Utilizando las ecuaciones (2.5) y (2.6), se tiene que

$$\log(r_{t,T}^2) = \log(\sigma_{t,T}^2) + \log(u_t^2),$$

$$\log(\sigma_{t,T}^2) = \log(\sigma^2) + \nu_{t,T}.$$

Sea $Y_{t,T} = \log(r_{t,T}^2)$, $\mu = \log(\sigma^2) + \mathbb{E}(\log(u_t^2))$, $\epsilon_t = \log(u_t^2) - \mathbb{E}(\log(u_t^2))$. Entonces,

$$Y_{t,T} = \mu + \nu_{t,T} + \epsilon_t. \quad (2.7)$$

El proceso $\{Y_{t,T}\}$ corresponde a la suma de un modelo localmente estacionario más una secuencia de ruido blanco.

Dado que $\{\epsilon_t\}$ y $\{\nu_{t,T}\}$ son ortogonales, la función de autocovarianza del proceso $\{Y_{t,T}\}$ es

$$\text{Cov}(Y_{t,T}, Y_{t+k,T}) = \text{Cov}(\nu_{t,T}, \nu_{t+k,T}) + \sigma_\epsilon^2 \delta_0(k), \quad \forall k \in \mathbb{N} \cup \{0\},$$

donde $\delta_0(k) = 1$ si $k = 0$ y $\delta_0(k) = 0$ si $k > 0$.

Usando la relación entre la covarianza de variación en el tiempo y la local, es fácil mostrar que

$$\text{Cov}(Y_{[uT],T}, Y_{[uT]+k,T}) = c_\nu(u, k) + \sigma_\epsilon^2 \delta_0(k) + \mathcal{O}(T^{-1}), \quad \forall u \in [0, 1].$$

Para determinar la densidad espectral del proceso $\{Y_{t,T}\}$ haremos uso del lema A.5,

$$f_Y(u, \lambda) = f_\nu(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi}, \quad (2.8)$$

donde $f_\nu(u, \lambda)$ es la densidad espectral límite de $\{\nu_{t,T}\}$ y σ_ϵ^2 la varianza de $\{\epsilon_t\}$.

Ejemplo 2.4.

(i) Suponga que $\{\nu_{t,T}\}$ es una secuencia autoregresiva de medias móviles con coeficientes de variación en el tiempo, es decir, es solución de

$$\sum_{j=0}^p a_j \left(\frac{t}{T}\right) \nu_{t-j,T} = \sum_{j=0}^q b_j \left(\frac{t}{T}\right) \sigma \left(\frac{t-j}{T}\right) \varepsilon_{t-j},$$

con $\{\varepsilon_t\}$ son *iid* con media cero y varianza 1, y $a_0(u) \equiv b_0(u) \equiv 1$. Si el proceso $\{\nu_{t,T}\}$ satisface la definición 2.1, entonces la densidad espectral de variación en el tiempo de $\{Y_{t,T}\}$ será

$$f_Y(u, \lambda) = \frac{\sigma(u)^2}{2\pi} \frac{\left| \sum_{j=0}^q b_j(u) \exp(-i\lambda j) \right|^2}{\left| \sum_{j=0}^p a_j(u) \exp(-i\lambda j) \right|^2} + \frac{\sigma_\epsilon^2}{2\pi}.$$

- (ii) Considere que $\{\nu_{t,T}\}$ es una clase de procesos localmente estacionarios, que está representado por una expansión de medias móviles infinita

$$\nu_{t,T} = \sigma \left(\frac{t}{T} \right) \sum_{j=0}^{\infty} \psi_j \left(\frac{t}{T} \right) \varepsilon_{t-j},$$

donde $\{\varepsilon_t\}$ es un proceso independiente idénticamente distribuido de media cero y varianza uno. En este ejemplo la función de transferencia del proceso está dada por $A_{t,T}^0(\lambda) = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \psi_j\left(\frac{t}{T}\right) e^{-i\lambda j} = A\left(\frac{t}{T}, \lambda\right)$, entonces la densidad espectral de $Y_{t,T}$ es

$$f_Y(u, \lambda) = \frac{\sigma(u)^2}{2\pi} \left| \sum_{j=0}^{\infty} \psi_j(u) \exp(-i\lambda j) \right|^2 + \frac{\sigma_\varepsilon^2}{2\pi}.$$

2.2.1. Propiedades del proceso de retornos localmente estacionario $\{r_{t,T}\}$

A continuación se presentan algunas propiedades que son relevantes para el estudio de retornos financieros:

- (i) $\mathbb{E}(r_{t,T}) = \mathbb{E}(\sigma_{t,T})\mathbb{E}(u_t) = 0$, dado que $\sigma_{t,T}$ y u_t son independientes.

- (ii) $\mathbb{V}(r_{t,T}) = \mathbb{E}(\sigma_{t,T}^2 u_t^2) = \mathbb{E}(\sigma_{t,T}^2)\mathbb{E}(u_t^2) = \mathbb{E}(\sigma_{t,T}^2)$.

- (iii) Para $k > 0$ entero se tiene que

$$\text{Cov}(r_{t,T}, r_{t+k,T}) = \mathbb{E}(u_t)\mathbb{E}(u_{t+k})\mathbb{E}(\exp\{\nu_{t,T} + \nu_{t+k,T}\}) = 0.$$

- (iii) Suponiendo que el proceso $\{\nu_{t,T}\}$ tiene media cero, entonces

$$\begin{aligned} \mathbb{E}(r_{t,T}^2) &= \sigma^2 \mathbb{E}(\exp\{\nu_{t,T}\}) \\ &= \sigma^2 \exp \left\{ \frac{1}{2} \mathbb{V}(\nu_{t,T}) \right\}. \end{aligned}$$

Por ejemplo, considere el modelo LS-AR($\phi(u)$) con $\sigma_\varepsilon^2 = 1$,

$$\mathbb{E}(r_{t,T}^2) = \sigma^2 \exp \left\{ \frac{1}{2} \frac{1}{1 - \phi(t/T)^2} + \mathcal{O}(T^{-1}) \right\}.$$

(iv) Si suponemos que el proceso $\{\nu_{t,T}\}$ es Gaussiano, la función de covarianza de los $\{r_{t,T}^2\}$ está dada por

$$\begin{aligned} \text{Cov}(r_{t,T}^2, r_{t+k,T}^2) &= \sigma^2 [\mathbb{E}(\sigma_{t,T}^2 \sigma_{t+k,T}^2) - \mathbb{E}(\sigma_{t,T}^2) \mathbb{E}(\sigma_{t+k,T}^2)] \\ &= \sigma^2 [\mathbb{E}(\exp\{\nu_{t,T} + \nu_{t+k,T}\}) - \mathbb{E}(\exp\{\nu_{t,T}\}) \mathbb{E}(\exp\{\nu_{t+k,T}\})] \\ &= \sigma^2 \left[\exp \left\{ \text{Cov}(\nu_{t,T}, \nu_{t+k,T}) + \frac{1}{2} \mathbb{V}(\nu_{t,T}) + \frac{1}{2} \mathbb{V}(\nu_{t+k,T}) \right\} \right. \\ &\quad \left. - \exp \left\{ \frac{1}{2} \mathbb{V}(\nu_{t,T}) + \frac{1}{2} \mathbb{V}(\nu_{t+k,T}) \right\} \right] \\ &= \sigma^2 \exp \left\{ \frac{1}{2} (\mathbb{V}(\nu_{t,T}) + \mathbb{V}(\nu_{t+k,T})) \right\} \left[\exp \left\{ \text{Cov}(\nu_{t,T}, \nu_{t+k,T}) \right\} - 1 \right]. \end{aligned}$$

(v) Si el proceso $\{\nu_{t,T}\}$ es Gaussiano, la función de curtosis es

$$\mathbf{K}_{t,T} = \frac{\mathbb{E}(r_{t,T}^4)}{(\mathbb{E}(r_{t,T}^2))^2} = \frac{3\sigma^4 \exp\{2\mathbb{V}(\nu_{t,T})\}}{\sigma^4 \exp\{\mathbb{V}(\nu_{t,T})\}} = 3 \exp\{\mathbb{V}(\nu_{t,T})\} > 3.$$

Al momento de proponer un modelo estadístico para la volatilidad, se deben revisar estas propiedades de momento, y verificar que estas sean coherentes con los hechos estilizados sobre retornos financieros.

2.3. Estimación de modelos paramétricos a procesos de volatilidad localmente estacionarios

Sea $\theta \in \Theta$ el vector de parámetros del proceso $\{Y_{t,T}\}$ especificado en (2.7), donde el espacio paramétrico Θ es un subconjunto de un espacio Euclidiano finito dimensional.

Dada una muestra $\{Y_{1,T}, \dots, Y_{t,T}\}$, el estimador del vector de parámetros del proceso puede ser obtenido minimizando la función de log verosimilitud local de Whittle

$$\mathcal{L}_T(\theta) = \frac{1}{4\pi} \frac{1}{M} \int_{-\pi}^{\pi} \sum_{j=1}^M \left[\log f_{\theta}(u_j, \lambda) + \frac{I_N(u_j, \lambda)}{f_{\theta}(u_j, \lambda)} \right] d\lambda, \quad (2.9)$$

donde $f_{\theta}(u, \lambda)$ corresponde a la densidad espectral de variación en el tiempo del proceso $\{Y_{t,T}\}$ que está especificada por el parámetro θ , $I_N(u_j, \lambda) = \frac{|D_N(u, \lambda)|^2}{2\pi H_{2,N}(0)}$ es un periodograma con taper, con

$$D_N(u, \lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) Y_{[uT]-N/2+s+1, T} e^{-i\lambda s}, \quad H_{k,N} = \sum_{s=0}^{N-1} h^k\left(\frac{s}{N}\right) e^{-i\lambda s},$$

$T = S(M - 1) + N$, $u_j = t_j/T$, $t_j = S(j - 1) + N/2$, $j = 1, \dots, M$ y $h(\cdot)$ es un taper para los datos. La intuición detrás de esta extensión del estimador de Whittle en (2.9) es la siguiente: La muestra $\{Y_{1,T}, \dots, Y_{t,T}\}$ es subdividida en M bloques de largo N con saltos cada S de bloque en bloque. A modo de ejemplo, suponga que se dispone de una serie de largo $T = 652$ observaciones, con $M = 100$ bloques y cada bloque de largo $N = 256$ y el salto de $S = 4$ entre un bloque y otro, es decir, $(Y_{1,652}, \dots, Y_{256,652}), \dots, (Y_{398,652}, \dots, Y_{652,652})$. De esta forma, el estimador local de Whittle del vector de parámetro θ esta dado por

$$\hat{\theta}_T = \arg \min \mathcal{L}_T(\theta), \quad (2.10)$$

donde la minimización es bajo el espacio paramétrico Θ .

Suponga que se dispone de una secuencia estocástica $\{r_{1,T}, \dots, r_{T,T}\}$ generada por un proceso localmente estacionario $\{\nu_{t,T}\}$ y un ruido independiente idénticamente distribuido $\{u_t\}$ definido por las ecuaciones (2.5) y (2.6). En lo que sigue de la sección, vamos a revisar un conjunto de condiciones de regularidad que fueron definidos por Dahlhaus (1997) (que afectan a $\nu_{t,T}$ y ϵ_t), para luego presentar las propiedades asintóticas de los estimadores.

Supuestos 2.1.

- (i) La secuencia estocástica $\nu_{1,T}, \dots, \nu_{T,T}$ corresponde a un proceso localmente estacionario con verdadera función de transferencia A_ν^0 y media $\mu_\nu(u)$, sin embargo, asumimos como verdadera la densidad espectral $f_\nu(u, \lambda) = |A_\nu(u, \lambda)|^2$ con A_ν en la definición 2.1. $A_\nu(u, \lambda)$ es diferenciable en u y λ con derivada uniformemente acotada $\frac{\partial^2}{\partial u \partial \lambda} A_\nu$, y $g_{4,\nu}$ es continua.
- (ii) Se ajusta un modelo a una clase de procesos localmente estacionario con densidad espectral $f_\theta(u, \lambda)$ (2.8), donde $\theta \in \Theta \subset \mathbb{R}^p$, Θ es compacto. La densidad espectral $f_{\nu,\theta}(u, \lambda)$ (de $\nu_{t,T}$) es uniformemente acotada por arriba y por abajo. Los componentes de $f_{\nu,\theta}(u, \lambda)$, $\nabla f_{\nu,\theta}(u, \lambda)$ y $\nabla^2 f_{\nu,\theta}(u, \lambda)$ son continuos en $\Theta \times [0, 1] \times [-\pi, \pi]$ (∇ denota el gradiente con respecto a θ). $\nabla f_{\nu,\theta}^{-1}$ y $\nabla^2 f_{\nu,\theta}^{-1}$ son diferenciables en u y λ con derivada uniformemente acotada $(\partial/\partial u)(\partial/\partial \lambda)g$ donde $g = (\partial/\partial \theta_i) f_{\nu,\theta}^{-1}$ o $g = (\partial/\partial \theta_i)(\partial/\partial \theta_j) f_{\nu,\theta}^{-1}$.
- (iii) θ_0 existe, es único y se encuentra en el interior de Θ .
- (iv) N , S y T satisfacen la relación $T^{1/4} \ll N \ll T^{1/2}/\ln T$ y $S = N$ o $S/N \rightarrow 0$.
- (v) El taper para los datos $h : \mathbb{R} \rightarrow \mathbb{R}$ con $h(x) = 0$ para todo $x \notin [0, 1]$ es continuo en \mathbb{R} y dos veces diferenciables.

Adicionalmente, vamos a suponer que todos los momentos de $\{\epsilon_t = \log(u_t^2) - \mathbb{E}(\log(u_t^2))\}$ existen y denotaremos $\kappa_4 = \text{cum}_4(\epsilon_t)/\sigma_\epsilon^4$ al cumulante de 4to orden.

Para la estimación se debe elegir una función taper para los datos, en este capítulo se trabajará con la función taper de Tukey-Hanning,

$$h(x) = \begin{cases} \frac{1}{2}[1 - \cos(2\pi x)] , & \text{si } x \in (0, 1), \\ 0, & \text{en otro caso.} \end{cases}$$

Los siguientes dos teoremas corresponden a propiedades fundamentales del estimador de quasi-verosimilitud de Whittle en (2.10), que son la consistencia y el teorema del límite central. Las demostraciones de estos resultados se presentan en el apéndice B.

Teorema 2.1. *Sea θ_0 el verdadero valor del parámetro θ . Bajo los supuestos 2.1, el estimador $\hat{\theta}_T$ satisface $\hat{\theta}_T \rightarrow \theta_0$, en probabilidad, cuando $T \rightarrow \infty$.*

Teorema 2.2. *Sea θ_0 el verdadero valor del parámetro θ . Si se tienen los supuestos 2.1, entonces el estimador de Whittle $\hat{\theta}_T$ satisface el teorema del límite central*

$$\sqrt{T} \left(\hat{\theta}_T - \theta_0 \right) \rightarrow N \left(0, \Gamma(\theta_0)^{-1} [\Gamma(\theta_0) + \Lambda(\theta_0) + \Omega(\theta_0)] \Gamma(\theta_0)^{-1} \right),$$

en distribución, cuando $T \rightarrow \infty$. Donde

$$\begin{aligned} \Gamma(\theta_0) &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla \log f_{\theta_0}(u, \lambda) [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du. \\ \Lambda(\theta_0) &= \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu du. \\ \Omega(\theta_0) &= \frac{1}{8\pi} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f_{\nu, \theta_0}(u, \lambda) f_{\nu, \theta_0}(u, \mu) \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' g_{\nu, 4}(\lambda, -\lambda, \mu) d\lambda d\mu du. \end{aligned}$$

Aquí se considera la densidad espectral de $\{Y_{t,T}\}$ como $f_{\theta_0}(u, \lambda) = f_{\nu, \theta_0}(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi}$.

Observación 2.1. Note que si el proceso latente que genera la volatilidad $\{\nu_{t,T}\}$ es Gaussiano, entonces $g_{\nu, 4}(\lambda, -\lambda, \mu) = 0$ y por lo tanto $\Omega(\theta_0) = 0$. En ese caso el teorema del límite central para $\hat{\theta}_T$ se reduce a

$$\sqrt{T} (\widehat{\theta}_T - \theta_0) \rightarrow N(0, \Gamma(\theta_0)^{-1} [\Gamma(\theta_0) + \Lambda(\theta_0)] \Gamma(\theta_0)^{-1}).$$

2.3.1. Selección de Modelos

En la práctica surge la necesidad de contar con un índice que permita medir la calidad del ajuste de los datos y comparar entre una colección de modelos. Para construir este índice se va a definir la siguiente función

$$\mathcal{L}(\theta) := \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \log f_{\theta}(u, \lambda) + \frac{f(u, \lambda)}{f_{\theta}(u, \lambda)} d\lambda du$$

que mide la distancia entre la densidad espectral del proceso verdadero $f(u, \lambda)$ y del modelo propuesto $f_{\theta}(u, \lambda)$, también conocida como la distancia de Kullback-Leibler. La mejor aproximación del valor del parámetro de nuestra clase de modelos es

$$\theta_0 := \arg \min_{\theta \in \Theta} \mathcal{L}(\theta).$$

Se asume que la densidad espectral del verdadero proceso es $f(u, \lambda) = f_{\theta^*}(u, \lambda)$, y si el modelo es el correcto, entonces $\theta_0 = \theta^*$.

El criterio para medir la calidad del ajuste es estimar el valor medio de la distancia de Kullback Leibler entre la densidad espectral del proceso verdadero y del modelo propuesto, es decir, medir la calidad a través de $\mathbb{E}\mathcal{L}(\widehat{\theta}_T)$, donde $\widehat{\theta}_T$ es el estimador de Whittle en (2.10). Para estimar este índice se van a realizar dos expansiones cuadráticas, una de $\mathcal{L}(\theta)$ alrededor de θ_0 y otra de $\mathcal{L}_T(\theta)$ alrededor de $\widehat{\theta}_T$,

$$\mathcal{L}(\hat{\theta}_T) \approx \mathcal{L}(\theta_0) + \frac{1}{2} (\hat{\theta}_T - \theta_0)' \nabla^2 \mathcal{L}(\theta_0) (\hat{\theta}_T - \theta_0), \quad (2.11)$$

$$\mathcal{L}_T(\theta_0) \approx \mathcal{L}_T(\hat{\theta}_T) + \frac{1}{2} (\hat{\theta}_T - \theta_0)' \nabla^2 \mathcal{L}_T(\hat{\theta}_T) (\hat{\theta}_T - \theta_0). \quad (2.12)$$

De la demostración del teorema 2.1 (en el apéndice B) se puede verificar que $\lim_{t \rightarrow \infty} \mathbb{E} \mathcal{L}_T(\theta_0) = \mathcal{L}(\theta_0)$ y es un estimador consistente, $\nabla^2 \mathcal{L}(\theta_0) = \Gamma(\theta_0)$ y del teorema 2.2 (en el apéndice B) se tiene que $\nabla^2 \mathcal{L}_T(\hat{\theta}_T) \rightarrow \Gamma(\theta_0)$ en probabilidad. Utilizando los argumentos anteriores se puede estimar $\mathbb{E} \mathcal{L}(\hat{\theta}_T)$ por

$$\begin{aligned} \widehat{\mathbb{E}} \mathcal{L}(\hat{\theta}_T) &\approx \mathcal{L}_T(\hat{\theta}_T) + \mathbb{E} \left[(\hat{\theta}_T - \theta_0)' \nabla^2 \mathcal{L}(\theta_0) (\hat{\theta}_T - \theta_0) \right] \\ &\approx \mathcal{L}_T(\hat{\theta}_T) + \frac{1}{T} \text{Tr} \left[\Gamma(\theta_0)^{-1} (\Gamma(\theta_0) + \Lambda(\theta_0) + \Omega(\theta_0)) \right] \\ &\approx \mathcal{L}_T(\hat{\theta}_T) + \frac{p}{T} + \frac{1}{T} \text{Tr} \left[\Gamma(\theta_0)^{-1} (\Lambda(\theta_0) + \Omega(\theta_0)) \right], \end{aligned}$$

donde p es el número de parámetros estimados y las matrices $\Gamma(\theta_0)$, $\Lambda(\theta_0)$ y $\Omega(\theta_0)$ están definidas en el teorema 2.2. En la práctica se desconoce el verdadero valor del parámetro θ_0 , por lo tanto θ_0 será reemplazado por el estimador de Whittle

$$\widehat{\mathbb{E}} \mathcal{L}(\hat{\theta}_T) \approx \mathcal{L}_T(\hat{\theta}_T) + \frac{p}{T} + \frac{1}{T} \text{Tr} \left[\Gamma(\hat{\theta}_T)^{-1} (\Lambda(\hat{\theta}_T) + \Omega(\hat{\theta}_T)) \right].$$

Finalmente para obtener el usual AIC (Akaike (1974)) multiplicamos por 2

$$\text{AIC} = 2\mathcal{L}_T(\hat{\theta}_T) + \frac{2p}{T} + \frac{2}{T} \text{Tr} \left[\Gamma(\hat{\theta}_T)^{-1} (\Lambda(\hat{\theta}_T) + \Omega(\hat{\theta}_T)) \right].$$

2.4. Experimento de Monte Carlo

Primer escenario de simulación: Considere que $\{\nu_{t,T}\} \sim \text{LS-AR}(\phi(u), \sigma(u))$ que satisface la ecuación

$$\nu_{t,T} = \phi\left(\frac{t}{T}\right) \nu_{t-1,T} + \sigma\left(\frac{t}{T}\right) \varepsilon_t, \quad (2.13)$$

donde $|\phi(u)| < 1$ para todo $u \in [0, 1]$ y $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$. De esta forma el proceso de retornos será

$$r_{t,T} = \sigma_{t,T} u_t, \quad (2.14)$$

$$\sigma_{t,T} = \exp(\nu_{t,T}/2), \quad (2.15)$$

y considere a $\{u_t\} \stackrel{iid}{\sim} N(0, 1)$. Con el proceso $\{\nu_{t,T}\}$ especificado por (2.13), tenemos que la densidad espectral de $\{Y_{t,T}\}$ será

$$f_\theta(u, \lambda) = \frac{\sigma(u)^2}{2\pi} \frac{1}{1 - 2\phi(u) \cos(\lambda) + \phi(u)^2} + \frac{\pi}{4}.$$

Para estudiar el comportamiento del estimador, se consideró 1000 repeticiones del proceso (2.14). Los tamaños muestrales utilizados son $T = 512, 1024, 2048$, y para estos, se ilustra el estimador con saltos $S = 20, 30, 40$ y largo de los bloques $N = 128, 256$. Los parámetros $\phi(u)$ y $\sigma(u)$ fueron configurados de la siguiente manera:

$$\phi(u) = \beta_0 + \beta_1 u,$$

$$\sigma(u) = \alpha_0.$$

En el cálculo de la varianza teórica del estimador de Whittle del vector de parámetros, se necesita conocer los momentos de la distribución $\log u_t^2 \sim \log \chi_1^2$, estos son:

$$\begin{aligned}\mathbb{E}(\log \chi_1^2) &= \psi(1/2) - \log(1/2), \\ \mathbb{V}(\log \chi_1^2) &= \psi'(1/2).\end{aligned}$$

El tercer y cuarto momento centrados son:

$$\begin{aligned}\mathbb{E}(\log \chi_1^2 - \mathbb{E}(\log \chi_1^2))^3 &= \psi''(1/2), \\ \mathbb{E}(\log \chi_1^2 - \mathbb{E}(\log \chi_1^2))^4 &= \left(\psi^{(3)}(1/2) + 3\right) (\psi'(1/2))^2,\end{aligned}$$

donde $\psi(z) = \frac{d}{dz} \log \Gamma(z)$ y $\psi^{(s)}(z) = \frac{d^s}{dz^s} \psi(z)$. Así

$$\begin{aligned}\mathbb{E}(\log \chi_1^2) &= -1.27, \\ \mathbb{V}(\log \chi_1^2) &= \pi^2/2.\end{aligned}$$

Para la simulación de procesos localmente estacionarios se probaron dos técnicas, la primera técnica fue el algoritmo de innovaciones (ver Brockwell and Davis (1991) página 172) y la segunda fue un algoritmo recursivo utilizando la representación estocástica del proceso. El segundo algoritmo fue más eficiente computacionalmente y no utilizaba la matriz de varianza covarianza exacta. A continuación se presentan las tablas con los resultados de las simulaciones para $N = 128$.

Tabla 2.1: Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 128$ y salto $S = 20$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.50	0.35	0.471	0.363	10.177	0.092	0.152	0.364	0.072	0.110	0.337
	0.35	0.45	0.332	0.448	10.130	0.101	0.172	0.366	0.079	0.123	0.335
	0.35	-0.25	0.321	-0.226	9.998	0.096	0.167	0.337	0.088	0.154	0.330
	0.40	0.20	0.384	0.198	10.083	0.102	0.183	0.380	0.081	0.136	0.334
1024	0.50	0.35	0.481	0.356	10.206	0.060	0.100	0.250	0.051	0.078	0.239
	0.35	0.45	0.335	0.447	10.141	0.066	0.111	0.251	0.056	0.087	0.237
	0.35	-0.25	0.332	-0.247	9.995	0.064	0.113	0.244	0.062	0.110	0.234
	0.40	0.20	0.379	0.205	10.088	0.068	0.120	0.252	0.058	0.096	0.236
2048	0.50	0.35	0.483	0.353	10.215	0.040	0.066	0.178	0.036	0.055	0.169
	0.35	0.45	0.333	0.453	10.144	0.043	0.068	0.176	0.040	0.061	0.168
	0.35	-0.25	0.333	-0.249	10.000	0.045	0.079	0.165	0.044	0.077	0.165
	0.40	0.20	0.384	0.197	10.100	0.045	0.077	0.176	0.041	0.068	0.167
4096	0.50	0.35	0.483	0.356	10.217	0.027	0.043	0.125	0.025	0.039	0.119
	0.35	0.45	0.332	0.456	10.151	0.030	0.049	0.126	0.028	0.044	0.119
	0.35	-0.25	0.331	-0.246	9.995	0.033	0.058	0.116	0.031	0.055	0.117
	0.40	0.20	0.382	0.201	10.108	0.029	0.051	0.123	0.029	0.048	0.118

Tabla 2.2: Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 128$ y salto $S = 30$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.50	0.35	0.471	0.363	10.170	0.092	0.160	0.372	0.072	0.110	0.337
	0.35	0.45	0.332	0.447	10.125	0.102	0.180	0.368	0.079	0.123	0.335
	0.35	-0.25	0.320	-0.222	10.002	0.096	0.171	0.334	0.088	0.154	0.330
	0.40	0.20	0.383	0.198	10.081	0.104	0.195	0.387	0.081	0.136	0.334
1024	0.50	0.35	0.481	0.356	10.204	0.060	0.100	0.250	0.051	0.078	0.239
	0.35	0.45	0.335	0.447	10.139	0.066	0.111	0.251	0.056	0.087	0.237
	0.35	-0.25	0.332	-0.246	9.996	0.064	0.113	0.243	0.062	0.110	0.234
	0.40	0.20	0.379	0.205	10.087	0.068	0.119	0.252	0.058	0.096	0.236
2048	0.50	0.35	0.483	0.353	10.215	0.039	0.065	0.178	0.036	0.055	0.169
	0.35	0.45	0.332	0.453	10.144	0.043	0.067	0.176	0.040	0.061	0.168
	0.35	-0.25	0.333	-0.249	10.000	0.045	0.078	0.165	0.044	0.077	0.165
	0.40	0.20	0.384	0.197	10.100	0.045	0.077	0.176	0.041	0.068	0.167
4096	0.50	0.35	0.483	0.356	10.217	0.027	0.043	0.125	0.025	0.039	0.119
	0.35	0.45	0.332	0.456	10.151	0.030	0.048	0.126	0.028	0.044	0.119
	0.35	-0.25	0.331	-0.246	9.995	0.033	0.058	0.116	0.031	0.055	0.117
	0.40	0.20	0.382	0.201	10.108	0.029	0.051	0.123	0.029	0.048	0.118

Tabla 2.3: Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 128$ y salto $S = 40$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.50	0.35	0.471	0.363	10.171	0.090	0.155	0.365	0.072	0.110	0.337
	0.35	0.45	0.332	0.447	10.126	0.100	0.177	0.364	0.079	0.123	0.335
	0.35	-0.25	0.321	-0.224	9.999	0.095	0.169	0.339	0.088	0.154	0.330
	0.40	0.20	0.384	0.197	10.080	0.102	0.191	0.382	0.081	0.136	0.334
1024	0.50	0.35	0.481	0.356	10.207	0.059	0.097	0.248	0.051	0.078	0.239
	0.35	0.45	0.335	0.447	10.141	0.064	0.107	0.249	0.056	0.087	0.237
	0.35	-0.25	0.332	-0.247	9.996	0.063	0.111	0.243	0.062	0.110	0.234
	0.40	0.20	0.379	0.205	10.088	0.067	0.116	0.251	0.058	0.096	0.236
2048	0.50	0.35	0.483	0.353	10.215	0.039	0.065	0.178	0.036	0.055	0.169
	0.35	0.45	0.333	0.453	10.144	0.042	0.067	0.175	0.040	0.061	0.168
	0.35	-0.25	0.333	-0.249	10.000	0.045	0.078	0.165	0.044	0.077	0.165
	0.40	0.20	0.384	0.197	10.100	0.045	0.077	0.176	0.041	0.068	0.167
4096	0.50	0.35	0.483	0.356	10.217	0.027	0.043	0.125	0.025	0.039	0.119
	0.35	0.45	0.332	0.456	10.151	0.030	0.048	0.126	0.028	0.044	0.119
	0.35	-0.25	0.331	-0.246	9.995	0.033	0.058	0.116	0.031	0.055	0.117
	0.40	0.20	0.383	0.201	10.108	0.029	0.050	0.123	0.029	0.048	0.118

Fijando el tamaño de ventana en $N = 128$ y variando el tamaño de salto S , se puede apreciar que las desviaciones de Monte Carlo son muy similares entre si, quizás para el valor de salto $S = 40$ es donde se observa la mayor precisión (recién en el tercer decimal). Sin embargo, estas diferencias en las desviaciones puede llegar a ser despreciable.

Si se analiza el sesgo cometido en las estimaciones por los diferentes tamaños de saltos S , se puede notar que no existen diferencia. Se debe tener en cuenta que las conclusiones pueden variar dependiendo de la estructura elegida para los parámetros, ya que aquí sólo se tienen 4 configuraciones paramétricas. A continuación se presenta el mismo ejercicio fijando el tamaño de la ventana en $N = 256$

Tabla 2.4: Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 256$ y salto $S = 20$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.50	0.35	0.487	0.348	10.062	0.122	0.220	0.425	0.072	0.110	0.337
	0.35	0.45	0.340	0.447	10.061	0.137	0.249	0.416	0.079	0.123	0.335
	0.35	-0.25	0.319	-0.206	9.998	0.127	0.232	0.384	0.088	0.154	0.330
	0.40	0.20	0.388	0.205	10.027	0.145	0.279	0.447	0.081	0.136	0.334
1024	0.50	0.35	0.490	0.355	10.094	0.067	0.115	0.261	0.051	0.078	0.239
	0.35	0.45	0.345	0.446	10.064	0.075	0.129	0.263	0.056	0.087	0.237
	0.35	-0.25	0.339	-0.245	9.992	0.077	0.135	0.253	0.062	0.110	0.234
	0.40	0.20	0.387	0.205	10.034	0.080	0.141	0.260	0.058	0.096	0.236
2048	0.50	0.35	0.493	0.349	10.106	0.041	0.069	0.179	0.036	0.055	0.169
	0.35	0.45	0.342	0.451	10.065	0.045	0.073	0.179	0.040	0.061	0.168
	0.35	-0.25	0.341	-0.249	10.001	0.049	0.087	0.171	0.044	0.077	0.165
	0.40	0.20	0.394	0.195	10.048	0.047	0.084	0.181	0.041	0.068	0.167
4096	0.50	0.35	0.492	0.353	10.106	0.027	0.043	0.126	0.025	0.039	0.119
	0.35	0.45	0.341	0.453	10.072	0.030	0.049	0.126	0.028	0.044	0.119
	0.35	-0.25	0.340	-0.248	9.995	0.034	0.060	0.118	0.031	0.055	0.117
	0.40	0.20	0.391	0.201	10.052	0.030	0.052	0.123	0.029	0.048	0.118

Tabla 2.5: Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 256$ y salto $S = 30$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.50	0.35	0.486	0.350	10.062	0.118	0.213	0.418	0.072	0.110	0.337
	0.35	0.45	0.350	0.432	10.056	0.134	0.246	0.430	0.079	0.123	0.335
	0.35	-0.25	0.326	-0.214	9.991	0.128	0.229	0.384	0.088	0.154	0.330
	0.40	0.20	0.388	0.198	10.027	0.135	0.261	0.417	0.081	0.136	0.334
1024	0.50	0.35	0.490	0.355	10.094	0.067	0.116	0.261	0.051	0.078	0.239
	0.35	0.45	0.345	0.446	10.063	0.074	0.130	0.263	0.056	0.087	0.237
	0.35	-0.25	0.339	-0.245	9.995	0.076	0.135	0.253	0.062	0.110	0.234
	0.40	0.20	0.387	0.205	10.034	0.079	0.142	0.260	0.058	0.096	0.236
2048	0.50	0.35	0.493	0.349	10.106	0.041	0.069	0.179	0.036	0.055	0.169
	0.35	0.45	0.342	0.451	10.065	0.045	0.073	0.179	0.040	0.061	0.168
	0.35	-0.25	0.341	-0.249	10.001	0.049	0.087	0.171	0.044	0.077	0.165
	0.40	0.20	0.394	0.195	10.048	0.047	0.084	0.182	0.041	0.068	0.167
4096	0.50	0.35	0.492	0.353	10.106	0.027	0.043	0.125	0.025	0.039	0.119
	0.35	0.45	0.341	0.453	10.072	0.030	0.049	0.126	0.028	0.044	0.119
	0.35	-0.25	0.340	-0.248	9.995	0.034	0.060	0.118	0.031	0.055	0.117
	0.40	0.20	0.391	0.201	10.052	0.030	0.052	0.123	0.029	0.048	0.118

Tabla 2.6: Estimación de Whittle: $\alpha_0 = 10$, tamaño de ventana $N = 256$ y salto $S = 40$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.50	0.35	0.490	0.345	10.058	0.120	0.214	0.423	0.072	0.110	0.337
	0.35	0.45	0.340	0.448	10.032	0.132	0.240	0.424	0.079	0.123	0.335
	0.35	-0.25	0.321	-0.205	9.995	0.125	0.225	0.378	0.088	0.154	0.330
	0.40	0.20	0.389	0.192	10.035	0.135	0.260	0.414	0.081	0.136	0.334
1024	0.50	0.35	0.490	0.354	10.095	0.066	0.112	0.258	0.051	0.078	0.239
	0.35	0.45	0.345	0.446	10.065	0.073	0.125	0.260	0.056	0.087	0.237
	0.35	-0.25	0.340	-0.246	9.993	0.075	0.131	0.252	0.062	0.110	0.234
	0.40	0.20	0.387	0.205	10.034	0.077	0.137	0.258	0.058	0.096	0.236
2048	0.50	0.35	0.493	0.349	10.105	0.041	0.069	0.179	0.036	0.055	0.169
	0.35	0.45	0.342	0.451	10.064	0.045	0.074	0.179	0.040	0.061	0.168
	0.35	-0.25	0.341	-0.249	10.001	0.049	0.087	0.172	0.044	0.077	0.165
	0.40	0.20	0.394	0.195	10.048	0.047	0.084	0.182	0.041	0.068	0.167
4096	0.50	0.35	0.492	0.353	10.106	0.027	0.043	0.125	0.025	0.039	0.119
	0.35	0.45	0.341	0.453	10.072	0.030	0.049	0.126	0.028	0.044	0.119
	0.35	-0.25	0.340	-0.248	9.995	0.034	0.059	0.118	0.031	0.055	0.117
	0.40	0.20	0.391	0.201	10.052	0.030	0.052	0.123	0.029	0.048	0.118

Al comparar la precisión muestral de las estimaciones en estas 3 tablas, se puede observar que para tamaños muestrales más pequeños el salto óptimo sería $S = 30$, y para tamaños muestrales más grandes el salto óptimo sería $S = 40$. Sin embargo, las diferencias son mínimas, y se aprecian sólo en algunos modelos.

Si se compara los resultados de las 6 tablas, se puede concluir que la elección de S y N afecta la precisión de los resultados, ya que cuando los tamaños muestrales son pequeños, por ejemplo en $T = 512$, el estimador se comporta mejor para un tamaño de ventana de $N = 128$, donde las estimaciones presentan una desviación muestral menor en un 30% que la obtenida con $N = 256$, sin embargo, a medida que aumente el largo de la muestra, las desviaciones de ambas tablas se asemejan. En Dahlhaus and Giraitis (1998) se determina el valor óptimo para S y N , sin embargo en los estudio empíricos los resultados pueden variar dependiendo del taper elegido.

Segundo escenario de simulación: Considere un proceso localmente estacionario $\{\nu_{t,T}\} \sim \text{LS-ARMA}(\phi(u), \vartheta(u), \sigma(u))$, es decir,

$$\nu_{t,T} = \phi\left(\frac{t}{T}\right) \nu_{t-1,T} + \sigma\left(\frac{t}{T}\right) \varepsilon_t + \vartheta\left(\frac{t}{T}\right) \sigma\left(\frac{t-1}{T}\right) \varepsilon_{t-1},$$

donde $|\phi(u)| < 1$ para todo $u \in (0, 1)$, $|\vartheta(u)| < 1$ y $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$. El proceso de retornos será

$$r_{t,T} = \sigma_{t,T} u_t,$$

$$\sigma_{t,T} = \exp(\nu_{t,T}/2),$$

y considere $\{u_t\} \stackrel{iid}{\sim} N(0, 1)$. La densidad espectral del proceso $\{Y_{t,T}\}$ es

$$f_{\theta}(u, \lambda) = \frac{\sigma(u)^2}{2\pi} \frac{1 + 2\vartheta(u) \cos(\lambda) + \vartheta(u)^2}{1 - 2\phi(u) \cos(\lambda) + \phi(u)^2} + \frac{\pi}{4}.$$

Para este experimento de Monte Carlo se consideraron 1000 repeticiones del proceso de retornos. Los parámetros $\phi(u)$ y $\sigma(u)$ se consideraron de la siguiente manera:

$$\phi(u) = \beta_0 + \beta_1 u,$$

$$\vartheta(u) = \vartheta,$$

$$\sigma(u) = \alpha_0 + \alpha_1 u.$$

A diferencia del primer experimento, aquí se presentan los resultados de Monte Carlo para una configuración fija de S y N . La elección del par S y N es aquella que minimiza el error cuadrático medio (para un valor θ_0 específico).

En la siguiente figura, se muestra la curva de contorno del error cuadrático empírico MSE del estimador de Whittle $\hat{\theta}$ como el promedio de la distancia Euclidiana $\|\hat{\theta} - \theta_0\|^2$ considerando 100 estimaciones de θ_0 , esta curva se construyó con la configuración $\theta_0 = (0.5, 0.3, -0.5, 2.0, 1.5)$, para un largo de serie $T = 512$. En la estimación se utilizó el taper de Tukey-Hanning. Cabe mencionar que los resultados fueron similares para otros valores paramétricos.

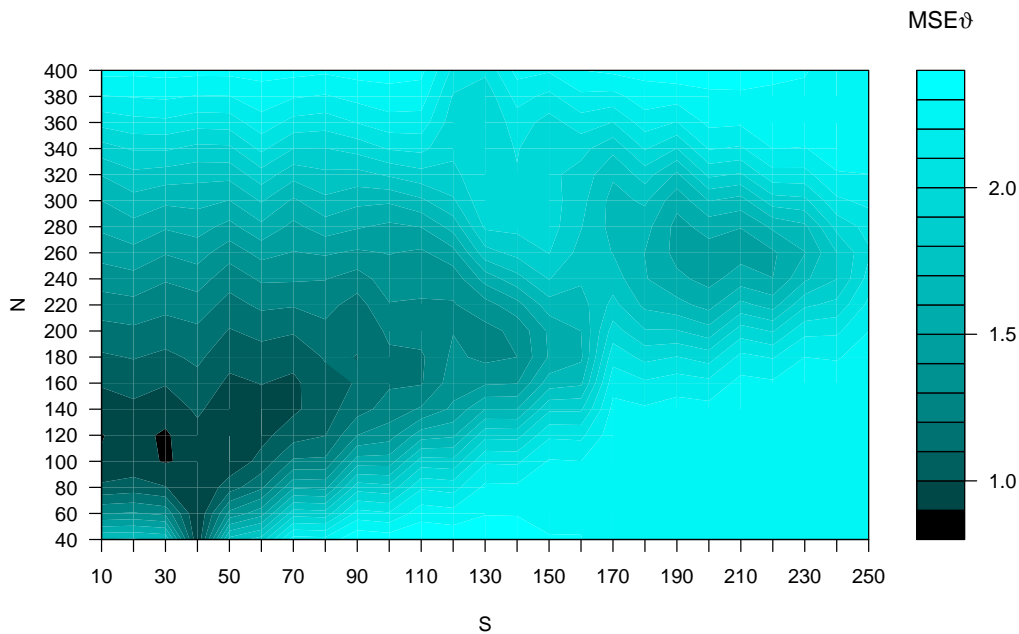


Figura 2.2: Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARMA con $T = 512$.

Note que para un tamaño de muestra $T = 512$, los valores que minimizan el error cuadrático medio se encuentran cercanos a $S \approx 30$ y $N \approx 120$. Utilizando estos valores, se presentan los resultados de la simulación de Monte Carlo basado en las 1000 repeticiones.

Tabla 2.7: Estimación de Whittle: tamaño muestral $T = 512$, tamaño de ventana $N = 120$ y salto $S = 30$

Caso	Parámetros					Estimaciones				
	β_0	β_1	ϑ	α_0	α_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\vartheta}$	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$
1	0.5	0.35	0.4	5	3	0.447	0.364	0.448	5.059	3.269
2	0.8	-0.3	0.4	5	3	0.756	-0.315	0.448	5.165	3.023
3	0.5	0.3	0.4	2	1.5	0.486	0.289	0.518	1.901	1.444
4	0.6	-0.3	0.4	7	3	0.542	-0.296	0.440	7.150	3.031
5	0.3	0.4	0.4	2	2	0.287	0.387	0.487	1.924	1.947
6	0.5	0.25	0.4	2	2	0.477	0.253	0.491	1.924	1.927
7	0.5	0.35	-0.5	5	3	0.334	0.409	-0.365	4.934	3.071
8	0.8	-0.3	-0.5	5	3	0.664	-0.347	-0.347	5.000	2.919
9	0.5	0.3	-0.5	2	1.5	0.403	0.312	-0.409	1.948	1.534
10	0.6	-0.3	-0.5	7	3	0.720	-0.284	-0.675	6.896	2.831
11	0.6	0.3	-0.5	2	2	0.539	0.309	-0.441	1.981	1.989
12	0.5	0.25	-0.5	2	2	0.268	0.401	-0.480	1.961	2.030

Caso	SD Teórica					SD Empírica				
	$\sigma(\widehat{\beta}_0)$	$\sigma(\widehat{\beta}_1)$	$\sigma(\widehat{\vartheta})$	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\vartheta})$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$
1	0.088	0.126	0.074	0.755	1.407	0.110	0.178	0.075	0.584	1.129
2	0.068	0.134	0.078	0.779	1.395	0.094	0.178	0.079	0.627	1.131
3	0.119	0.168	0.187	0.394	0.623	0.137	0.215	0.284	0.466	0.719
4	0.086	0.151	0.071	0.775	1.416	0.115	0.197	0.072	0.713	1.360
5	0.154	0.210	0.157	0.386	0.660	0.167	0.253	0.236	0.449	0.758
6	0.119	0.172	0.162	0.389	0.638	0.147	0.227	0.249	0.474	0.752
7	0.154	0.136	0.125	0.495	0.943	0.181	0.185	0.160	0.512	1.002
8	0.115	0.140	0.157	0.490	0.922	0.190	0.203	0.187	0.506	0.941
9	0.239	0.198	0.195	0.342	0.595	0.258	0.269	0.241	0.400	0.701
10	0.323	0.154	0.334	0.512	0.937	0.256	0.202	0.293	0.537	0.982
11	0.144	0.149	0.114	0.347	0.610	0.185	0.202	0.152	0.407	0.735
12	0.269	0.196	0.230	0.341	0.597	0.274	0.272	0.271	0.359	0.642

Los resultados presentados en este estudio de simulación sugieren que el desempeño del estimador propuesto es bastante bueno considerando que la transformación $\log(r_{i,T}^2)$ tiene un ruido aditivo.

La siguiente figura corresponde a la curva de nivel del error cuadrático medio MSE cuando el tamaño muestral de la serie es $T = 1024$. Aquí nuevamente se ilustra el caso $\theta_0 = (0.5, 0.3, -0.5, 2.0, 1.5)$ (para ser consecuente con el gráfico anterior). De la misma manera que en el gráfico anterior, la curva se construye con el promedio la distancia $\|\hat{\theta} - \theta_0\|^2$ del estimador local de Whittle $\hat{\theta}$ para 100 repeticiones.

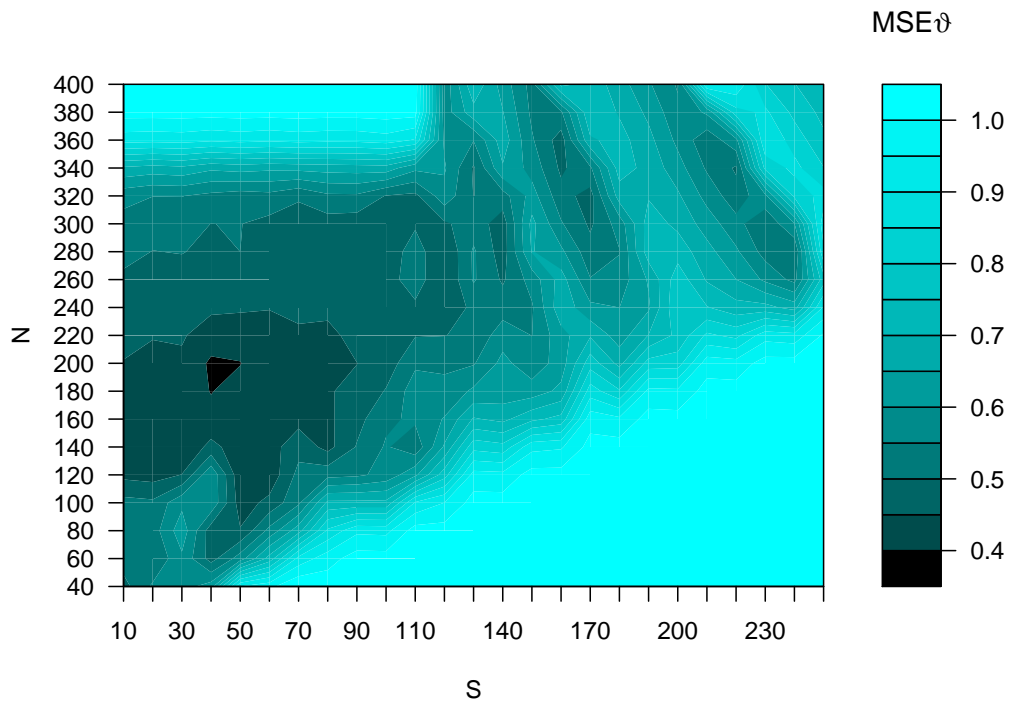


Figura 2.3: Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARMA con $T = 1024$.

De este gráfico ilustrativo se desprende que los valores que minimizan el error cuadrático medio son $S \approx 45$ y $N \approx 210$. La siguiente tabla muestra los resultados obtenidos en la simulación de Monte Carlo utilizando los valores óptimos en la curva de nivel.

Tabla 2.8: Estimación de Whittle: tamaño muestral $T = 1024$, tamaño de ventana $N = 210$ y salto $S = 45$

Caso	Parámetros					Estimaciones				
	β_0	β_1	ϑ	α_0	α_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\vartheta}$	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$
1	0.5	0.35	0.4	5	3	0.470	0.359	0.425	5.042	3.131
2	0.8	-0.3	0.4	5	3	0.777	-0.305	0.425	5.114	2.978
3	0.5	0.3	0.4	2	1.5	0.492	0.295	0.456	1.952	1.462
4	0.6	-0.3	0.4	7	3	0.573	-0.311	0.425	7.089	2.982
5	0.3	0.4	0.4	2	2	0.282	0.402	0.434	1.974	1.966
6	0.5	0.25	0.4	2	2	0.484	0.253	0.447	1.960	1.971
7	0.5	0.35	-0.5	5	3	0.413	0.384	-0.431	4.967	3.064
8	0.8	-0.3	-0.5	5	3	0.726	-0.332	-0.411	4.998	2.984
9	0.5	0.3	-0.5	2	1.5	0.436	0.320	-0.444	1.981	1.504
10	0.6	-0.3	-0.5	7	3	0.674	-0.279	-0.604	6.949	2.919
11	0.6	0.3	-0.5	2	2	0.571	0.307	-0.477	1.986	1.997
12	0.5	0.25	-0.5	2	2	0.267	0.398	-0.474	1.968	2.018

Caso	SD Teórica					SD Empírica				
	$\sigma(\widehat{\beta}_0)$	$\sigma(\widehat{\beta}_1)$	$\sigma(\widehat{\vartheta})$	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\vartheta})$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$
1	0.062	0.089	0.053	0.534	0.995	0.071	0.109	0.055	0.399	0.763
2	0.048	0.094	0.055	0.551	0.986	0.063	0.117	0.056	0.425	0.761
3	0.084	0.118	0.132	0.279	0.441	0.093	0.141	0.210	0.342	0.516
4	0.061	0.107	0.050	0.548	1.001	0.071	0.128	0.054	0.506	0.930
5	0.109	0.149	0.111	0.273	0.467	0.113	0.161	0.148	0.300	0.504
6	0.084	0.122	0.115	0.275	0.451	0.098	0.143	0.171	0.317	0.498
7	0.109	0.096	0.088	0.350	0.667	0.125	0.126	0.101	0.362	0.701
8	0.082	0.099	0.111	0.347	0.652	0.127	0.130	0.136	0.359	0.691
9	0.169	0.140	0.138	0.242	0.420	0.209	0.186	0.177	0.285	0.505
10	0.228	0.109	0.236	0.362	0.663	0.236	0.155	0.265	0.425	0.783
11	0.101	0.105	0.080	0.245	0.432	0.127	0.132	0.101	0.283	0.515
12	0.189	0.139	0.163	0.241	0.422	0.245	0.176	0.237	0.244	0.431

Al analizar los resultados de la tabla 2.8, las conclusiones en cuanto a la precisión del estimador son similares a las realizadas sobre la tabla 2.7, es decir, se observa que el estimador tiene un buen desempeño para muestras finitas, considerando que el proceso es la suma de un proceso latente no observado más un ruido blanco. Note que al aumentar el tamaño muestral de 512 a 1024, se produce automáticamente una reducción en el sesgo (que ya era muy ligero considerando la estructura).

En líneas generales el experimento de Monte Carlo muestra que el desempeño del estimador es bastante bueno, ya que en el primero escenario de simulación se consideró un parámetro de dispersión muy alto, y en segundo escenario se redujo considerablemente dicho valor, aún así los resultados empíricos en ambos casos fueron buenos.

Capítulo 3

Procesos de Volatilidad Estocástica Localmente Estacionarios de Larga Memoria

En el presente capítulo se introduce una nueva clase de modelos de volatilidad estocástica, donde el procesos latentes no observado que generar la volatilidad corresponde a un proceso localmente estacionario de larga memoria.

La densidad espectral de los procesos de larga memoria tiene una singularidad en la frecuencia cero, lo que aumenta la complejidad del problema. Es por este motivo que se trabajara con procesos Gaussianos y se utilizarán los supuestos de Palma and Olea (2010).

En la literatura actual, se encuentran otros trabajos en procesos localmente estacionarios de larga memoria, como Beran (2009) y Roueff and von Sachs (2011). En el primero, se discute un estimador por mínimos cuadrados utilizando una representación autoregresiva infinita del proceso, y en el segundo, se utiliza un estimador del tipo wavelet. El estimador propuesto en este capítulo, es similar al utilizado en el capítulo 2, y corresponde al estimador por bloques de Whittle.

3.1. Procesos Localmente Estacionarios de Larga Memoria

Una clase de procesos Gaussiano localmente estacionario con función de transferencia A^0 puede ser definido por la siguiente representación espectral

$$X_{t,T} = \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^0(\lambda) d\xi(\lambda), \quad (3.1)$$

para $t = 1, \dots, T$, donde $\xi(\lambda)$ es un movimiento Browniano en $[-\pi, \pi]$ y existe una constante positiva K y una función de transferencia 2π periódica $A : (0, 1] \times \mathbb{R} \rightarrow \mathbb{C}$ con $A(u, \lambda) = A(u, -\lambda)$ tal que

$$\sup_{t,\lambda} \left| A_{t,T}^0(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq KT^{-1}, \quad (3.2)$$

para todo T . La función de transferencia $A_{t,T}^0(\lambda)$ de esta clase de procesos no estacionarios cambia suavemente a través del tiempo, y es por esto que en una vecindad puede ser aproximada por un proceso estacionario. Un ejemplo de esta clase de procesos localmente estacionarios está dado por la expansión de medias móviles infinita

$$X_{t,T} = \sigma \left(\frac{t}{T} \right) \sum_{j=0}^{\infty} \psi_j \left(\frac{t}{T} \right) \varepsilon_{t-j}, \quad (3.3)$$

donde $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$ y los coeficientes $\{\psi_j(u)\}$ satisfacen $\sum_{j=0}^{\infty} \psi_j(u)^2 < \infty$ para todo $u \in [0, 1]$. En este ejemplo, la función de transferencia del proceso está dada por $A_{t,T}^0(\lambda) = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \psi_j\left(\frac{t}{T}\right) e^{-i\lambda j} = A\left(\frac{t}{T}, \lambda\right)$, así, la condición (3.2) se satisface de manera trivial.

Ejemplo 3.1. Ejemplos de procesos localmente estacionarios de larga memoria:

- (i) Suponga que $\{Y_t\}$ es un proceso Gaussiano ARFIMA(p, d, q) estacionario e invertible con representación espectral

$$Y_t = \int_{-\pi}^{\pi} \exp(i\lambda t) A(\lambda) d\xi(\lambda),$$

donde $\xi(\lambda)$ es un movimiento Browniano en $[-\pi, \pi]$ y la función de transferencia está dada por $A(\lambda) = \theta(e^{-i\lambda})\phi(e^{-i\lambda})^{-1}(1 - e^{-i\lambda})^{-d}$, además $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ y $\phi(B) = 1 + \phi_1 B + \dots + \phi_p B^p$ son los polinomios de medias móviles y autoregresivos respectivamente (con ceros fuera del disco unitario) y $d \in (-1, 0.5)$. Suponga que $\sigma : [0, 1] \rightarrow \mathbb{R}^+$ es continuo. Entonces

$$X_{t,T} = \sigma(t/T)Y_t$$

es un proceso localmente estacionario con $A_{t,T}^0(\lambda) = A(t/T, \lambda) = \sigma(t/T)A(\lambda)$.

- (ii) Considere a $\{Y_t\} \sim \text{LS-FN}(d(u), \sigma(u))$, como la versión generalizada del ruido fraccionado Gaussiano descrito por la ecuación

$$X_{t,T} = \sigma\left(\frac{t}{T}\right) (1 - B)^{-d(t/T)} \varepsilon_t = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \eta_j\left(\frac{t}{T}\right) \varepsilon_{t-j}, \quad (3.4)$$

para $t = 1, \dots, T$, donde $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$ y los coeficientes $\{\eta_j(u)\}$ están dados por

$$\eta_j(u) = \frac{\Gamma[j + d(u)]}{\Gamma[j + 1]\Gamma[d(u)]}, \quad (3.5)$$

aquí $\Gamma(\cdot)$ corresponde a la función Gamma y $d(\cdot)$ es el parámetro de larga memoria de variación suave. En este ejemplo $A_{t,T}^0(\lambda) = A(t/T, \lambda) = \sigma\left(\frac{t}{T}\right) (1 - e^{-i\lambda})^{-d(t/T)}$.

(iii) Suponga que $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$ y

$$\begin{aligned} X_{t,T} &= \sigma \left(\frac{t}{T} \right) (1 - \vartheta B)(1 - B)^{-d(t/T)} \varepsilon_t \\ &= \sigma \left(\frac{t}{T} \right) \sum_{j=0}^{\infty} \varphi_j \left(\frac{t}{T} \right) \varepsilon_{t-j}, \end{aligned}$$

donde

$$\varphi_j(u) = \begin{cases} 1, & \text{si } j = 0, \\ \eta_j(u) - \vartheta \eta_{j-1}(u), & \text{si } j > 0. \end{cases} \quad (3.6)$$

y los coeficientes $\{\eta_j(u)\}$ están definidos en (3.5). Aquí la función de transferencia está dada por

$$A_{t,T}^0(\lambda) = A(t/T, \lambda) = \sigma \left(\frac{t}{T} \right) \sum_{j=0}^{\infty} \varphi_j(t/T) \exp(-i\lambda j).$$

En el capítulo 2 se presentó una definición de la densidad espectral de variación en el tiempo para un proceso localmente estacionario. Esta definición, fue realizada por Dahlhaus (1996) y corresponde a la densidad espectral límite. En este capítulo se utilizará la misma definición

$$f(u, \lambda) := |A(u, \lambda)|^2.$$

Note que en los ejemplos de procesos localmente estacionarios de larga memoria, siempre se utilizó la función de transferencia exacta del proceso, es decir, se consideró $A_{t,T}^0(\lambda) = A(t/T, \lambda)$.

Ejemplo 3.2. Sea $X_{t,T} = \sigma(t/T)(1 - B)^{-d(t/T)}\varepsilon_t = \sigma(t/T) \sum_{j=0}^{\infty} \eta_j(t/T)\varepsilon_{t-j}$, donde $\varepsilon_t \stackrel{iid}{\sim} N(0, 1)$, $\eta_j(u) = \frac{\Gamma[d(u)]}{\Gamma[j+1]\Gamma[d(u)]}$ y $d(u) = \beta_0 + \beta_1 u + \beta_2 u^2$. Para este proceso la función de covarianza exacta está dada por

$$\text{Cov}(X_{t,T}, X_{s,T}) = \sigma\left(\frac{t}{T}\right) \sigma\left(\frac{s}{T}\right) \frac{\Gamma[1 - d(t/T) - d(s/T)]\Gamma[t - s + d(t/T)]}{\Gamma[1 - d(t/T)]\Gamma[d(t/T)]\Gamma[t - s + 1 - d(s/T)]}$$

donde $t, s = 1, \dots, T$ y $t \geq s$. La densidad espectral del ejemplo es

$$f(u, \lambda) = \frac{1}{2\pi} \left(2 \sin \frac{\lambda}{2} \right)^{-2d(u)}.$$

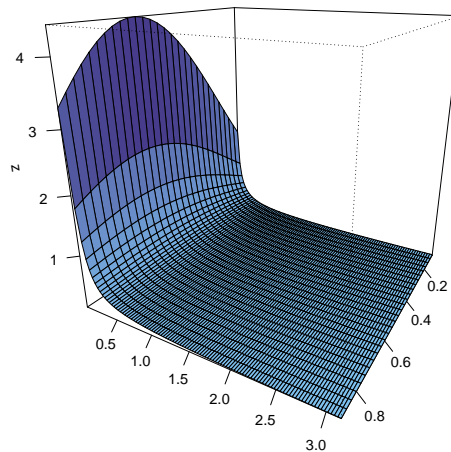


Figura 3.1: Densidad espectral del proceso LS-FN donde $\beta_0 = 0.3$, $\beta_1 = 0.6$ y $\beta_2 = 0.5$.

En la figura se puede observar que a medida que frecuencia λ se acerca a cero, la densidad comienza a crecer rápidamente. Esta característica es particular de los procesos de larga memoria.

3.2. Modelos de Volatilidad Estocástica de Larga Memoria

Suponga que se observa de una serie de retornos $\{r_t; t \in \{1, \dots, T\}\}$, el índice será reescalado por el largo de la serie, es decir, la serie de retornos de largo T se va a denotar como $\{r_{t,T}; t \in \{1, \dots, T\}\}$. Diremos que el proceso de retornos $\{r_{t,T}\}$ sigue un proceso de volatilidad estocástica localmente estacionario de larga memoria, si está definido por las siguientes ecuaciones

$$r_{t,T} = \sigma_{t,T} u_t, \quad (3.7)$$

$$\sigma_{t,T} = \sigma \exp(\nu_{t,T}/2), \quad (3.8)$$

donde $\sigma > 0$, $\{u_t\} \stackrel{iid}{\sim} (0, 1)$ y $\{\nu_{t,T}\}$ es un proceso Gaussiano localmente estacionario de larga memoria independiente de $\{u_t\}$. De la misma manera que en el capítulo de corta memoria, se utilizan las ecuaciones (3.7) y (3.8), para obtener

$$\log(r_{t,T}^2) = \log(\sigma_{t,T}^2) + \log(u_t^2),$$

$$\log(\sigma_{t,T}^2) = \log(\sigma^2) + \nu_{t,T},$$

luego, sumando y restando la media de $\{\log(u_t^2)\}$ se tiene

$$Y_{t,T} = \mu + \nu_{t,T} + \epsilon_t, \quad (3.9)$$

donde $Y_{t,T} = \log(r_{t,T}^2)$, $\mu = \log(\sigma^2) + \mathbb{E}(\log(u_t^2))$, $\epsilon_t = \log(u_t^2) - \mathbb{E}(\log(u_t^2))$. Aquí la transformación $\{Y_{t,T}\}$ corresponde a la suma de un proceso localmente estacionario Gaussiano de larga memoria más una secuencia de ruido blanco.

La función de autocovarianza exacta para el proceso $\{Y_{t,T}\}$ es

$$\text{Cov}(Y_{t,T}, Y_{t+k,T}) = \text{Cov}(\nu_{t,T}, \nu_{t+k,T}) + \sigma_\epsilon^2 \delta_0(k), \quad \forall k \in \mathbb{N},$$

donde $\delta_0(k) = 1$ si $k = 0$ y $\delta_0(k) = 0$ si $k > 0$. Utilizando el lema A.5 se puede obtener la densidad espectral de $\{Y_{t,T}\}$, que está dada por

$$f_Y(u, \lambda) = f_\nu(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi}, \quad (3.10)$$

aquí $f_\nu(u, \lambda)$ corresponde a la densidad espectral de $\{\nu_{t,T}\}$ y σ_ϵ^2 es la varianza de $\{\epsilon_t\}$ (este resultado es valido para procesos localmente estacionarios de corta y larga memoria).

Ejemplo 3.3. Sea $\{\nu_{t,T}\}$ es una secuencia LS-ARFIMA definido por

$$\nu_{t,T} = \sigma \left(\frac{t}{T} \right) \left[1 - \vartheta \left(\frac{t}{T} \right) \right] (1 - B)^{-d(t/T)} \epsilon_t,$$

para todo $t = 1, 2, \dots, T$ donde $(d(u), \vartheta(u), \sigma(u)) \in (0, 0.5) \times (-1, 1) \times \mathbb{R}^+$ para todo $u \in [0, 1]$ (parámetros de variación suave). Suponga que $\{u_t\} \stackrel{iid}{\sim} N(0, 1)$ y $\{\epsilon_t\} \stackrel{iid}{\sim} N(0, 1)$.

En este ejemplo la función de autocovarianza exacta es

$$\begin{aligned} \text{Cov}(Y_{t,T}, Y_{s,T}) &= \sigma \left(\frac{t}{T} \right) \sigma \left(\frac{s}{T} \right) \frac{\Gamma[1 - d(t/T) - d(s/T)] \Gamma[t - s + d(t/T)]}{\Gamma[1 - d(t/T)] \Gamma[d(t/T)] \Gamma[t - s + 1 - d(s/T)]} \\ &\times \left[1 + \vartheta \left(\frac{t}{T} \right) \vartheta \left(\frac{s}{T} \right) \frac{t - s - d(s/T)}{t - s - 1 + d(t/T)} - \vartheta \left(\frac{s}{T} \right) \frac{t - s + d(t/T)}{t - s + 1 - d(s/T)} \right] \\ &+ \sigma_\epsilon^2 \delta_0(t - s), \end{aligned}$$

para todo $t \geq s$.

La densidad espectral del ejemplo está dada por

$$f_Y(u, \lambda) = \frac{\sigma(u)^2}{2\pi} (1 - 2\vartheta(u) \cos(\lambda) + \vartheta(u)^2) \left(2 \sin \frac{\lambda}{2} \right)^{-2d(u)} + \frac{\pi}{4}.$$

3.2.1. Propiedades del proceso de retornos localmente estacionario $\{r_{t,T}\}$

A continuación se presentan algunas propiedades para el proceso $\{r_{t,T}\}$, cuando la secuencia $\{\nu_{t,T}\}$ corresponde a un proceso Gaussiano localmente estacionario de larga memoria.

(i) Los procesos $\{\sigma_{t,T}\}$ y $\{u_t\}$ son independientes, entonces

$$\mathbb{E}(r_{t,T}) = \mathbb{E}(\sigma_{t,T})\mathbb{E}(u_t) = 0.$$

(ii) $\mathbb{V}(r_{t,T}) = \mathbb{E}(\sigma_{t,T}^2 u_t^2) = \mathbb{E}(\sigma_{t,T}^2)\mathbb{E}(u_t^2) = \mathbb{E}(\sigma_{t,T}^2) = \sigma^2 \exp\{\mathbb{V}(\nu_{t,T})/2\}$. Por ejemplo, suponga que $\{\nu_{t,T}\} \sim \text{LS-FN}(d(u))$ con $\sigma(u) \equiv 1$, entonces

$$\mathbb{E}(r_{t,T}^2) = \sigma^2 \exp\left\{ \frac{1}{2} \frac{\Gamma[1 - 2d(t/T)]}{\Gamma[1 - d(t/T)]^2} \right\}.$$

(iii) Para todo $k > 0$ se tiene que

$$\text{Cov}(r_{t,T}, r_{t+k,T}) = \mathbb{E}(u_t)\mathbb{E}(u_{t+k})\mathbb{E}(\exp\{\nu_{t,T} + \nu_{t+k,T}\}) = 0.$$

(iv) La función de autocovarianza del proceso $\{r_{t,T}^2\}$ está dada por

$$\begin{aligned}
\text{Cov}(r_{t,T}^2, r_{s,T}^2) &= \sigma^2 [\mathbb{E}(\sigma_{t,T}^2 \sigma_{s,T}^2) - \mathbb{E}(\sigma_{t,T}^2) \mathbb{E}(\sigma_{s,T}^2)] \\
&= \sigma^2 [\mathbb{E}(\exp\{\nu_{t,T} + \nu_{s,T}\}) - \mathbb{E}(\exp\{\nu_{t,T}\}) \mathbb{E}(\exp\{\nu_{s,T}\})] \\
&= \sigma^2 \left[\exp\left\{ \text{Cov}(\nu_{t,T}, \nu_{s,T}) + \frac{1}{2} \mathbb{V}(\nu_{t,T}) + \frac{1}{2} \mathbb{V}(\nu_{s,T}) \right\} \right. \\
&\quad \left. - \exp\left\{ \frac{1}{2} \mathbb{V}(\nu_{t,T}) + \frac{1}{2} \mathbb{V}(\nu_{s,T}) \right\} \right] \\
&= \sigma^2 \exp\left\{ \frac{1}{2} (\mathbb{V}(\nu_{t,T}) + \mathbb{V}(\nu_{s,T})) \right\} \left[\exp\left\{ \text{Cov}(\nu_{t,T}, \nu_{s,T}) \right\} - 1 \right].
\end{aligned}$$

Si suponemos el ejemplo 3.1 (ii) y $t \geq s$, entonces

$$\begin{aligned}
\text{Cov}(r_{t,T}^2, r_{s,T}^2) &= \\
&\sigma^2 \exp \left\{ \frac{1}{2} \left(\sigma^2 \left(\frac{t}{T} \right) \frac{\Gamma[1 - 2d(t/T)]}{\Gamma[1 - d(t/T)]^2} + \sigma^2 \left(\frac{s}{T} \right) \frac{\Gamma[1 - 2d(s/T)]}{\Gamma[1 - d(s/T)]^2} \right) \right\} \\
&\times \left[\exp \left\{ \sigma \left(\frac{t}{T} \right) \sigma \left(\frac{s}{T} \right) \frac{\Gamma[1 - d(t/T) - d(s/T)] \Gamma[t - s + d(t/T)]}{\Gamma[1 - d(t/T)] \Gamma[d(t/T)] \Gamma[t - s + 1 - d(s/T)]} \right\} - 1 \right].
\end{aligned}$$

(v) La función de curtosis del proceso es

$$\mathbf{K}_{t,T} = \frac{\mathbb{E}(r_{t,T}^4)}{(\mathbb{E}(r_{t,T}^2))^2} = \frac{3\sigma^4 \exp\{2\mathbb{V}(\nu_{t,T})\}}{\sigma^4 \exp\{\mathbb{V}(\nu_{t,T})\}} = 3 \exp\{\mathbb{V}(\nu_{t,T})\} > 3.$$

Utilizando el ejemplo 3.1 (ii) tenemos

$$\mathbf{K}_{t,T} = 3 \exp \left\{ \sigma^2 \left(\frac{t}{T} \right) \frac{\Gamma[1 - 2d(t/T)]}{\Gamma[1 - d(t/T)]^2} \right\}.$$

Como ya se mencionó en el capítulo 2, el modelo propuesto, debe ser apropiado para los datos, y por lo tanto cumplir con los hechos estilizados de los retornos.

3.2.2. Estimación de modelos paramétricos a procesos de volatilidad localmente estacionarios de larga memoria

El mecanismo de estimación utilizado corresponde a la log verosimilitud local de Whittle que fue presentado en el capítulo 2, la diferencia se encuentra en los supuestos utilizados (que serán presentados más adelante). A continuación un breve repaso del mecanismo de estimación.

Considere a $\theta \in \Theta$ como el vector de los parámetros del proceso $\{Y_{t,T}\}$ en (3.9), donde el espacio paramétrico Θ es un subconjunto de un espacio Euclidiano finito dimensional. Dada una muestra $\{Y_{1,T}, \dots, Y_{t,T}\}$, el vector de parámetros estimado del proceso puede ser obtenido minimizando la función de log verosimilitud local de Whittle

$$\mathcal{L}_T(\theta) = \frac{1}{4\pi} \frac{1}{M} \int_{-\pi}^{\pi} \sum_{j=1}^M \left[\log f_{\theta}(u_j, \lambda) + \frac{I_N(u_j, \lambda)}{f_{\theta}(u_j, \lambda)} \right] d\lambda, \quad (3.11)$$

donde $f_{\theta}(u, \lambda)$ corresponde a la densidad espectral de variación en el tiempo del proceso $\{Y_{t,T}\}$ definida en (3.10) que se encuentra caracterizada por el parámetro θ , $I_N(u_j, \lambda) = \frac{|D_N(u, \lambda)|^2}{2\pi H_{2,N}(0)}$ es un periodograma local con taper,

$$D_N(u, \lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) Y_{[uT]-N/2+s+1, T} e^{-i\lambda s}, \quad H_{k,N} = \sum_{s=0}^{N-1} h^k\left(\frac{s}{N}\right) e^{-i\lambda s},$$

$T = S(M-1) + N$, $u_j = t_j/T$, $t_j = S(j-1) + N/2$, $j = 1, \dots, M$ y $h(\cdot)$ es la función taper para los datos. La intuición sobre el estimador de Whittle en (3.11) se puede revisar en el capítulo 2.

El estimador local de Whittle del vector de parámetro θ está dado por

$$\widehat{\theta}_T = \arg \min \mathcal{L}_T(\theta), \quad (3.12)$$

donde la minimización es bajo el espacio paramétrico Θ . A continuación se revisa un conjunto de supuestos que fueron definidos por Palma and Olea (2010) y que afectan al estimador y al proceso $\{r_{t,T}\}$. Posteriormente, se presentan las propiedades asintóticas de los estimadores.

Supuestos 3.1.

El primer supuesto hace referencia a la densidad espectral de variación en el tiempo del proceso $\{\nu_{t,T}\}$. El segundo supuesto está relacionado con la función Taper para los datos y el tercer supuesto se relaciona con el esquema de muestras por bloques. Se asume que el espacio paramétrico Θ es compacto y que θ_0 existe, es único y se encuentra en el interior de Θ . En los siguientes supuestos, K siempre será una constante positiva que será diferente línea a línea.

A1. La densidad espectral de variación en el tiempo de la secuencia $\{\nu_{t,T}\}$ en el proceso (3.9) es estrictamente positiva y satisface

$$f_{\nu,\theta}(u, \lambda) \sim \mathcal{C}_f(\theta, u)|\lambda|^{-2d_\theta(u)},$$

cuando $|\lambda| \rightarrow 0$, donde $\mathcal{C}_f(\theta, u) > 0$, $0 < \inf_{\theta,u} d_\theta(u), \sup_{\theta,u} d_\theta(u) < 1/2$ y $d_\theta(u)$ tiene primera derivada acotada con respecto a u . Existe una función integrable $g(\lambda)$ tal que $|\nabla_\theta \log f_{\nu,\theta}(u, \lambda)| \leq g(\lambda)$ para todo $\theta \in \Theta$, $u \in [0, 1]$ y $\lambda \in [-\pi, \pi]$. La función $|A_\nu(u, \lambda)|$ es dos veces diferenciable con respecto a u y satisface

$$\int_{-\pi}^{\pi} A_\nu(u, \lambda) A_\nu(v, -\lambda) \exp(ik\lambda) d\lambda \sim \mathcal{C}(\theta, u, v) k^{d_\theta(u)+d_\theta(v)-1} \quad (3.13)$$

cuando $k \rightarrow \infty$, donde $|\mathcal{C}(\theta, u, v)| \leq K$ para $u, v \in [0, 1]$ y $\theta \in \Theta$. La función $f_{\nu, \theta}(u, \lambda)^{-1}$ es dos veces diferenciable con respecto a θ , u y λ .

A2. El taper de los datos $h(u)$ es positivo, es una función acotada para $u \in [0, 1]$ y simétrico en $1/2$ con derivada acotada.

A3. Cuando el tamaño muestral T y las subdivisiones N , S y M tienden a infinito satisfacen $S/N \rightarrow 0$, $\sqrt{T} \log^2(N)/N \rightarrow 0$ y $N^3 \log^2(N)/T^2 \rightarrow 0$.

Adicionalmente, vamos a suponer que todos los momentos de $\{\epsilon_t\}$ existen y denotamosremos a $\kappa_4 = \text{cum}_4(\epsilon_t)/\sigma_\epsilon^4$ como el cumulante de 4to orden. Utilizaremos el taper presentado en el capítulo de corta memoria,

$$h(x) = \begin{cases} \frac{1}{2} [1 - \cos(2\pi x)] \cdot, & \text{si } x \in (0, 1), \\ 0, & \text{en otro caso.} \end{cases}$$

Bajo los supuestos 3.1 se obtienen dos propiedades fundamentales para el estimador (3.12), que son las consistencia y el teorema del límite central. Las demostraciones se encuentran en el apéndice C. Cabe mencionar que la variación de las demostraciones está dada por el no acotamiento de la densidad espectral, es por esto que para las demostraciones se utilizó el coeficiente de Fourier de manera directa.

Teorema 3.1. *Sea θ_0 el verdadero valor del parámetro θ . Bajo los supuestos 3.1, el estimador $\hat{\theta}_T$ satisface $\hat{\theta}_T \rightarrow \theta_0$, en probabilidad, cuando $T \rightarrow \infty$.*

Teorema 3.2. *Sea θ_0 el verdadero valor del parámetro θ . Si se tiene los supuestos 3.1, entonces el estimador de Whittle $\hat{\theta}_T$ satisface el teorema del límite central, es decir,*

$$\sqrt{T} \left(\hat{\theta}_T - \theta_0 \right) \rightarrow N \left(0, \Gamma(\theta_0)^{-1} [\Gamma(\theta_0) + \Lambda(\theta_0)] \Gamma(\theta_0)^{-1} \right),$$

en distribución, cuando $T \rightarrow \infty$. Donde

$$\begin{aligned}\Gamma(\theta_0) &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla \log f_{\theta_0}(u, \lambda) [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du. \\ \Lambda(\theta_0) &= \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu du.\end{aligned}$$

Observación 3.1. Recuerde que aquí suponemos que el proceso $\{\nu_{t,T}\}$ es Gaussiano, por lo tanto hay una diferencia con en el teorema del límite central del estimador en el capítulo 2, dado que el triespectrum para procesos Gaussianos es $g_{\nu,4}(\lambda, -\lambda, \mu) = 0$.

3.2.3. Selección de Modelos

Para construir un índice de selección de modelos, se utilizará la misma deducción del capítulo 2. Primero se define la función de distancia de Kullback Leibler entre la densidad espectral del proceso verdadero $f(u, \lambda)$ y del modelo propuesto $f_\theta(u, \lambda)$

$$\mathcal{L}(\theta) := \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \log f_\theta(u, \lambda) + \frac{f(u, \lambda)}{f_\theta(u, \lambda)} d\lambda du$$

Se asume que la mejor aproximación del valor del parámetro es

$$\theta_0 := \arg \min_{\theta \in \Theta} \mathcal{L}(\theta).$$

Si el modelo es el correcto y suponemos que la densidad espectral del verdadero proceso es $f_{\theta^*}(u, \lambda)$, entonces $\theta_0 = \theta^*$. El índice que mide la calidad del ajuste será $\mathbb{E}\mathcal{L}(\widehat{\theta}_T)$, donde $\widehat{\theta}_T$ es el estimador de Whittle en (3.12). Para deducir una estimación del índice primero se realizan dos expansiones cuadráticas, una de $\mathcal{L}(\theta)$ alrededor de θ_0 y otra de $\mathcal{L}_T(\theta)$ alrededor de $\widehat{\theta}_T$,

$$\mathcal{L}(\hat{\theta}_T) \approx \mathcal{L}(\theta_0) + \frac{1}{2} (\hat{\theta}_T - \theta_0)' \nabla^2 \mathcal{L}(\theta_0) (\hat{\theta}_T - \theta_0), \quad (3.14)$$

$$\mathcal{L}_T(\theta_0) \approx \mathcal{L}_T(\hat{\theta}_T) + \frac{1}{2} (\hat{\theta}_T - \theta_0)' \nabla^2 \mathcal{L}_T(\hat{\theta}_T) (\hat{\theta}_T - \theta_0). \quad (3.15)$$

Por las propiedades del estimador de Whittle $\hat{\theta}_T$ (ver apéndice C) se tiene que $\mathcal{L}_T(\theta_0)$ es un estimador consistente, $\nabla^2 \mathcal{L}(\theta_0) = \Gamma(\theta_0)$ y $\nabla^2 \mathcal{L}_T(\hat{\theta}_T) \rightarrow \Gamma(\theta_0)$ en probabilidad. Entonces el estimador de $\mathbb{E}\mathcal{L}(\hat{\theta}_T)$ se puede construir como

$$\begin{aligned} \widehat{\mathbb{E}}\mathcal{L}(\hat{\theta}_T) &\approx \mathcal{L}_T(\hat{\theta}_T) + \mathbb{E} \left[(\hat{\theta}_T - \theta_0)' \nabla^2 \mathcal{L}(\theta_0) (\hat{\theta}_T - \theta_0) \right] \\ &\approx \mathcal{L}_T(\hat{\theta}_T) + \frac{p}{T} + \frac{1}{T} \text{Tr} [\Gamma(\theta_0)^{-1} (\Lambda(\theta_0))], \end{aligned}$$

donde p es el número de parámetros y las matrices $\Gamma(\theta_0)$ y $\Lambda(\theta_0)$ y se encuentran en el teorema 3.2. Finalmente multiplicamos por 2 y reemplazamos θ_0 por $\hat{\theta}_T$

$$\text{AIC} = 2\mathcal{L}_T(\hat{\theta}_T) + \frac{2p}{T} + \frac{2}{T} \text{Tr} [\Gamma(\hat{\theta}_T)^{-1} (\Lambda(\hat{\theta}_T))].$$

3.3. Experimento de Monte Carlo

Primer escenario de simulación: Para este experimento se considera un ruido fraccionado localmente estacionario $\{\nu_{t,T}\} \sim \text{LS-FN}(d(u), \sigma(u))$, es decir,

$$\nu_{t,T} = \sigma(t/T)(1 - B)^{-d(t/T)} \varepsilon_t,$$

donde $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$, $d(u) \in (0, 0.5)$ para todo $u \in [0, 1]$ y $\{u_t\} \stackrel{iid}{\sim} N(0, 1)$. El proceso $\{r_{t,T}\}$ está definido por las siguientes ecuaciones

$$\begin{aligned}
r_{t,T} &= \sigma_{t,T} u_t, \\
\sigma_{t,T} &= \exp(\nu_{t,T}/2).
\end{aligned}
\tag{3.16}$$

Por lo anterior, la densidad espectral del proceso $\log r_{t,T}^2$ está dada por

$$f_{\theta}(u, \lambda) = \frac{\sigma(u)^2}{2\pi} [2(1 - \cos \lambda)]^{-d(u)} + \frac{\pi}{4}.$$

Para ser consistente con el experimento del capítulo 2, también se realizan 1000 repeticiones del proceso (3.16) para estudiar el comportamiento del estimador (3.12).

Los tamaños muestrales utilizados en este experimento son $T = 512, 1024, 2048, 4096$ y para estos valores se consideran saltos $S = 20, 30, 40$ y largo de los bloques $N = 128, 256$. Los parámetros $d(u)$ y $\sigma(u)$ estudiados son:

$$\begin{aligned}
d(u) &= \beta_0 + \beta_1 u, \\
\sigma(u) &= \alpha_0.
\end{aligned}$$

Para la simulación del proceso $\{\nu_{t,T}\}$ se utilizó el algoritmo de innovaciones (ver página 172 de Brockwell and Davis (1991)), dado que en el trabajo de Palma and Olea (2010) se mostró empíricamente que este operaba igual que el algoritmo recursivo, pero era mucho más eficiente en tiempo de computo.

Observación 3.2. Las desviaciones asintóticas necesitan los momentos poblacionales de la distribución $\log \chi_1^2$, estas pueden ser consultadas en el capítulo 2, en la sección de experimento de Monte Carlo.

A continuación se presentan las tablas con los resultados de las simulaciones.

Tabla 3.1: Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 128$ y salto $S = 20$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.15	0.15	0.143	0.123	1.975	0.128	0.225	0.197	0.108	0.178	0.193
	0.25	0.15	0.203	0.112	1.984	0.136	0.230	0.210	0.100	0.163	0.198
	0.35	-0.10	0.316	-0.078	1.991	0.133	0.231	0.206	0.099	0.169	0.196
	0.20	0.20	0.203	0.155	1.993	0.126	0.215	0.200	0.102	0.167	0.197
1024	0.15	0.15	0.139	0.136	1.989	0.098	0.175	0.133	0.077	0.126	0.137
	0.25	0.15	0.182	0.149	1.994	0.100	0.173	0.144	0.071	0.115	0.140
	0.35	-0.10	0.330	-0.093	1.992	0.095	0.166	0.142	0.070	0.120	0.138
	0.20	0.20	0.183	0.202	1.995	0.095	0.160	0.144	0.072	0.118	0.139
2048	0.15	0.15	0.133	0.154	1.996	0.069	0.120	0.092	0.054	0.089	0.097
	0.25	0.15	0.187	0.147	1.994	0.068	0.116	0.098	0.050	0.081	0.099
	0.35	-0.10	0.337	-0.104	2.002	0.067	0.118	0.096	0.050	0.085	0.098
	0.20	0.20	0.181	0.206	1.998	0.070	0.111	0.095	0.051	0.083	0.098
4096	0.15	0.15	0.133	0.155	1.995	0.049	0.083	0.066	0.038	0.063	0.068
	0.25	0.15	0.182	0.159	1.998	0.048	0.080	0.068	0.035	0.058	0.070
	0.35	-0.10	0.338	-0.103	2.002	0.045	0.080	0.066	0.035	0.060	0.069
	0.20	0.20	0.182	0.206	2.002	0.048	0.079	0.068	0.036	0.059	0.070

Tabla 3.2: Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 128$ y salto $S = 30$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.15	0.15	0.142	0.122	1.975	0.127	0.229	0.200	0.108	0.178	0.193
	0.25	0.15	0.202	0.111	1.984	0.135	0.235	0.213	0.100	0.163	0.198
	0.35	-0.10	0.316	-0.077	1.992	0.133	0.236	0.208	0.099	0.169	0.196
	0.20	0.20	0.203	0.153	1.993	0.125	0.221	0.201	0.102	0.167	0.197
1024	0.15	0.15	0.139	0.136	1.989	0.097	0.175	0.133	0.077	0.126	0.137
	0.25	0.15	0.183	0.148	1.994	0.100	0.173	0.143	0.071	0.115	0.140
	0.35	-0.10	0.330	-0.093	1.992	0.094	0.166	0.142	0.070	0.120	0.138
	0.20	0.20	0.184	0.201	1.995	0.095	0.160	0.144	0.072	0.118	0.139
2048	0.15	0.15	0.133	0.154	1.996	0.068	0.119	0.092	0.054	0.089	0.097
	0.25	0.15	0.187	0.147	1.994	0.067	0.116	0.098	0.050	0.081	0.099
	0.35	-0.10	0.337	-0.104	2.002	0.067	0.117	0.096	0.050	0.085	0.098
	0.20	0.20	0.181	0.206	1.998	0.070	0.110	0.095	0.051	0.083	0.098
4096	0.15	0.15	0.133	0.155	1.995	0.049	0.083	0.066	0.038	0.063	0.068
	0.25	0.15	0.182	0.159	1.998	0.048	0.080	0.068	0.035	0.058	0.070
	0.35	-0.10	0.338	-0.103	2.002	0.045	0.080	0.066	0.035	0.060	0.069
	0.20	0.20	0.182	0.206	2.002	0.048	0.079	0.068	0.036	0.059	0.070

Tabla 3.3: Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 128$ y salto $S = 40$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.15	0.15	0.142	0.123	1.976	0.125	0.226	0.197	0.108	0.178	0.193
	0.25	0.15	0.202	0.111	1.985	0.134	0.233	0.211	0.100	0.163	0.198
	0.35	-0.10	0.317	-0.078	1.992	0.132	0.233	0.207	0.099	0.169	0.196
	0.20	0.20	0.202	0.155	1.993	0.124	0.219	0.200	0.102	0.167	0.197
1024	0.15	0.15	0.138	0.137	1.989	0.096	0.170	0.133	0.077	0.126	0.137
	0.25	0.15	0.182	0.149	1.994	0.098	0.168	0.143	0.071	0.115	0.140
	0.35	-0.10	0.330	-0.094	1.993	0.093	0.162	0.140	0.070	0.120	0.138
	0.20	0.20	0.183	0.202	1.995	0.094	0.157	0.143	0.072	0.118	0.139
2048	0.15	0.15	0.133	0.155	1.996	0.068	0.119	0.092	0.054	0.089	0.097
	0.25	0.15	0.186	0.147	1.994	0.067	0.115	0.097	0.050	0.081	0.099
	0.35	-0.10	0.337	-0.104	2.002	0.067	0.117	0.095	0.050	0.085	0.098
	0.20	0.20	0.181	0.206	1.998	0.069	0.110	0.095	0.051	0.083	0.098
4096	0.15	0.15	0.133	0.155	1.995	0.049	0.083	0.066	0.038	0.063	0.068
	0.25	0.15	0.182	0.159	1.998	0.048	0.080	0.068	0.035	0.058	0.070
	0.35	-0.10	0.338	-0.103	2.002	0.045	0.079	0.066	0.035	0.060	0.069
	0.20	0.20	0.182	0.206	2.002	0.048	0.078	0.068	0.036	0.059	0.070

Se puede observar que las desviaciones de Monte Carlo son similares en las tres tablas, donde la mayor precisión muestral se aprecia en el valor de salto $S = 40$ (muy similar al $S = 20$), sin embargo estas diferencias son despreciable. Al analizar el sesgo en las estimaciones puntuales, no se observan grandes diferencias, logrando el menor sesgo cuando el tamaño del saltos es $S = 40$. Como conclusión de las 3 tablas los resultados fueron buenos, ya que las estimaciones puntuales tienen un leve sesgo y las desviaciones de Monte Carlo se asemejan a las teóricas.

Recuerde que estas conclusiones están sujetas a la elección de los parámetros, y no son imputables a otras configuraciones. El ejercicio que se presenta a continuación es similar, pero se considera un tamaño de ventana $N = 256$

Tabla 3.4: Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 256$ y salto $S = 20$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.15	0.15	0.162	0.093	1.972	0.146	0.264	0.228	0.108	0.178	0.193
	0.25	0.15	0.216	0.095	1.983	0.152	0.268	0.237	0.100	0.163	0.198
	0.35	-0.10	0.309	-0.054	1.994	0.151	0.269	0.238	0.099	0.169	0.196
	0.20	0.20	0.220	0.125	1.996	0.142	0.260	0.229	0.102	0.167	0.197
1024	0.15	0.15	0.147	0.132	1.992	0.104	0.190	0.139	0.077	0.126	0.137
	0.25	0.15	0.189	0.147	1.994	0.105	0.180	0.150	0.071	0.115	0.140
	0.35	-0.10	0.336	-0.092	1.992	0.101	0.179	0.149	0.070	0.120	0.138
	0.20	0.20	0.192	0.196	1.997	0.098	0.165	0.151	0.072	0.118	0.139
2048	0.15	0.15	0.138	0.156	1.997	0.068	0.121	0.096	0.054	0.089	0.097
	0.25	0.15	0.194	0.144	1.996	0.067	0.117	0.100	0.050	0.081	0.099
	0.35	-0.10	0.342	-0.101	2.004	0.066	0.114	0.099	0.050	0.085	0.098
	0.20	0.20	0.188	0.203	2.000	0.068	0.110	0.099	0.051	0.083	0.098
4096	0.15	0.15	0.140	0.153	1.996	0.046	0.078	0.067	0.038	0.063	0.068
	0.25	0.15	0.189	0.157	2.000	0.045	0.076	0.069	0.035	0.058	0.070
	0.35	-0.10	0.344	-0.103	2.003	0.042	0.075	0.066	0.035	0.060	0.069
	0.20	0.20	0.190	0.203	2.003	0.046	0.075	0.069	0.036	0.059	0.070

Tabla 3.5: Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 256$ y salto $S = 30$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.15	0.15	0.161	0.095	1.974	0.145	0.261	0.225	0.108	0.178	0.193
	0.25	0.15	0.216	0.096	1.984	0.150	0.266	0.235	0.100	0.163	0.198
	0.35	-0.10	0.310	-0.057	1.994	0.149	0.266	0.236	0.099	0.169	0.196
	0.20	0.20	0.219	0.128	1.995	0.140	0.256	0.226	0.102	0.167	0.197
1024	0.15	0.15	0.147	0.132	1.992	0.103	0.190	0.139	0.077	0.126	0.137
	0.25	0.15	0.189	0.147	1.994	0.105	0.181	0.150	0.071	0.115	0.140
	0.35	-0.10	0.336	-0.092	1.992	0.100	0.180	0.149	0.070	0.120	0.138
	0.20	0.20	0.192	0.195	1.997	0.098	0.166	0.151	0.072	0.118	0.139
2048	0.15	0.15	0.138	0.156	1.997	0.068	0.121	0.096	0.054	0.089	0.097
	0.25	0.15	0.194	0.144	1.996	0.067	0.117	0.100	0.050	0.081	0.099
	0.35	-0.10	0.342	-0.101	2.004	0.066	0.114	0.099	0.050	0.085	0.098
	0.20	0.20	0.188	0.203	2.000	0.067	0.110	0.099	0.051	0.083	0.098
4096	0.15	0.15	0.140	0.153	1.996	0.045	0.078	0.067	0.038	0.063	0.068
	0.25	0.15	0.189	0.157	2.000	0.045	0.075	0.069	0.035	0.058	0.070
	0.35	-0.10	0.344	-0.103	2.003	0.042	0.074	0.066	0.035	0.060	0.069
	0.20	0.20	0.190	0.203	2.003	0.046	0.075	0.069	0.036	0.059	0.070

Tabla 3.6: Estimación de Whittle: $\alpha_0 = 2$, tamaño de ventana $N = 256$ y salto $S = 40$

T	Parámetros		Estimación			SD Empírica			SD Teórica		
	β_0	β_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}_0$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\beta_0)$	$\widehat{\sigma}(\beta_1)$	$\widehat{\sigma}(\alpha_0)$
512	0.15	0.15	0.160	0.098	1.974	0.143	0.258	0.223	0.108	0.178	0.193
	0.25	0.15	0.215	0.098	1.984	0.149	0.263	0.234	0.100	0.163	0.198
	0.35	-0.10	0.311	-0.059	1.994	0.147	0.264	0.233	0.099	0.169	0.196
	0.20	0.20	0.218	0.131	1.995	0.138	0.253	0.223	0.102	0.167	0.197
1024	0.15	0.15	0.146	0.134	1.992	0.102	0.185	0.138	0.077	0.126	0.137
	0.25	0.15	0.189	0.148	1.994	0.103	0.175	0.149	0.071	0.115	0.140
	0.35	-0.10	0.336	-0.093	1.992	0.098	0.175	0.148	0.070	0.120	0.138
	0.20	0.20	0.191	0.198	1.997	0.096	0.162	0.150	0.072	0.118	0.139
2048	0.15	0.15	0.138	0.156	1.997	0.068	0.121	0.096	0.054	0.089	0.097
	0.25	0.15	0.194	0.144	1.996	0.067	0.117	0.100	0.050	0.081	0.099
	0.35	-0.10	0.342	-0.101	2.004	0.066	0.114	0.099	0.050	0.085	0.098
	0.20	0.20	0.188	0.203	2.000	0.067	0.110	0.099	0.051	0.083	0.098
4096	0.15	0.15	0.140	0.153	1.996	0.045	0.077	0.067	0.038	0.063	0.068
	0.25	0.15	0.189	0.157	2.000	0.045	0.075	0.069	0.035	0.058	0.070
	0.35	-0.10	0.344	-0.103	2.003	0.042	0.074	0.066	0.035	0.060	0.069
	0.20	0.20	0.190	0.203	2.003	0.046	0.075	0.069	0.036	0.059	0.070

Aquí se observa que para muestras de tamaño $T = 512, 1024$ si se aumenta el tamaño del salto S se aumenta levemente la precisión muestral del estimador. El sesgo cometido no varía por la elección de S .

Si comparamos las desviaciones muestrales de los estimadores en estas 3 tablas ($N = 256$) y las comparamos con las desviaciones de las 3 tablas anteriores ($N = 128$), podemos notar que para tamaños muestrales pequeños las desviaciones de Monte Carlo son menores cuando la ventana es de $N = 128$ y para tamaños muestrales mayores, las desviaciones de Monte Carlo son menores cuando el tamaño de la ventana es $N = 256$.

Considerando los resultados de este experimento, el estimador de Whittle es bastante bueno, ya que el proceso es una secuencia latente no observada con un ruido aditivo.

Segundo escenario de simulación: Se considera que el proceso latente $\{\nu_{t,T}\}$ que genera la volatilidad está definido por una secuencia ARFIMA($d(u), \vartheta(u), \sigma(u)$) localmente estacionario,

$$\nu_{t,T} = \sigma \left(\frac{t}{T} \right) (1 - \vartheta(t/T)B) (1 - B)^{-d(t/T)} \varepsilon_t,$$

para $t = 1, \dots, T$ con $d(u) = \beta_0 + \beta_1 u$, $\sigma(u) = \alpha_0 + \alpha_1 u$, $\vartheta(u) = \vartheta$ y $\{\varepsilon_t\} \stackrel{iid}{\sim} N(0, 1)$.

Con la especificación anterior y suponiendo que la secuencia de ruido $\{u_t\} \stackrel{iid}{\sim} N(0, 1)$ en (3.7), se tiene que la densidad espectral de $\{\log r_{t,T}^2\}$ es

$$f_\theta(u, \lambda) = \frac{\sigma(u)^2}{2\pi} (1 - 2\vartheta \cos \lambda + \vartheta^2) \left(2 \sin \frac{\lambda}{2} \right)^{-2d(u)} + \frac{\pi}{4}.$$

Este estudio de Monte Carlo considera 1000 repeticiones del estimador por bloques de Whittle, con la función taper de Tukey-Hanning.

En este experimento se presentan los resultados para una elección de S y N fijos (similar al capítulo 2). La elección del S y N será aquella que minimice el error cuadrático medio para alguna configuración específica θ_0 que sea representativo con los otros parámetros presentados en las tablas.

Sea $\theta = (\beta_0, \beta_1, \vartheta, \alpha_0, \alpha_1)$ el vector de parámetros del modelo. A continuación se presenta la curva de contorno del error cuadrático medio MSE del estimador de Whittle $\hat{\theta}$, como el promedio de 100 repeticiones de la distancia euclidiana $\|\hat{\theta} - \theta_0\|^2$ considerando θ_0 el verdadero valor del parámetro, para diferentes valores de S y N . Para esta curva se consideró $\theta_0 = (0.20, 0.25, 0.5, 1.5, 2.0)$ y un tamaño muestral de $T = 512$

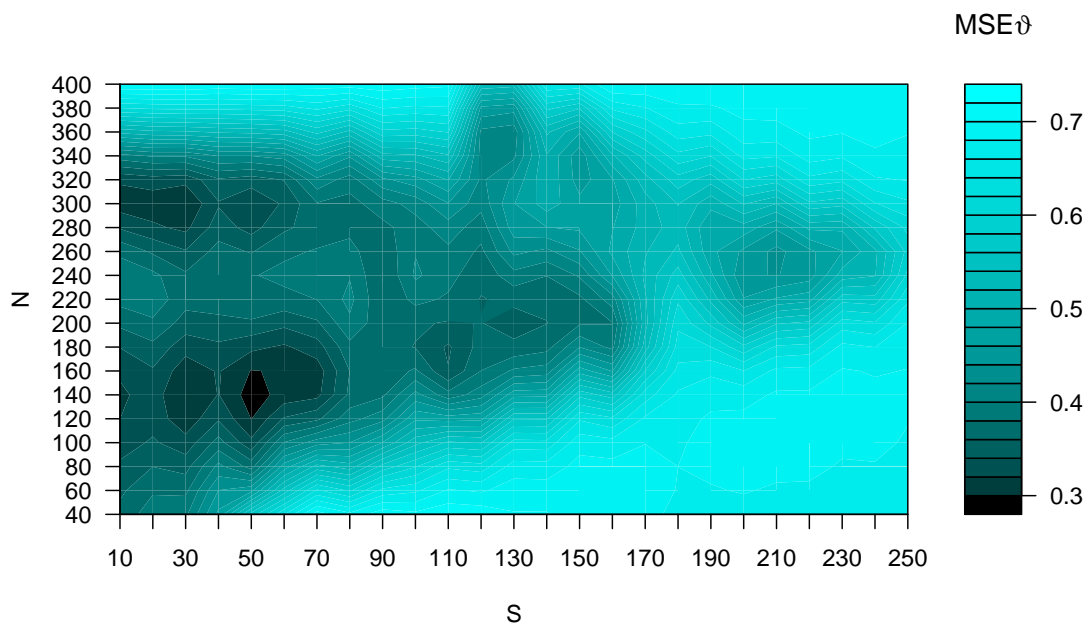


Figura 3.2: Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARFIMA con $T = 512$.

Los valores S y N que minimizan el error son $S \approx 45$ y $N \approx 130$. Si se compara esta curva de contorno con la presentada en el capítulo 2 (para $T = 512$), se puede observar que los valores de S y N que minimizan el MSE son mayores aquí, esto se debe a que el orden de magnitud de los estimadores es menor cuando $\{\nu_{t,T}\}$ es un proceso localmente estacionario de corta memoria, y por lo tanto, el estimador de Whittle converge a mayor velocidad. El orden de magnitud de los estimadores se encuentra en el apéndice.

A continuación se presentan los resultados de la simulación de Monte Carlo, utilizando los valores S y N que minimizan el MSE.

Tabla 3.7: Estimación de Whittle: tamaño muestral $T = 512$, tamaño de ventana $N = 130$ y salto $S = 40$

Caso	Parámetros					Estimaciones				
	β_0	β_1	ϑ	α_0	α_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\vartheta}$	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$
1	0.15	0.20	0.50	3.00	1.50	0.120	0.161	0.459	2.980	1.490
2	0.15	0.20	0.50	4.50	-1.50	0.125	0.160	0.468	4.459	-1.500
3	0.20	0.25	0.50	6.00	-2.00	0.160	0.205	0.448	5.973	-2.024
4	0.20	0.25	0.50	1.50	2.00	0.16	0.186	0.436	1.481	2.013
5	0.40	-0.15	0.50	2.00	1.50	0.313	-0.117	0.443	1.971	1.500
6	0.40	-0.15	0.50	7.00	-2.00	0.340	-0.139	0.449	6.987	-2.028
7	0.15	0.20	-0.40	3.00	1.50	0.151	0.175	-0.414	2.956	1.515
8	0.15	0.20	-0.40	4.50	-1.50	0.145	0.175	-0.424	4.457	-1.497
9	0.20	0.25	-0.40	6.00	-2.00	0.193	0.222	-0.418	5.954	-1.992
10	0.20	0.25	-0.40	1.50	2.00	0.195	0.211	-0.457	1.454	1.987
11	0.40	-0.15	-0.40	2.00	1.50	0.364	-0.118	-0.462	1.965	1.438
12	0.40	-0.15	-0.40	7.00	-2.00	0.378	-0.132	-0.411	6.968	-2.002

Caso	SD Teórica					SD Empírica				
	$\sigma(\widehat{\beta}_0)$	$\sigma(\widehat{\beta}_1)$	$\sigma(\widehat{\vartheta})$	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\vartheta})$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$
1	0.152	0.165	0.148	0.346	0.608	0.137	0.189	0.136	0.293	0.489
2	0.149	0.159	0.147	0.358	0.597	0.137	0.190	0.138	0.300	0.501
3	0.130	0.138	0.126	0.412	0.674	0.128	0.162	0.124	0.352	0.601
4	0.18	0.205	0.175	0.358	0.600	0.141	0.197	0.144	0.278	0.446
5	0.151	0.173	0.164	0.336	0.577	0.162	0.226	0.153	0.289	0.480
6	0.124	0.133	0.121	0.460	0.754	0.134	0.182	0.126	0.441	0.765
7	0.091	0.147	0.092	0.362	0.635	0.111	0.195	0.117	0.388	0.722
8	0.088	0.147	0.092	0.380	0.634	0.107	0.191	0.114	0.437	0.754
9	0.083	0.137	0.076	0.436	0.723	0.093	0.158	0.079	0.428	0.747
10	0.115	0.182	0.151	0.384	0.634	0.119	0.196	0.194	0.332	0.557
11	0.102	0.168	0.132	0.381	0.611	0.121	0.212	0.203	0.401	0.666
12	0.080	0.131	0.069	0.490	0.797	0.100	0.177	0.080	0.547	0.975

Al observar los resultados presentados en la tabla 3.7, uno puede notar un leve sesgo en el parámetro de larga memoria para lagunas configuraciones, sin embargo, un tamaño $T = 512$ puede ser considerado pequeño dado que el proceso es localmente estacionario de larga memoria y tiene un ruido de manera aditiva. Las desviaciones de Monte Carlo de las estimaciones se comportan similar a las teóricas.

Considerando el largo de las trayectorias y los argumentos presentados, el estimador de Whittle presenta un buen desempeño aún para muestras pequeñas.

La proxima figura corresponde a la curva de contorno del error cuadrático medio MSE, cuando el tamaño muestral de la serie es $T = 1024$. Para la ilustración se consideró la misma configuración paramétrica $\theta_0 = (0.20, 0.25, 0.5, 1.5, 2.0)$ y 100 repeticiones.

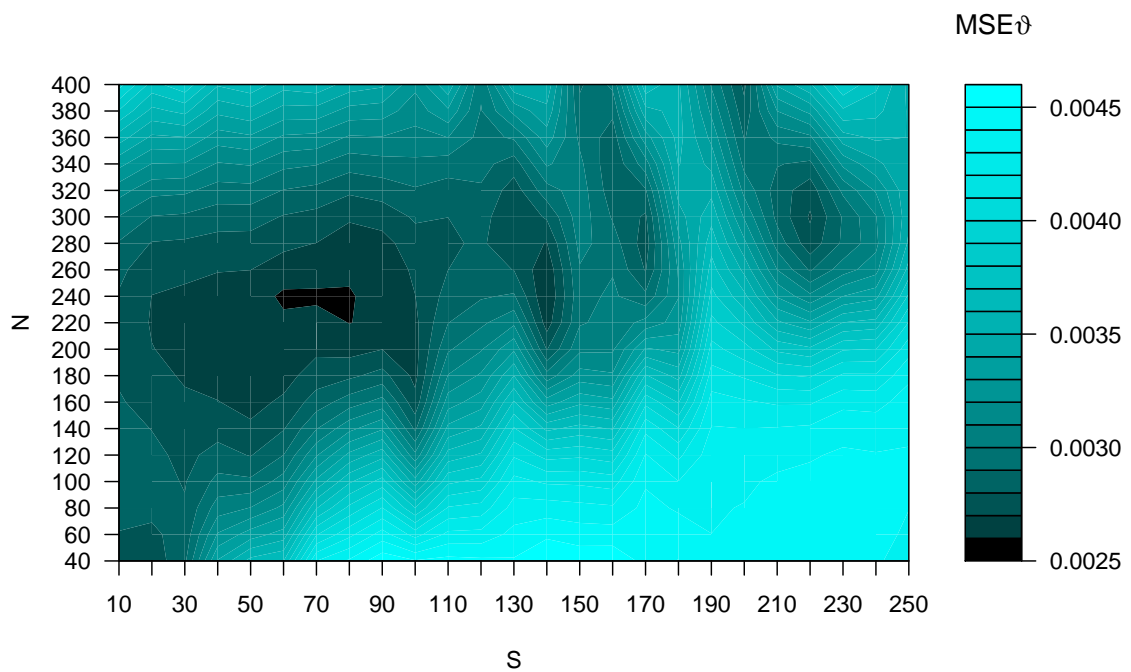


Figura 3.3: Curva de contorno del MSE empírico del estimador de Whittle para el proceso LS-ARFIMA con $T = 1024$.

Los valores que minimizan el error cuadrático medio son $S \approx 60$ y $N \approx 230$. El objetivo de este gráfico es ilustrar de manera empírica una elección para S y N , en la práctica de deben evaluar otras herramientas. Los resultados obtenidos para otras configuraciones fueron similares y dado el gasto computacional, solo se presenta el gráfico para un valor específico de θ_0 (que es representativo).

Tabla 3.8: Estimación de Whittle: tamaño muestral $T = 1024$, tamaño de ventana $N = 230$ y salto $S = 60$

Caso	Parámetros					Estimaciones				
	β_0	β_1	ϑ	α_0	α_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\vartheta}$	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$
1	0.15	0.20	0.50	3.00	1.50	0.126	0.186	0.474	2.984	1.500
2	0.15	0.20	0.50	4.50	-1.50	0.128	0.189	0.476	4.486	-1.508
3	0.20	0.25	0.50	6.00	-2.00	0.166	0.231	0.456	5.996	-2.018
4	0.20	0.25	0.50	1.50	2.00	0.163	0.225	0.454	1.490	2.004
5	0.40	-0.15	0.50	2.00	1.50	0.343	-0.134	0.451	1.994	1.488
6	0.40	-0.15	0.50	7.00	-2.00	0.363	-0.145	0.468	6.980	-2.004
7	0.15	0.20	-0.40	3.00	1.50	0.145	0.191	-0.413	2.974	1.503
8	0.15	0.20	-0.40	4.50	-1.50	0.142	0.199	-0.410	4.460	-1.474
9	0.20	0.25	-0.40	6.00	-2.00	0.196	0.233	-0.411	5.964	-1.969
10	0.20	0.25	-0.40	1.50	2.00	0.198	0.232	-0.429	1.488	1.987
11	0.40	-0.15	-0.40	2.00	1.50	0.387	-0.147	-0.431	1.981	1.464
12	0.40	-0.15	-0.40	7.00	-2.00	0.391	-0.150	-0.410	6.969	-1.998

Caso	SD Teórica					SD Empírica				
	$\sigma(\widehat{\beta}_0)$	$\sigma(\widehat{\beta}_1)$	$\sigma(\widehat{\vartheta})$	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\vartheta})$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$
1	0.107	0.117	0.104	0.245	0.430	0.122	0.144	0.117	0.223	0.393
2	0.105	0.112	0.104	0.253	0.422	0.122	0.145	0.120	0.242	0.413
3	0.092	0.097	0.089	0.291	0.476	0.111	0.113	0.099	0.280	0.482
4	0.12	0.145	0.124	0.253	0.424	0.127	0.150	0.119	0.217	0.346
5	0.107	0.122	0.116	0.237	0.408	0.133	0.172	0.132	0.232	0.403
6	0.088	0.094	0.086	0.325	0.533	0.106	0.130	0.101	0.340	0.595
7	0.064	0.104	0.065	0.256	0.449	0.083	0.135	0.077	0.308	0.552
8	0.062	0.104	0.065	0.268	0.449	0.079	0.141	0.071	0.308	0.534
9	0.058	0.097	0.054	0.308	0.511	0.070	0.114	0.057	0.312	0.541
10	0.081	0.129	0.106	0.272	0.448	0.091	0.141	0.138	0.251	0.444
11	0.072	0.119	0.093	0.269	0.432	0.086	0.150	0.131	0.306	0.500
12	0.056	0.093	0.048	0.346	0.564	0.068	0.123	0.056	0.415	0.723

Si se analiza la tabla anterior y la comparamos con la tabla 3.7 (cuando $T = 512$), podemos notar que se reduce el sesgo del estimador para el parámetro de larga memoria y que las desviaciones de Monte Carlo son más cercanas las teóricas, entonces el tamaño de muestra parece ser más apropiado.

Como conclusión general de ambas tablas, el estimador tiene un buen desempeño, sin embargo se debe tener en cuenta el tamaño muestral, ya que en estos procesos el largo de la trayectoria debe ser mayor que en el caso de corta memoria.

En todos los experimentos de Monte Carlo se asumió conocido el parámetro σ_ϵ^2 . Si se desea estimar el parámetro σ_ϵ^2 es recomendado es utilizar una reparametrización de la densidad espectral y minimizar la función de log verosimilitud condicional.

Para presentar el procedimiento vamos a asumir que la densidad espectral se puede reparametrizar de la siguiente manera

$$\begin{aligned} f_Y(u, \lambda) &= f_\nu(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi} \\ &= \sigma(u)^2 f_0(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi} \\ &= \frac{\sigma_\epsilon^2}{2\pi} \left[\frac{2\pi}{\sigma_\epsilon^2} \sigma(u)^2 f_0(u, \lambda) + 1 \right] \\ &= \sigma^2 [\psi(u)^2 f_0(u, \lambda) + 1]. \end{aligned}$$

Con esta reparametrización, la función objetivo a minimizar es

$$\tilde{\mathcal{L}}(\vartheta) = \frac{1}{4} \log(\hat{\sigma}^2) + \frac{1}{4\pi} \frac{1}{M} \int_{-\pi}^{\pi} \sum_{j=1}^M \log(\psi(u_j)^2 f_0(u_j, \lambda) + 1)$$

donde

$$\hat{\sigma}^2 = \frac{1}{\pi} \frac{1}{M} \int_{-\pi}^{\pi} \sum_{j=1}^M \frac{I_N(u_j, \lambda)}{\psi(u_j)^2 f_0(u_j, \lambda) + 1}.$$

El vector de parametros estimados es $\hat{\vartheta} = (\hat{\vartheta}, \hat{\sigma}^2)'$, donde $\hat{\vartheta}$ es el argumento que minimiza $\tilde{\mathcal{L}}$. En el caso estacionario se realizar un procedimiento similar, que puede ser consultado en Hurvich and Ray (2003).

Como ilustración del comportamiento del estimador con este procedimiento, se considerará $\{\nu_t, T\} \sim \text{LS-FN}(d(u), \sigma(u))$ y $\{u_t\} \stackrel{iid}{\sim} N(0, 1)$, de esta forma, la densidad espectral del proceso será

$$f_{\theta}(u, \lambda) = \frac{\sigma(u)^2}{2\pi} [2(1 - \cos \lambda)]^{-d(u)} + \frac{\sigma_{\epsilon}^2}{2\pi}$$

En este experimento se va a el comportamiento del estimador cuando se estima σ_{ϵ}^2 . Vamos a utilizar los mismos valores de que se obtuvieron en el segundo experimento para S y N . Los parámetros son configurados como $d(u) = \beta_0 + \beta_1 u$ y $\sigma(u) = \alpha_0 + \alpha_1$.

A continuación los resultados de la simulación de Monte Carlo

Tabla 3.9: Estimación de Whittle: tamaño muestral $T = 1024$, tamaño de ventana $N = 230$ y salto $S = 60$ con $\sigma_{\epsilon} = \pi/\sqrt{2} \approx 2.221$

Caso	Parámetros					Estimaciones				
	β_0	β_1	σ_{ϵ}	α_0	α_1	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\sigma}_{\epsilon}$	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$
1	0.15	0.20	2.221	3.00	1.50	0.155	0.181	1.901	3.100	1.522
2	0.15	0.20	2.221	4.50	-1.50	0.140	0.193	1.860	4.628	-1.472
3	0.20	0.25	2.221	6.00	-2.00	0.190	0.229	1.679	6.173	-1.942
4	0.20	0.25	2.221	1.50	2.00	0.164	0.241	1.777	1.808	2.019
5	0.40	-0.15	2.221	2.00	1.50	0.353	-0.12	1.840	2.292	1.396
6	0.40	-0.15	2.221	7.00	-2.00	0.383	-0.136	1.754	7.132	-2.004
7	0.15	0.20	2.221	1.50	3.00	0.135	0.179	1.695	1.766	3.096
8	0.25	0.20	2.221	2.00	2.00	0.209	0.209	1.747	2.304	1.973

Caso	SD Teórica					SD Empírica				
	$\sigma(\widehat{\beta}_0)$	$\sigma(\widehat{\beta}_1)$	$\sigma(\widehat{\sigma}_{\epsilon})$	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\sigma}_{\epsilon})$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$
1	0.106	0.098	1.149	1.223	0.521	0.120	0.149	1.099	0.851	0.618
2	0.060	0.129	1.162	0.705	0.563	0.079	0.148	1.036	0.532	0.596
3	0.055	0.113	1.151	0.650	0.603	0.072	0.122	0.922	0.441	0.608
4	0.126	0.123	0.961	0.915	0.572	0.146	0.175	0.771	0.691	0.606
5	0.119	0.126	0.852	0.749	0.477	0.118	0.155	0.886	0.642	0.532
6	0.075	0.093	1.519	0.978	0.600	0.073	0.132	1.091	0.571	0.706
7	0.144	0.142	1.534	1.311	0.898	0.147	0.182	0.860	0.790	0.714
8	0.110	0.110	1.090	0.844	0.507	0.129	0.165	0.845	0.649	0.546

En la tabla 3.9 podemos notar que la estimación de σ_{ϵ} tiene un leve sesgo, sin embargo esto no es una sorpresa, ya que en el caso estacionario se ocurre un patrón similar. Al observar las estimaciones de los parámetros de larga memoria y de escala del proceso $\nu_{t,T}$, se tiene un comportamiento similar al observado en las tablas

anteriores, un muy leve sesgo que se compensa entre parámetros. Las desviaciones muestrales para estas configuraciones paramétricas se aproximan a las teóricas, y los casos donde se observa mayor diferencia es cuando la estimación de $\hat{\sigma}_\epsilon$ presenta mayor sesgo.

Capítulo 4

Ilustración

En este capítulo se utilizará la metodología propuesta en el capítulo 3 para procesos de volatilidad estocástica de larga memoria localmente estacionarios en el análisis de un índice real. La ilustración consiste en utilizar la serie del índice bursátil S&P500 a nivel semanal.

Se propondrá un modelo apropiado para los datos. No es el espíritu establecer este modelo como el único aplicable a esta serie. Lo que se busca es ofrecer una clase de modelos más flexibles que permitan considerar cambios suaves en la estructura paramétrica.

4.1. Índice Bursátil S&P500

El índice bursátil Standard & Poor's 500 (o abreviado S&P500) es uno de los índices globales más importantes ya que representa la situación real del mercado de los Estados Unidos. Este índice considera a las 500 empresas que más acciones cotizan en las bolsas de NASDAQ (bolsa de valores electrónica y automatizada) y NYSE (bolsa de New York).

La siguiente figura presenta la serie del índice S&P500 y de los retornos logarítmicos a nivel semanal sobre el valor ajustado de cierre (fuente de información fue el sitio web <http://finance.yahoo.com/>) desde el 3 de Enero de 1950 hasta 2 de Febrero del año 2007.

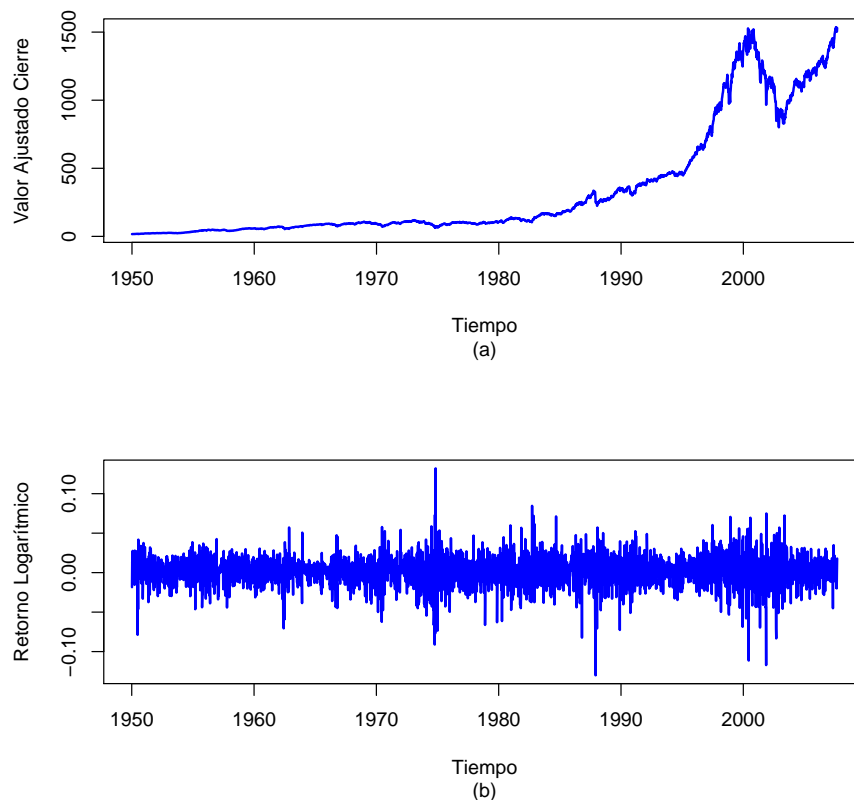


Figura 4.1: (a) Índice S&P500 considerando valor de cierre ajustado (b) Retornos logarítmicos de la serie (a).

Lo primero que se debe hacer antes de aplicar un modelo de volatilidad estocástica localmente estacionario es identificar si la serie de retornos presenta autocorrelación, para esto se calcula la función ACF y PACF del proceso.

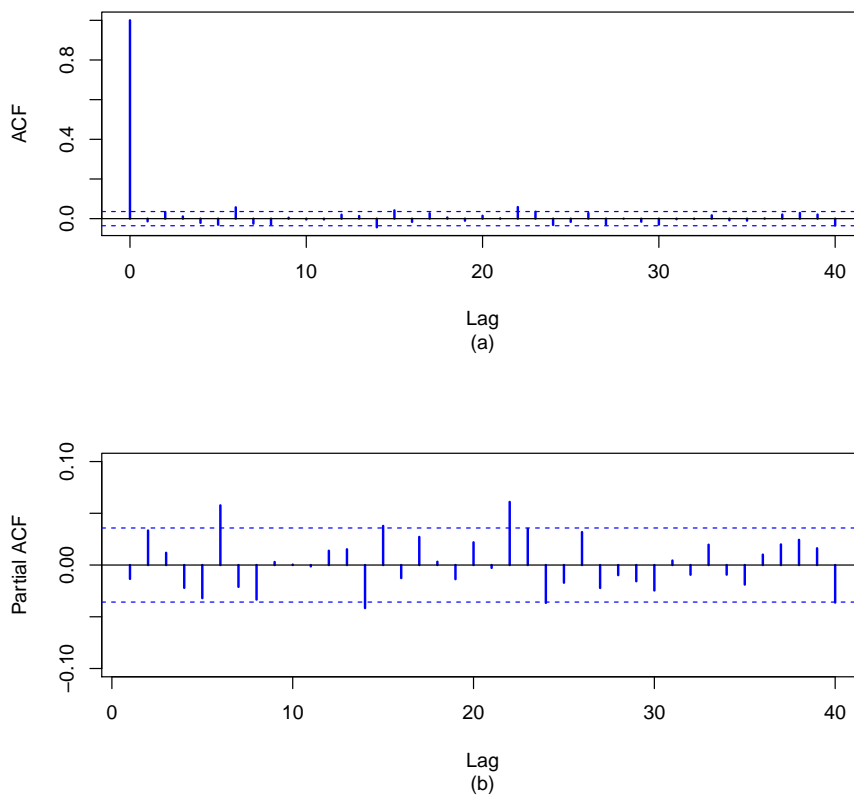


Figura 4.2: (a) Función de autocorrelación muestral (b) Función de autocorrelación parcial muestral.

Los retornos de la serie deben ser un secuencia de ruido blanco, sin embargo de la figura 4.2 podemos ver que existen algunos rezagos que pueden alterar el cumplimiento de este supuesto. Para verificar se utilizará el test de “Ljung Box” en diferentes lag.

La siguiente tabla presenta el test de “Ljung Box” aplicado a los retornos semanales en los rezagos 5, 10, 15, 20 y 25.

Tabla 4.1: Test de Ljung Box sobre la serie de retornos

Rezagos	Estadístico	P Valor
5	8.439	0.134
10	22.718	0.012
15	35.525	0.002
20	39.612	0.006
25	57.539	< 0.001

A partir de la tabla 4.1 se puede ver que los retornos de la serie no cumplen con el supuesto de ruido blanco, es por esto que se ajustará un modelo de la familia ARMA para remover la correlación serial. Se evaluaron diferentes configuraciones, de estos el modelo que resultó más adecuado por elección de modelos AIC y cumplimiento de ruido blanco fue

$$r_t = z_t + \psi_2 e_{t-2} + \psi_6 e_{t-6} + \psi_{15} e_{t-15} + \psi_{22} e_{t-22}, \quad \{e_t\} \sim RB(0, \sigma_z^2). \quad (4.1)$$

Los parámetros estimados son:

Tabla 4.2: Parámetros estimados

Coefficiente	Estimación	Error Estándar	P valor
ψ_2	0.0443	0.0176	0.0121
ψ_6	0.0621	0.0176	0.0004
ψ_{15}	0.0449	0.0188	0.0167
ψ_{22}	0.0642	0.0176	0.0002

Con $\hat{\sigma}_z^2 = 0.00038$ y $AIC = 7551.7$. A partir del modelo (4.1) se calculan los residuos, los cuales se encuentran alineados con los “Hechos estilizados” mencionados en la introducción.

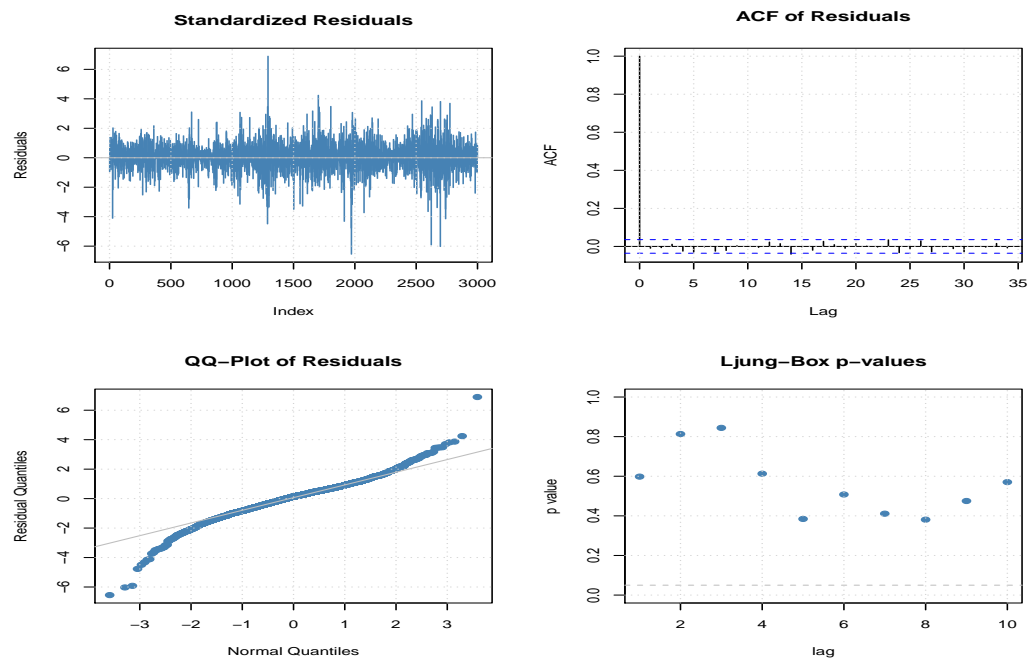


Figura 4.3: Diagnóstico del modelo.

El diagnóstico de los residuos $\{r_t - \hat{r}_t\}$ del modelo ajustado está presentado en la figura 4.3. De los gráficos de ACF residual y “Ljung Box” podemos concluir que los residuos son no autocorrelacionados y que no se rechazó el test de blancura de “Ljung Box”. Sin embargo, los datos son semanales, por lo cual se debe tener cuidado con posibles efectos estacionales que no son observables en el diagnóstico gráfico.

Para tener un diagnóstico más amplio se presenta un gráfico con los valores-p del test de blancura de “Ljung Box” para valores mayores de 10.

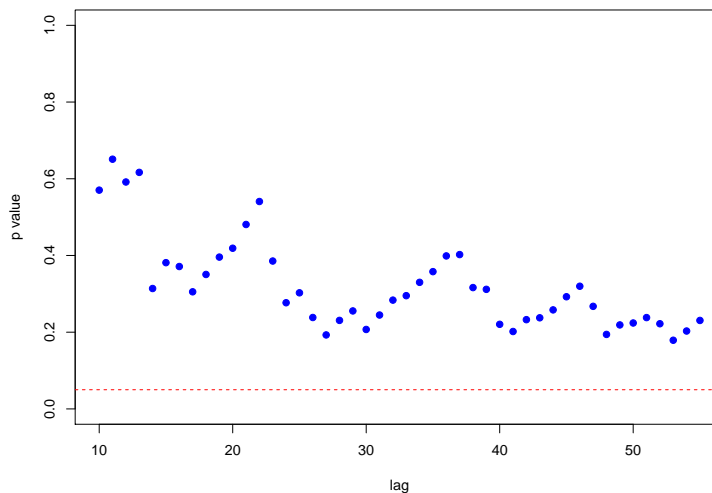


Figura 4.4: Gráfico del p valor del test de "Ljung Box".

De la figura 4.4 se puede concluir que no hay efecto estacional inadvertido por el primer diagnóstico. De esta manera, los residuos del ajuste parecen ser adecuados para aplicar un modelo de volatilidad estocástica localmente estacionario.

En primer lugar se definirá la secuencia de residuos del modelo por $\{z_t := r_t - \hat{r}_t\}$. Luego, como se mencionó en capítulos anteriores, se realizará un reescalamiento en el índice de la serie $\{z_t\}$ para utilizar la nomenclatura estándar de procesos localmente estacionarios, es decir, se denotará a la secuencia residuos del proceso de retorno por $\{z_{t,T}\}$.

Como primer análisis se graficará la serie $\{z_{t,T}^2\}$ y su función de autocorrelación.

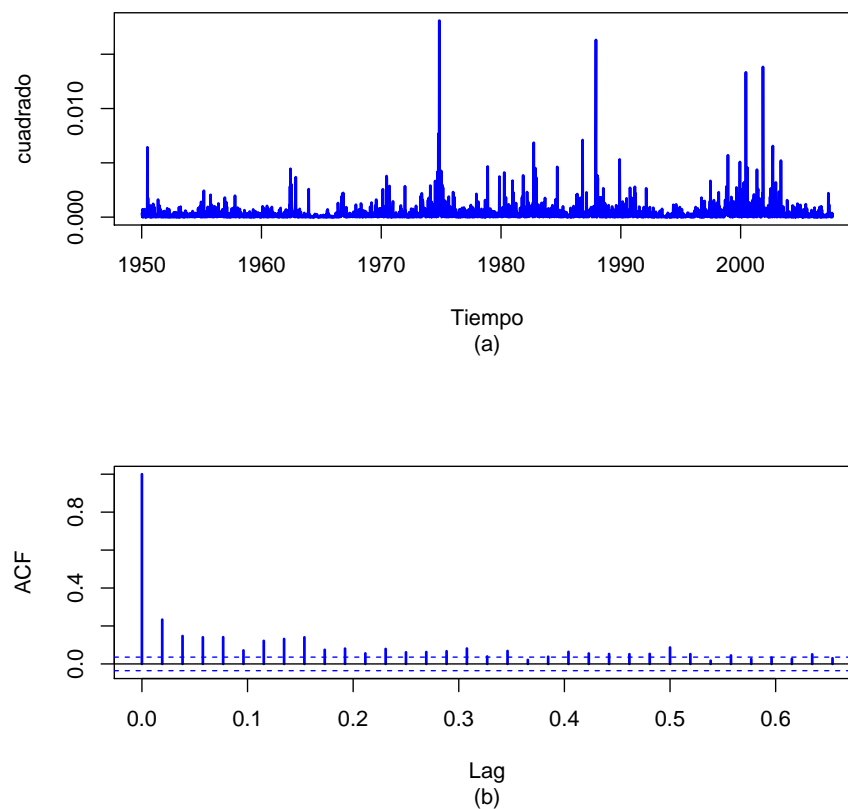


Figura 4.5: (a) Cuadrado de los residuos (b) Función de autocorrelación muestral.

Notamos de la figura 4.5 que la escala de la secuencia $\{z_{t,T}^2\}$ es muy baja, y la autocorrelación presenta un claro grado de persistencia, lo que es muy común al analizar este tipo de índices.

En los modelos propuestos en el capítulo 2 y 3, la función de autocorrelación del proceso se ve afectada por la incertidumbre del error, ya que a mayor dispersión en el ruido aditivo, menor será el nivel en la autocorrelación total (efecto que también aplica a los procesos en su versión estacionaria).

Como segundo análisis se graficará la función de autocorrelación de $\{z_{t,T}^2\}$ para diferentes tramos de la serie.

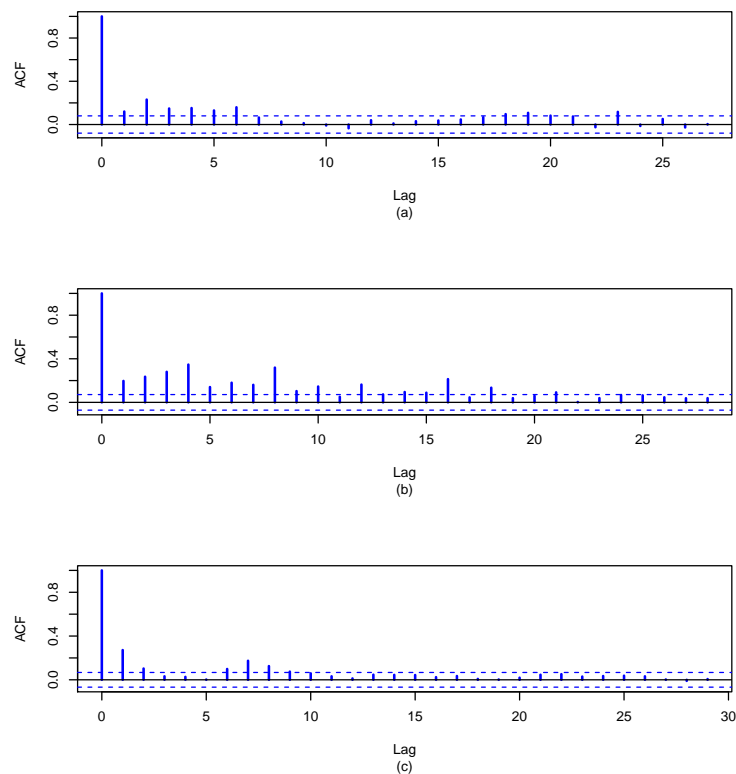


Figura 4.6: Función de autocorrelación (a) intervalo del 26/11/1951 al 27/05/1963 (b) intervalo del 11/05/1964 al 25/09/1978 (c) intervalo del 10/09/1979 al 26/12/1995.

De la figura 4.6 se puede apreciar que la estructura de dependencia cambia a través del tiempo, mostrando la mayor correlación en el intervalo de tiempo de 11/05/1964 al 25/09/1978, de esta manera resulta razonable considerar un proceso de larga memoria.

En esta clase de procesos se trabaja con la transformación logaritmo natural al cuadrado de la serie $\{\log z_{t,T}^2\}$, la que se encuentra en la figura 4.7.

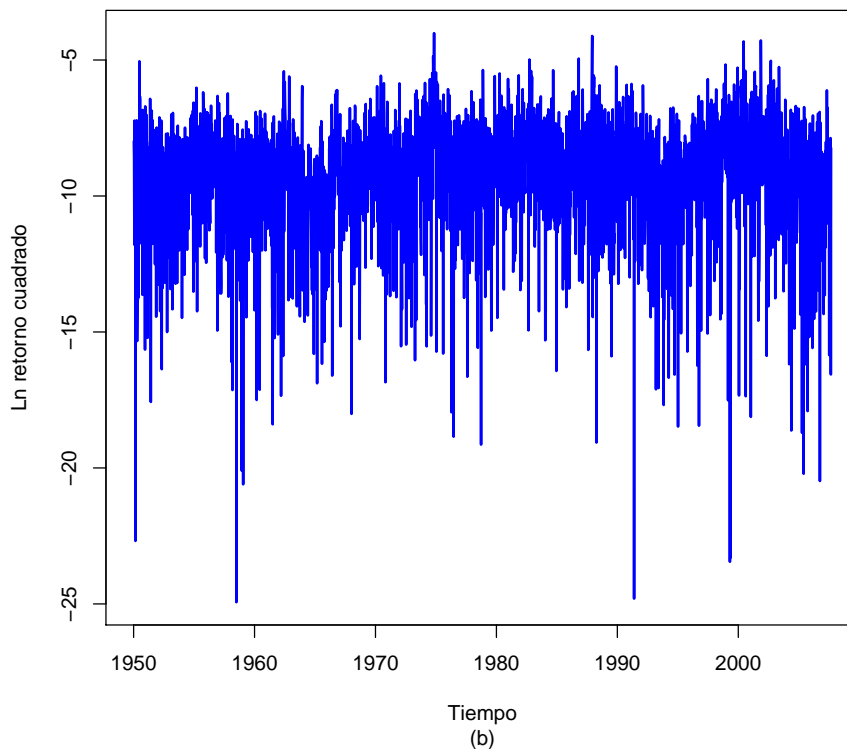


Figura 4.7: Gráfico de la serie transformada.

Note que la escala de la secuencia $\{\log z_{t,T}^2\}$ es alta, esto se debe al ajuste inicial sobre los retornos de la serie, ya que los residuos tenían un parámetro de dispersión muy bajo y al aplicar la transformación logaritmo sobre el cuadrado aumenta la escala de la variable.

Para los mismos segmentos presentados en la figura 4.6 se calculará la función de autorrelación sobre la serie transformada $\{\log z_{t,T}^2\}$, ya que los mecanismos de estimación presentados en esta tesis se basan en un cambio suave sobre los parámetros identificados en la transformación.

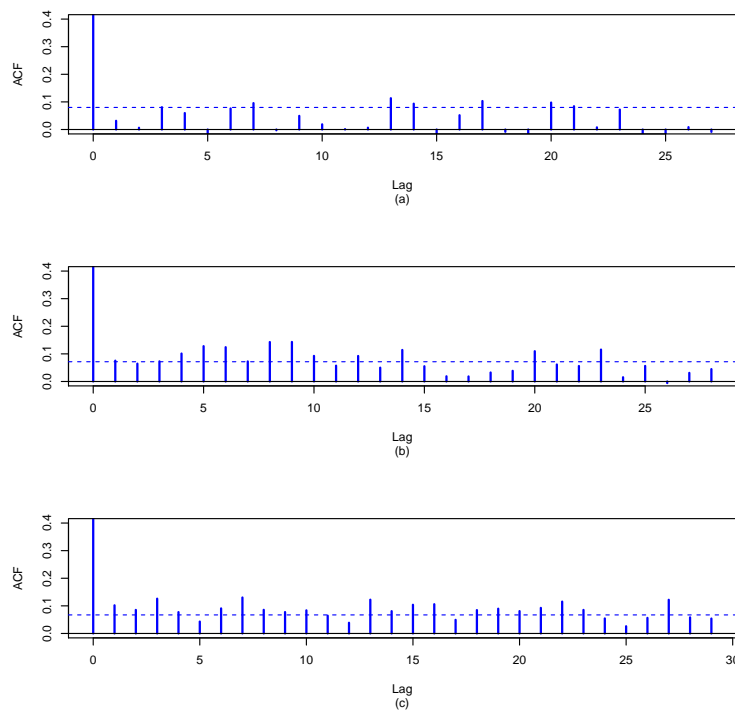


Figura 4.8: Función de autocorrelación (a) intervalo del 26/11/1951 al 27/05/1963 (b) intervalo del 11/05/1964 al 25/09/1978 (c) intervalo del 10/09/1979 al 26/12/1995.

El mismo fenómeno que se observó con la función de autocorrelación muestral de $\{z_{t,T}^2\}$ se observa en la figura 4.8, es decir, hay un cambio en la dependencia del proceso $\{\log z_{t,T}^2\}$ a través del tiempo.

Para esta ilustración se consideró una clase de procesos de volatilidad estocástica localmente estacionarios para $\{\nu_{t,T}\}$ de la forma LS-ARFIMA(0, $d(u)$, $\vartheta_1(u)$, $\vartheta_2(u)$, $\sigma(u)$) con las siguiente configuración paramétrica:

$$d(u) = \beta_0 + \beta_1 u + \cdots + \beta_p u^p,$$

$$\vartheta_1(u) = \vartheta_1,$$

$$\vartheta_2(u) = \vartheta_2,$$

$$\sigma(u) = \alpha_0 + \alpha_1 u + \cdots + \alpha_q u^q.$$

Para comprender de mejor manera la evolución de los parámetros involucrados en la volatilidad del modelo, se realizó una estimación por bloques de tamaño $N = 750$ con saltos $S = 200$ entre cada bloque, y se consideró en cada bloque a la serie como un proceso de volatilidad estacionario, es decir, en cada uno de ellos se ajustó un proceso ARFIMA(0, d , 2). La siguiente figura muestra la evolución de los parámetros

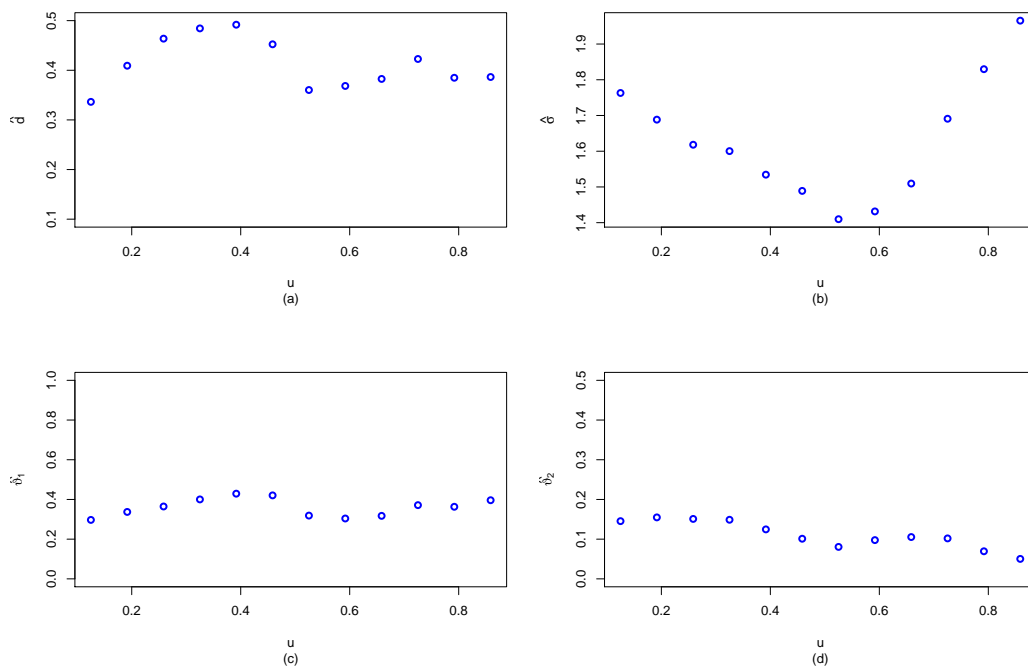


Figura 4.9: Estimación de los parámetros por bloque.

El modelo propuesto para los datos tiene la siguiente función de densidad espectral

$$f_{\theta}(u, \lambda) = \frac{\sigma(u)^2}{2\pi} (1 + \vartheta_1^2 + \vartheta_2^2 - 2\vartheta_1(1 - \vartheta_2) \cos \lambda - 2\vartheta_2 \cos 2\lambda) \left(2 \sin \frac{\lambda}{2}\right)^{-2d(u)} + \frac{\sigma_{\epsilon}^2}{2\pi}.$$

Para la elección del modelo se utilizó el índice AIC desarrollado en el capítulo 2.

$$\text{AIC} = 2\mathcal{L}_T(\hat{\theta}_T) + \frac{2p}{T} + \frac{2}{T} \text{Tr} \left[\Gamma(\hat{\theta}_T)^{-1} \left(\Lambda(\hat{\theta}_T) \right) \right].$$

De los modelos evaluados el que entregó mejor bondad de ajuste y mejor índice AIC fue

Tabla 4.3: Parámetros estimados

Coficiente	Estimación	Error Estándar	valor p
β_0	0.1301	0.0965	0.1777
β_1	2.2132	0.8475	0.0090
β_2	-4.5371	1.8224	0.0128
β_3	2.6938	1.1299	0.0171
ϑ_1	0.3584	0.0803	< 0.0001
ϑ_2	0.1106	0.0328	0.0007
α_0	2.1415	0.3096	< 0.0001
α_1	-2.9224	0.9997	0.0035
α_2	3.1500	0.9882	0.0014
σ_{ϵ}	1.3944	0.3906	0.0003

El índice AIC fue $\text{AIC} = 0.3589$. El parámetro β_0 parece ser no significativo, sin embargo se optó por considerarlo, ya que generó más estabilidad y el excluirlo en la estimación empeoraba las estimaciones de los otros coeficientes en el polinomio asociado al parámetros de larga memoria.

El siguiente gráfico muestra el comportamiento de los parámetros del proceso de volatilidad estocástica localmente estacionario y las estimaciones anteriores por bloques.

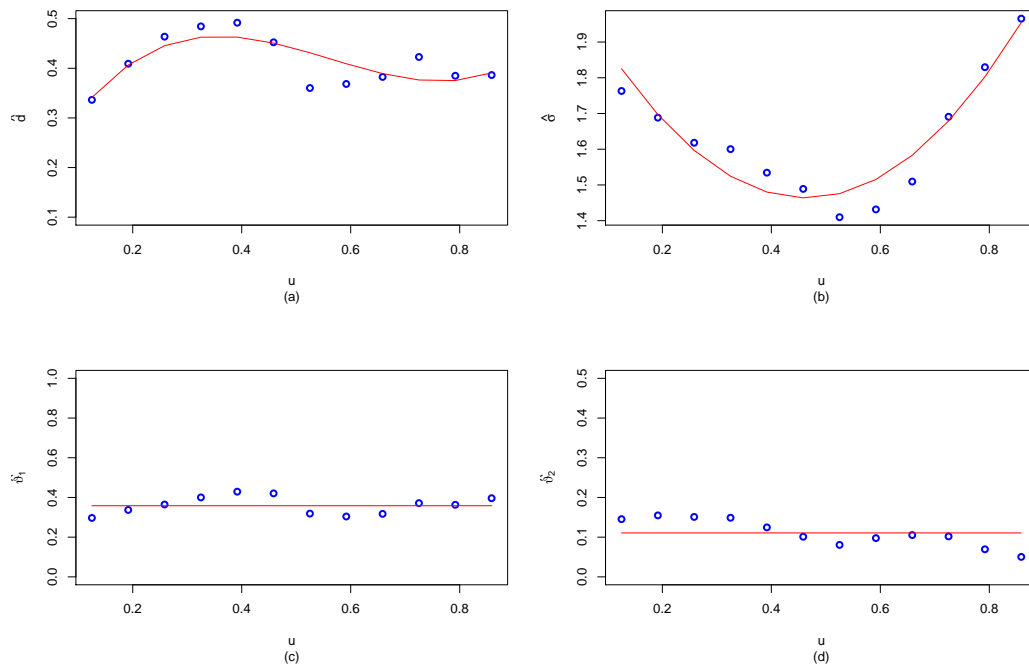


Figura 4.10: Gráfico de la evolución de los parámetros estimados.

En la figura 4.10 se puede ver la evolución de los parámetros estimados del proceso de volatilidad estocástica localmente estacionario y se puede apreciar que ésta logra capturar la tendencia por las estimaciones por bloques. Este proceso permite cambios suaves en la estructura paramétrica, y lo hace más flexible. En los gráficos presentados anteriormente vimos que en ciertas ventanas de la serie había un aumento y disminución en el grado de dependencia, efecto que sí es capturado por este ajuste.

Una vez ajustado el modelo, se utilizó el mecanismo de predicción descrito en la

sección 3 de Fryzlewicz et al. (2003), es decir, dada t observaciones de un proceso localmente estacionario $Y_{1,T}, \dots, Y_{t,T}$, el mejor predictor lineal a h pasos de $Y_{t+h,T}$ está dado por

$$\hat{Y}_{t+h,T} = \sum_{s=1}^t b_{t-1-s,T}^{(h)} Y_{s,T},$$

donde los coeficientes $b_{t-1-s,T}^{(h)}$ son aquellos que minimizan el error cuadrático medio de predicción (ECMP). El ECMP se define como

$$\text{ECMP}(\hat{Y}_{t-1+h,T}, Y_{t-1+h,T}) = \mathbb{E}(\hat{Y}_{t-1+h,T} - Y_{t-1+h,T})^2.$$

En esta ilustración se considera a $Y_{t,T} = \log z_{t,Y}^2 - \mu$, donde μ es la media constante del proceso y los coeficientes $b_{t-1-s,T}^{(h)}$ son estimados por el modelo ajustado.

La siguiente figura muestra la serie $\log z_{t,T}^2$ y su predicción a un paso.

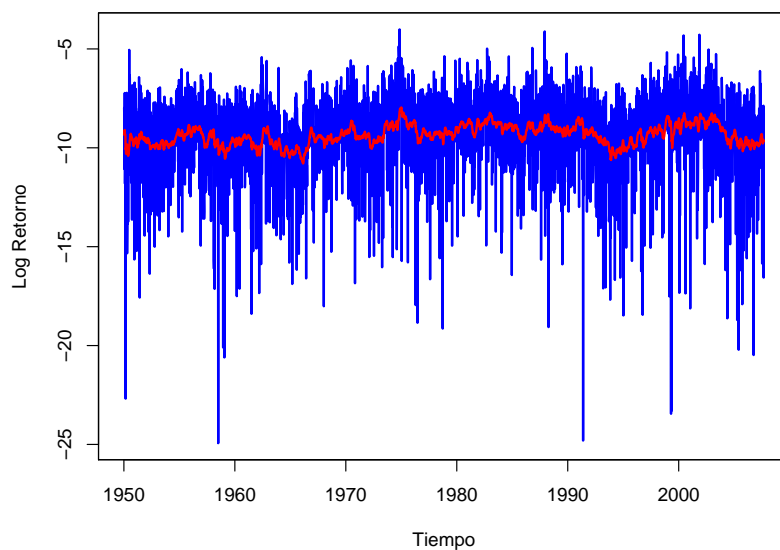


Figura 4.11: Serie $\log z_{t,T}^2$ y predicción a un paso.

A continuación se presenta el diagnóstico sobre los residuos del modelo

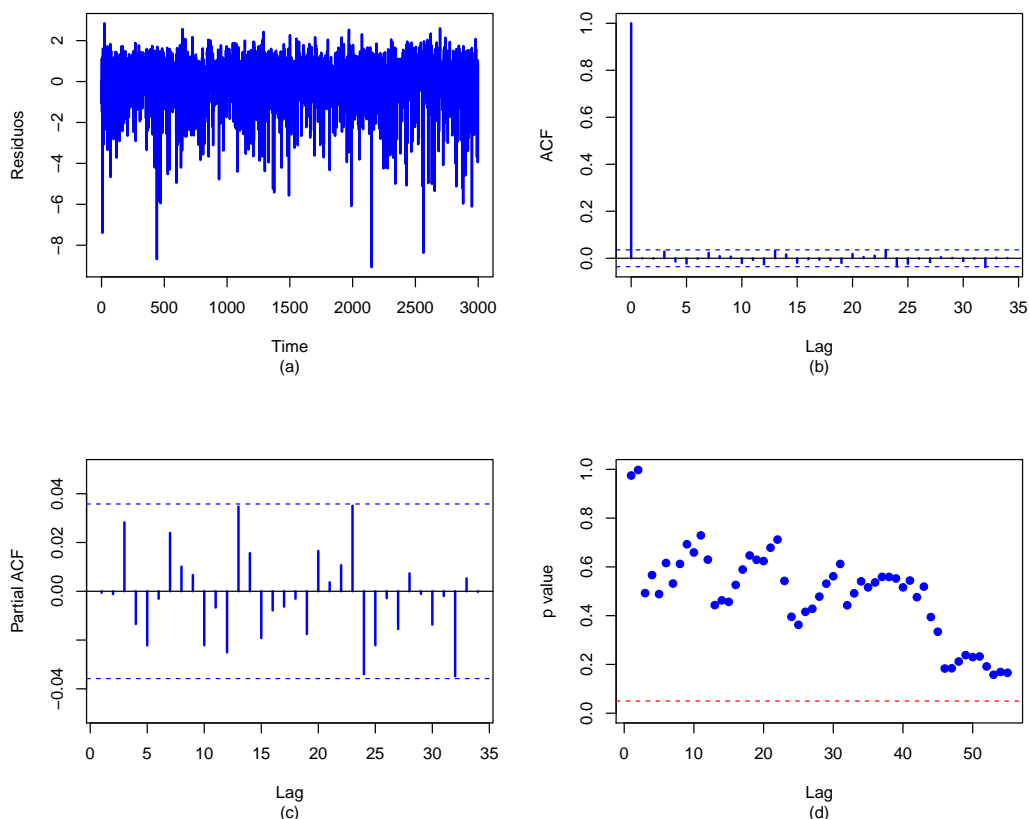


Figura 4.12: (a) Gráfico de los Residuos (b) Función de autocorrelación muestral (c) Función de autocorrelación parcial muestral (d) Gráfico de los valores p del test de “Ljung Box”.

Los residuos están centrados en el cero con un bajo grado de asimetría negativo, lo que es esperable dada la transformación $\log(\cdot)$. Los gráficos de autorrelación y autocorrelación parcial evidencian que la correlación serial fue removida, lo que es validado por los valores p del test de “Ljung Box”.

Como conclusión de la ilustración, el modelo propuesto para los datos es adecuado según los diagnósticos residuales y permite considerar una clase de procesos muy

amplia, que además es más flexible a lo disponible en la literatura. Note que la clase de modelo de volatilidad estocástica estacionaria está anidada por esta extensión.

Capítulo 5

Conclusiones y Trabajos Futuros

En esta tesis se ha introducido una nueva clase de procesos de volatilidad estocástica, donde la secuencia latente no observada que genera la volatilidad, es un proceso localmente estacionario, abordando dos casos: dependencia débil y fuerte. Para esta nueva clase se presentan algunas propiedades de momento, que son fundamentales para la verificación de los hechos estilizados sobre retornos financieros.

El mecanismo de estimación presentado, corresponde a la log verosimilitud local de Whittle, que tiene la ventaja de ser computacionalmente eficiente, es decir, es un algoritmo rápido al momento de su ejecución y además fácil de programar. Para los estimadores, se presentaron dos propiedades estadísticas relevantes, la consistencia y el teorema del límite central, que permiten tomar una decisión al momento de ajustar un modelo. Para verificar el comportamiento del estimador en muestras finitas, se desarrollaron diferentes estudios de simulación de Monte Carlo, donde se analizó el sesgo y la precisión, los resultados fueron favorable, considerando el aumento en la estructura paramétrica y el ruido aditivo presente.

Al revisar los trabajos desarrollados sobre procesos de volatilidad localmente estacionarios, encontramos que la mayoría están orientados a modelar la volatilidad

de manera determinística, utilizando procesos LS-ARCH(p) y LS-ARCH(∞), sin embargo, las restricciones paramétricas para garantizar los momentos del proceso y las propiedades asintóticas de los estimadores restringen mucho esta clase de proceso, no permitiendo una estructura de dependencia fuerte. El estimador presentado en este trabajo, relaja todos los supuestos paramétricos de los procesos de volatilidad determinística y permite considerar una clase de proceso localmente estacionario de fuerte dependencia en la volatilidad.

Finalmente, este trabajo propone una nueva herramienta estadística para modelar retornos financieros, que es más flexible a lo disponible en la literatura.

Como trabajos futuros existen tres líneas de investigación que son atractivas para estudiar.

- Estimador de Kalman para procesos de volatilidad estocástica localmente estacionarios. Al utilizar la representación estocástica del proceso en la ecuación (2.7) (o la ecuación (3.9)) $Y_{t,T} = \mu + \nu_{t,T} + \epsilon_t$, parece natural utilizar una representación de espacio estado del proceso (considerando que el proceso es Gaussiano), y como mecanismo de estimación el filtro de Kalman, sin embargo, el ruido aditivo de $\{\epsilon_t\}$ no es Gaussiano. Existen trabajos como el artículo de Ruiz (1994) donde relajan el supuesto sobre el ruido aditivo y para la estimación asumen que es aproximadamente Gaussiano, es decir, se considera que $\{\epsilon_t\} \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2)$. Por ejemplo, si se supone que $\{u_t\} \stackrel{iid}{\sim} N(0, 1)$, entonces $\epsilon_t = \log(u_t^2) - \mathbb{E}(\log(u_t^2))$ tiene distribución $\log \chi_1^2$ centrada en cero y con $\mathbb{V}(\epsilon_t) = \pi^2/2$, al momento de estimar Ruiz (1994) asume que $\epsilon_t \stackrel{iid}{\sim} N(0, \pi^2/2)$ y utiliza el filtro de Kalman. Bajo ese mecanismo de estimación, muestra que el estimador es consistente y

presenta el teorema del límite central.

Considerando lo anterior, se puede realizar un trabajo de investigación en esa dirección, es decir, aproximar la distribución de $\{\epsilon_t\}$ por una densidad normal para utilizar una representación de espacio estado de un proceso localmente estacionario y operará con el filtro de Kalman. Además se puede estudiar las propiedades asintóticas del estimador y comparar esos resultados con los obtenidos en esta tesis.

- Estimación de Kalman para procesos de volatilidad localmente estacionarios utilizando mezclas de distribuciones normales. La idea general aquí consiste en utilizar la representación de espacio estado del proceso $\{\log r_{t,T}^2\}$ y aproximar la ley de la secuencia de ruido blanco $\{\epsilon_t\}$ mediante mezclas de distribuciones normales. Existen trabajos que utilizan aproximaciones de mezclas para el ruido aditivo $\{\epsilon_t\}$, como Kim et al. (1998) y, las aproximaciones presentadas ahí son bastante buenas y es más verosímil que aproximar la ley de $\{\epsilon_t\}$ directamente con la distribución normal como Ruiz (1994).

Para determinar la recursión de Kalman, se debe considerar la representación de espacio estado de $\{\nu_{t,T}\}$ más la representación de la mezcla para $\{\epsilon_t\}$, luego determinar los componentes del filtro y construir la función de verosimilitud. Al momento de maximizar se debe recurrir al algoritmo EM para estimar los parámetros de las mezclas y actualizar los valores del filtro.

En esta línea se pueden realizar dos trabajos interesantes, el primero desde la perspectiva computacional, ya que se debe recurrir a métodos de Monte Carlo

para estimar las medias en el algoritmo EM, por lo tanto un desarrollo computacionalmente eficientes puede ser un gran aporte. El segundo trabajo es estudiar las propiedades del estimador y compararlas con los resultados presentados en esta tesis.

- Estimación para procesos de volatilidad localmente estacionarios en presencia de datos faltantes. Un trabajo que puede ser de interés en aplicaciones prácticas es desarrollar un mecanismo de estimación ante la presencia de datos faltantes. Existen al menos dos posible estrategia para abordar este problema, la primera es utilizar una representación de espacio estado y determinar el filtro de Kalman para observaciones missing, aproximando la distribución de $\{\epsilon_t\}$. La segunda es utilizar el estimador local de Whittle y corregir el periodograma ante la presencia de datos faltantes. Para ambas estrategias se pueden realizar estudios de simulación sobre el impacto de los missing y estudiar las propiedades asintóticas.

Apéndice A

Lemas Generales

Esté apéndice contiene varios lemas auxiliares, que son utilizados para probar las proposiciones y teoremas de ambos procesos. Hay lemas generales que son utilizados para ambos procesos como Dahlhaus (1996, 1997) y algunos específicos dependiendo de los supuestos del proceso Dahlhaus (1997) y Palma and Olea (2010).

Sea $L_T : \mathbb{R} \rightarrow \mathbb{R}$, $T \in \mathbb{R}^+$ una función 2π periódica definida por

$$L_T(\alpha) := \begin{cases} T, & |\alpha| \leq 1/T, \\ 1/|\alpha|, & 1/T \leq |\alpha| \leq \pi. \end{cases}$$

Lema A.1. (Dahlhaus (1997)). Sea $k, \ell, S, M, T \in \mathbb{N}$, $\alpha, \beta, \gamma, \nu, v, \mu, x \in \mathbb{R}$ y $\Pi := (-\pi, \pi]$

Se tienen los siguientes resultados con K constante independiente de T .

(a) $L_T(\alpha)$ es monótona creciente en T y decreciente en $\alpha \in [0, \pi]$.

(b) $\int_{\Pi} L_T(\alpha)^k d\alpha \leq KT^{k-1}$ para todo $k \geq 1$.

(c) $\int_{\Pi} L_T(\alpha) d\alpha \leq K \log(T)$ para $T \geq 1$.

(d) $|\alpha|L_T(\alpha) \leq K$.

(e) $\int_{\Pi} L_T(\beta - \alpha)L_T(\alpha + \gamma) d\alpha \leq KL_T(\beta + \gamma) \log(T)$.

$$(f) \quad L_T(\nu)^k L_T(\mu)^\ell \leq L_T((\nu - \mu)/2)^k L_T(\mu)^\ell + L_T((\nu - \mu)/2)^\ell L_T(\nu)^k.$$

$$(g) \quad L_T(c\alpha) \leq K_c L_T(\alpha) \text{ para } |c\alpha| \leq \pi.$$

$$(h) \quad \int_{\Pi} L_N(\alpha)^\ell L_M(S(\alpha - \beta))^k d\alpha \leq K(N^\ell M^{k-1}/S) \log(M)\{k = 1\} \log(S)\{\ell = 1\}.$$

$$(i) \quad \int_{\Pi} L_N(\lambda - x) L_N(x - \mu) L_M(S(\alpha - x)) L_M(S(x - \beta)) dx \leq K(N/S) \log(M) \log(S) L_N(\lambda - \mu) L_M(\alpha - \beta).$$

$$(j) \quad \int_{\Pi} L_N(\lambda - x) L_N(x - \mu) L_M(S(\alpha - x)) dx \leq K(N/S) \log(M) \log(S) L_N(\lambda - \mu). \quad \square$$

Sea f una función compleja, se define $H_N(f(\cdot), \lambda)$ como

$$H_N(f(\cdot), \lambda) := \sum_{s=0}^{N-1} f(s) \exp(-i\lambda s),$$

y para el taper de los datos $h(x)$, se define

$$H_{k,N}(\lambda) := H_N\left(h^k\left(\frac{\cdot}{N}\right), \lambda\right), \quad H_N(\lambda) = H_{1,N}(\lambda).$$

Lema A.2. Sea $N, T \in \mathbb{N}$. Suponga que el taper h cumple con todos el supuesto 2.1 (v), y $\psi : [0, 1] \rightarrow \mathbb{R}$ es diferenciable con derivada acotada. Entonces, para todo $0 \leq t \leq N$ se tiene

$$\begin{aligned} H_N\left(\psi\left(\frac{\cdot}{T}\right), \lambda\right) &= \psi\left(\frac{t}{T}\right) H_N(\lambda) + \mathcal{O}\left(\sup_u |\psi'(u)| \frac{N}{T} L_N(\lambda)\right) \\ &= \mathcal{O}\left(\sup_{u \leq N/T} |\psi(u)| L_N(\lambda) + \sup_u |\psi'(u)| L_N(\lambda)\right). \end{aligned}$$

Lema A.3. Sea ψ continua con derivada acotada y $t_j = S(j - 1) + N/2$, $u_j = t_j/T$ con N, M, S y T que satisfacen el supuesto 2.1 (iv), entonces

$$\left| \sum_{j=1}^M \psi(u_j) \exp(i\lambda S_j) \right| \leq K \left(\sup_u |\psi(u)| + \sup_u |\psi'(u)| \right) L_M(S\lambda).$$

Lema A.4. *Suponga que el taper h cumple con el supuesto 2.1 (v). Entonces*

$$|H_N(\lambda)| \leq K L_N(\lambda),$$

y además

$$|H_N(\lambda)| \leq K N^{-1} L_N(\lambda)^2.$$

Lema A.5. *Considere la función $\phi : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{C}$, tal que $\partial\phi(u, \gamma)/\partial u$ existe y $|\partial\phi(u, \gamma)/\partial u| \leq K|\gamma|^{-2d(u)}$, donde $0 \leq d(u) \leq d$ para todo $u \in [0, 1]$. Entonces, para todo $0 \leq t \leq N$ se tiene que*

$$H_N \left[\phi \left(\frac{\cdot}{T}, \gamma \right), \lambda \right] = \phi \left(\frac{t}{T}, \gamma \right) H_N(\lambda) + \mathcal{O} \left[\frac{N}{T} |\gamma|^{-2d} L_N(\lambda) \right].$$

Lema A.6. *Suponga que $X_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(k)})'$ es un proceso multivariado localmente estacionario con densidad espectral $f_X(u, \lambda)$. Sea $\zeta = (\zeta_1, \dots, \zeta_k)' \in \mathbb{C}^k$ y considere el proceso univariado $Y_{t,T} = \zeta' X_{t,T}$, la densidad espectral de variación en el tiempo esta dada por:*

$$f_Y(u, \lambda) = \zeta' f_X(u, \lambda) \bar{\zeta}, \quad \forall (u, \lambda) \in [0, 1] \times [-\pi, \pi]$$

DEMOSTRACIÓN. Sin pérdida de generalidad, suponga que el vector de media del proceso $X_{t,T} = (X_{t,T}^{(1)}, \dots, X_{t,T}^{(k)})'$ es igual a cero, así, el proceso tiene la siguiente representación:

$$X_{t,T} = \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^0(\lambda) d\xi(\lambda)$$

Por lo tanto, el proceso univariado $Y_{t,T} = \zeta' X_{t,T}$ puede ser representado por

$$Y_{t,T} = \zeta' X_{t,T} = \int_{-\pi}^{\pi} \exp(i\lambda t) \zeta' A_{t,T}^0(\lambda) d\xi(\lambda),$$

de esta forma, denotamos $\tilde{A}_{t,T}^0(\lambda) = \zeta' A_{t,T}^0(\lambda)$ a la función de transferencia univariada. Con esta representación la densidad espectral del proceso univariado $\{Y_{t,T}\}$ está dada por

$$\begin{aligned} f_Y(u, \lambda) &= \tilde{A}_{t,T}^0(\lambda) \overline{\tilde{A}_{t,T}^0(\lambda)'} \\ &= \zeta' A_{t,T}^0(\lambda) \overline{A_{t,T}^0(\lambda)'} \bar{\zeta} \\ &= \zeta' f_X(u, \lambda) \bar{\zeta}. \end{aligned}$$

Apéndice B

Volatilidad Estocástica Corta Memoria

B.1. Proposiciones

En esta sección se demuestran tres proposiciones, que serán fundamentales para los dos grandes teoremas del capítulo 2.

Considere la función $\phi : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{R}$ y defina el siguiente operador funcional.

$$J(\phi) = \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) f(u, \lambda) d\lambda du, \quad (\text{B.1})$$

donde $f(u, \lambda)$ corresponde a la densidad espectral límite de variación en el tiempo de la secuencia definida en (2.7), ésta es,

$$f_Y(u, \lambda) = f_\nu(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi} \quad (\text{B.2})$$

donde $f_\nu(u, \lambda)$ es la densidad espectral de variación en tiempo del proceso $\{\nu_{t,T}\}$ y σ_ϵ^2 es la varianza del ruido blanco $\{\epsilon_t = \log(u_t^2) - \mathbb{E}(\log(u_t^2))\}$, ambos definidos en (2.7). Considere la versión muestral de $J(\cdot)$ como:

$$J_T(\phi) = \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) I_N(u_j, \lambda) d\lambda, \quad (\text{B.3})$$

donde M , $I_N(u, \lambda)$ y u_j , $j = 1, \dots, M$, están definidas en la sección 2.3 del capítulo 2. Sin pérdida de generalidad vamos a suponer que $\{Y_{t,T}\}$ esta centrado en cero.

Proposición B.1. *Suponga que se tienen todos los supuestos 2.1, además, la función $\phi(u, \lambda)$ es 2π periódica en λ , diferenciable con respecto a u y λ , con derivada uniformemente acotada $(\partial/\partial u)(\partial/\partial \lambda)\phi(u, \lambda)$. Entonces*

$$\mathbb{E}[J_T(\phi)] = J(\phi) + \mathcal{O}(M^{-1}) + \mathcal{O}(N^{-1}) + \mathcal{O}\left(\frac{N}{T} \log N\right).$$

DEMOSTRACIÓN. Directamente de la definición, se tiene lo siguiente

$$\begin{aligned} \mathbb{E}[J_T(\phi)] &= \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \mathbb{E}[I_N(u_j, \lambda)] d\lambda \\ &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \mathbb{E}|D_N(u_j, \lambda)|^2 d\lambda \\ &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s,t=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \mathbb{C}_Y(u_j, t, s) e^{i\lambda(s-t)} d\lambda, \end{aligned}$$

Donde

$$\begin{aligned} \mathbb{C}_Y(u, t, s) &= \mathbf{Cov}(Y_{[uT]-N/2+t+1,T}; Y_{[uT]-N/2+s+1,T}) \\ &= \mathbf{Cov}(\nu_{[uT]-N/2+t+1,T}; \nu_{[uT]-N/2+s+1,T}) + \mathbf{Cov}(\epsilon_{[uT]-N/2+t+1}; \epsilon_{[uT]-N/2+s+1}) \\ &= \mathbb{C}_\nu(u_j, t, s) + \mathbf{Cov}(\epsilon_{[uT]-N/2+t+1}; \epsilon_{[uT]-N/2+s+1}), \end{aligned}$$

como $\{\epsilon_t\}$ es una secuencia de ruido blanco, se tiene

$$\text{Cov}(\epsilon_{[uT]-N/2+t+1}; \epsilon_{[uT]-N/2+s+1}) = \begin{cases} \sigma_\epsilon^2, & \text{si } s = t, \\ 0, & \text{en otro caso.} \end{cases}$$

Utilizando la descomposición de la covarianza y reemplazando, se obtiene

$$\begin{aligned} \mathbb{E}[J_T(\phi)] &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s,t=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \mathbb{C}_\nu(u_j, t, s) e^{i\lambda(s-t)} d\lambda \\ &\quad + \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right)^2 \sigma_\epsilon^2 d\lambda. \end{aligned}$$

Para el primer término de la suma

$$\frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s,t=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \mathbb{C}_\nu(u_j, t, s) e^{i\lambda(s-t)} d\lambda$$

se utiliza el Lema A.8 de Dahlhaus (1997), donde se muestra en que para un proceso localmente estacionario $\{\nu_{t,T}\}$, que satisface los supuestos 2.1, la ecuación anterior es igual a

$$\int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) f_\nu(u, \lambda) d\lambda du + \mathcal{O}(M^{-1}) + \mathcal{O}(N^{-2}) + \mathcal{O}\left(\frac{N}{T} \log N\right).$$

Por otro lado, el segundo término puede ser escrito como

$$\begin{aligned} \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right)^2 \sigma_\epsilon^2 d\lambda &= \frac{\sigma_\epsilon^2}{2\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) d\lambda \\ &= \frac{1}{M} \sum_{j=1}^M A(u_j), \end{aligned}$$

donde $A(u) = \int_{-\pi}^{\pi} \phi(u, \lambda) \frac{\sigma_{\epsilon}^2}{2\pi} d\lambda$, entonces

$$\begin{aligned} \frac{1}{M} \sum_{j=1}^M A(u_j) &= \int_0^1 A(u) du + \mathcal{O}\left(\frac{1}{M}\right) \\ &= \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) \frac{\sigma_{\epsilon}^2}{2\pi} d\lambda du + \mathcal{O}\left(\frac{1}{M}\right). \end{aligned}$$

Finalmente, sumando ambos resultados, se tiene que:

$$\begin{aligned} \mathbb{E}[J_T(\phi)] &= \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) \left\{ f_{\nu}(u, \lambda) + \frac{\sigma_{\epsilon}^2}{2\pi} \right\} d\lambda du \\ &\quad + \mathcal{O}(M^{-1}) + \mathcal{O}(N^{-2}) + \mathcal{O}\left(\frac{N}{T} \log N\right). \end{aligned}$$

Proposición B.2. *Suponga que se cumplen los supuestos 2.1, y considere que $\phi_1(u, \lambda)$ y $\phi_2(u, \lambda)$ son 2π periódica en λ , diferenciable con respecto a u y λ , con derivada uniformemente acotada $(\partial/\partial u)(\partial/\partial \lambda)\phi(u, \lambda)$, entonces*

$$\begin{aligned} \lim_{T \rightarrow \infty} T \mathbf{Cov}[J_T(\phi_1), J_T(\phi_2)] &= \\ &4\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \lambda) \left(f_{\nu}(u, \lambda) + \frac{\sigma_{\epsilon}^2}{2\pi} \right)^2 d\lambda du \\ &+ 2\pi \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) f_{\nu}(u, \lambda) f_{\nu}(u, \mu) g_{\nu,4}(\lambda, -\lambda, \mu) d\lambda d\mu du \\ &+ \frac{\kappa_4 \sigma_{\epsilon}^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) d\lambda d\mu du. \end{aligned}$$

DEMOSTRACIÓN. Utilizando cálculo directo, se tiene que

$$T \mathbf{Cov}[J_T(\phi_1), J_T(\phi_2)] = \frac{T}{M^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \mathbf{Cov}[I_N(u_j, \lambda), I_N(u_k, \mu)] d\lambda d\mu.$$

Donde

$$\begin{aligned}
& \mathbf{Cov} [I_N(u_j, \lambda), I_N(u_k, \lambda)] \\
&= \frac{1}{[2\pi H_{2,N}(0)]^2} \mathbf{Cov} (|D_N(u_j, \lambda)|^2, |D_N(u_k, \mu)|^2) \\
&= \frac{1}{[2\pi H_{2,N}(0)]^2} \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\
&\quad \times e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbf{Cov} \left(Y_{[u_j T]-N/2+s+1, T} \bar{Y}_{[u_j T]-N/2+t+1, T}; Y_{[u_k T]-N/2+p+1, T} \bar{Y}_{[u_k T]-N/2+m+1, T} \right).
\end{aligned}$$

Se asume que el proceso $\{Y_{t,T}\}$ esta centrado en cero, entonces la covarianza se descompone de la siguiente manera

$$\begin{aligned}
& \mathbf{Cov} \left(Y_{[u_j T]-N/2+s+1, T} \bar{Y}_{[u_j T]-N/2+t+1, T}; Y_{[u_k T]-N/2+p+1, T} \bar{Y}_{[u_k T]-N/2+m+1, T} \right) \\
&= \mathbf{Cov} \left(\nu_{[u_j T]-N/2+s+1, T} \bar{\nu}_{[u_j T]-N/2+t+1, T}; \nu_{[u_k T]-N/2+p+1, T} \bar{\nu}_{[u_k T]-N/2+m+1, T} \right) \\
&\quad + \mathbb{E} \left(\nu_{[u_j T]-N/2+s+1, T} \bar{\nu}_{[u_k T]-N/2+p+1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T]-N/2+t+1} \bar{\epsilon}_{[u_k T]-N/2+m+1} \right) \\
&\quad + \mathbb{E} \left(\nu_{[u_j T]-N/2+s+1, T} \bar{\nu}_{[u_k T]-N/2+m+1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T]-N/2+t+1} \bar{\epsilon}_{[u_k T]-N/2+p+1} \right) \\
&\quad + \mathbb{E} \left(\nu_{[u_j T]-N/2+t+1, T} \bar{\nu}_{[u_k T]-N/2+p+1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T]-N/2+s+1} \bar{\epsilon}_{[u_k T]-N/2+m+1} \right) \\
&\quad + \mathbb{E} \left(\nu_{[u_j T]-N/2+t+1, T} \bar{\nu}_{[u_k T]-N/2+m+1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T]-N/2+s+1} \bar{\epsilon}_{[u_k T]-N/2+p+1} \right) \\
&\quad + \mathbf{Cov} \left(\epsilon_{[u_j T]-N/2+s+1} \bar{\epsilon}_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+p+1} \bar{\epsilon}_{[u_k T]-N/2+m+1} \right).
\end{aligned}$$

Se definen las siguientes cantidades:

$$\begin{aligned}
B_N^{(1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbf{Cov} \left(\nu_{[u_j T]-N/2+s+1, T} \bar{\nu}_{[u_j T]-N/2+t+1, T}; \nu_{[u_k T]-N/2+p+1, T} \bar{\nu}_{[u_k T]-N/2+m+1, T} \right) d\lambda d\mu,
\end{aligned}$$

$$\begin{aligned}
B_N^{(2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbb{E}\left(\nu_{[u_j T]-N/2+s+1, T} \nu_{[u_k T]-N/2+p+1, T}\right) \mathbb{E}\left(\epsilon_{[u_j T]-N/2+t+1} \epsilon_{[u_k T]-N/2+m+1}\right) d\lambda d\mu, \\
B_N^{(3)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbb{E}\left(\nu_{[u_j T]-N/2+s+1, T} \nu_{[u_k T]-N/2+m+1, T}\right) \mathbb{E}\left(\epsilon_{[u_j T]-N/2+t+1} \epsilon_{[u_k T]-N/2+p+1}\right) d\lambda d\mu, \\
B_N^{(4)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbb{E}\left(\nu_{[u_j T]-N/2+t+1, T} \nu_{[u_k T]-N/2+p+1, T}\right) \mathbb{E}\left(\epsilon_{[u_j T]-N/2+s+1} \epsilon_{[u_k T]-N/2+m+1}\right) d\lambda d\mu, \\
B_N^{(5)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbb{E}\left(\nu_{[u_j T]-N/2+t+1, T} \nu_{[u_k T]-N/2+m+1, T}\right) \mathbb{E}\left(\epsilon_{[u_j T]-N/2+s+1} \epsilon_{[u_k T]-N/2+p+1}\right) d\lambda d\mu, \\
B_N^{(6)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbf{Cov}\left(\epsilon_{[u_j T]-N/2+s+1} \epsilon_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+p+1} \epsilon_{[u_k T]-N/2+m+1}\right) d\lambda d\mu.
\end{aligned}$$

Note que $B_N^{(2)} = B_N^{(3)} = B_N^{(4)} = B_N^{(5)}$, entonces

$$T \mathbf{Cov} [J_T(\phi_1), J_T(\phi_2)] = B_N^{(1)} + 4B_N^{(2)} + B_N^{(6)}.$$

Para el desarrollo de la covarianza $B_N^{(1)}$ se puede utilizar el lema A.9 de Dahlhaus (1997). Por lo tanto, primero nos enfocaremos en desarrollar el término $B_N^{(2)}$.

Considere la siguiente función Q , que corresponde al producto de dos covarianzas

$$Q_{u_j, u_k}(s, p, t, m) = \mathbb{E} \left(\nu_{[u_j T] - N/2 + s + 1, T} \nu_{[u_k T] - N/2 + p + 1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T] - N/2 + t + 1} \epsilon_{[u_k T] - N/2 + m + 1} \right),$$

entonces, sumando sobre los índices s, p, t, m y la función taper $h(\cdot)$ se tiene que la cantidad $B_N^{(2)}$ es igual a

$$\begin{aligned} & \sum_{t, s, p, m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t) + i\mu(m-p)} Q_{u_j, u_k}(s, p, t, m) \\ &= \sum_{t, s, p, m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t) + i\mu(m-p)} \\ & \quad \times \int_{-\pi}^{\pi} A_{\nu; t_j - N/2 + s + 1, T}^0(x) \overline{A_{\nu; t_k - N/2 + p + 1, T}^0(x)} e^{ix(t_j - t_k + s - p)} dx \times \int_{-\pi}^{\pi} \frac{\sigma_\epsilon^2}{2\pi} e^{iy(t_j - t_k + t - m)} dy \\ &= \int_{-\pi}^{\pi} \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) A_{\nu; t_j - N/2 + s + 1, T}^0(x) e^{is(x+\lambda)} \sum_{p=0}^{N-1} h\left(\frac{p}{N}\right) \overline{A_{\nu; t_k - N/2 + p + 1, T}^0(x)} e^{-ip(x+\mu)} \\ & \quad \times e^{ix(t_j - t_k)} dx \int_{-\pi}^{\pi} \sum_{t=0}^{N-1} h\left(\frac{t}{N}\right) \frac{\sigma_\epsilon}{\sqrt{2\pi}} e^{-it(\lambda-y)} \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right) \frac{\sigma_\epsilon}{\sqrt{2\pi}} e^{-im(y-\mu)} e^{iy(t_j - t_k)} dy \\ &= \int_{-\pi}^{\pi} H_N \left(A_{\nu; t_j - N/2 + 1 + \cdot, T}^0(x) h\left(\frac{\cdot}{N}\right), -x - \lambda \right) H_N \left(\overline{A_{\nu; t_k - N/2 + 1 + \cdot, T}^0(x)} h\left(\frac{\cdot}{N}\right), x + \mu \right) \\ & \quad \times e^{ix(t_j - t_k)} dx \int_{-\pi}^{\pi} H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), \lambda - y \right) H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), y - \mu \right) e^{iy(t_j - t_k)} dy \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} H_N \left(A_{\nu; t_j - N/2 + 1 + \cdot, T}^0(x) h\left(\frac{\cdot}{N}\right), -x - \lambda \right) H_N \left(\overline{A_{\nu; t_k - N/2 + 1 + \cdot, T}^0(x)} h\left(\frac{\cdot}{N}\right), x + \mu \right) \\ & \quad \times H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), \lambda - y \right) H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), y - \mu \right) e^{i(x+y)(t_j - t_k)} dx dy, \end{aligned}$$

con este resultado la cantidad $B_N^{(2)}$ es

$$\begin{aligned}
B_N^{(2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times H_N \left(A_{\nu; t_j - N/2 + 1 + \cdot, T}^0(x) h \left(\frac{\cdot}{N} \right), -x - \lambda \right) H_N \left(\overline{A_{\nu; t_k - N/2 + 1 + \cdot, T}^0(x) h \left(\frac{\cdot}{N} \right)}, x + \mu \right) \\
&\quad \times H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h \left(\frac{\cdot}{N} \right), \lambda - y \right) H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h \left(\frac{\cdot}{N} \right), y - \mu \right) e^{i(x+y)(t_j - t_k)} dx dy d\mu d\lambda \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times H_N \left(A_{\nu; t_j - N/2 + 1 + \cdot, T}^0(x) h \left(\frac{\cdot}{N} \right), -x - \lambda \right) H_N \left(\overline{A_{\nu; t_k - N/2 + 1 + \cdot, T}^0(x) h \left(\frac{\cdot}{N} \right)}, x + \mu \right) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dx dy d\mu d\lambda.
\end{aligned}$$

Ahora considere el siguiente término

$$\begin{aligned}
&\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} H_N \left(A_{\nu; t_j - N/2 + 1 + \cdot, T}^0(x) h \left(\frac{\cdot}{N} \right), -x - \lambda \right) H_N \left(\overline{A_{\nu; t_k - N/2 + 1 + \cdot, T}^0(x) h \left(\frac{\cdot}{N} \right)}, x + \mu \right) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dx dy,
\end{aligned}$$

que por el Lema A.2, se puede aproximar a

$$\begin{aligned}
&\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} A_\nu(u_j, x) H_N(-x - \lambda) A_\nu(u_k, -x) H_N(x + \mu) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dx dy,
\end{aligned}$$

mientras que el resto $R_{j,k}$ de la expansión anterior puede ser acotado por

$$\begin{aligned}
\left| \sum_{j,k=1}^M \phi_1(u_j, \lambda) \phi_2(u_k, \mu) R_{j,k} \right| &\leq KM \frac{N}{T} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} L_N(-x - \lambda) L_N(x + \mu) \\
&\quad \times L_N(\lambda - y) L_N(y - \mu) L_M(S(x + y)) dx dy.
\end{aligned}$$

Esta cota es el resultado de aplicar el Lema A.3 de la siguiente manera,

$$\sum_{j=1}^M \phi_1(u_j, \lambda) A_\nu(u_j, x) \exp\{iS(x+y)j\} = \mathcal{O}(L_M(S(x+y))).$$

Por el Lema A.1(j) y A.1(e), se tiene que la desigualdad es acotada por

$$KM \frac{N}{T} \frac{N}{S} \log(M) \log(N) \log(S) L_N(\lambda - \mu)^2.$$

Finalmente, integrando con respecto a μ y λ y aplicando el factor $T^2 / [2\pi M H_{2,N}(0)]^2$ el resto de la expansión está acotado superiormente por

$$K \frac{N}{T} \log(M) \log(N) \log(S) \rightarrow 0.$$

Para resolver el termino principal se utilizará un procedimiento similar

$$\begin{aligned} & \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\ & \quad \times A_\nu(u_j, x) H_N(-x - \lambda) A_\nu(u_k, -x) H_N(x + \mu) \\ & \quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dx dy, \end{aligned}$$

primero, se reemplaza $\phi_1(u_j, \lambda)$ por $\phi_1(u_j, y)$ y $\phi_2(u_k, \mu)$ por $\phi_2(u_k, y)$. Luego se aplica el Lema A.3 para las diferencias

$$\begin{aligned} & \left| \sum_{j=1}^M (\phi_1(u_j, \lambda) - \phi_1(u_j, y)) A_\nu(u_j, x) \exp\{iS(x+y)j\} \right| \\ & \leq K |\lambda - y| L_M(S(x+y)), \\ & \left| \sum_{k=1}^M (\phi_2(u_k, \mu) - \phi_2(u_k, y)) A_\nu(u_k, -x) \exp\{iS(x+y)k\} \right| \\ & \leq K |y - \mu| L_M(S(x+y)), \end{aligned}$$

Después se aplican el lema A.1(d) a ambos restos y se tiene la misma cota superior (no se desarrolla el cruzado por que es menor a ambos), que es

$$\begin{aligned} & K \frac{T}{M^2 N^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} L_N(x + \mu) L_N(y - \mu) L_N(-x - \lambda) \\ & \quad \times L_M(S(x + y))^2 dx dy d\lambda d\mu \\ & \leq K \frac{T}{M^2 N^2} \log^2(N) \frac{NM}{S} \log(S) \leq K \frac{\log^2(N)}{N} \log(S) \rightarrow 0. \end{aligned}$$

Segundo, se sustituye la función de transferencia $A_\nu(u_j, x)$ por $A_\nu(u_j, -y)$ y $A_\nu(u_k, -x)$ por $A_\nu(u_k, y)$, de esta forma se tiene

$$\begin{aligned} & \left| \sum_{j=1}^M (A_\nu(u_j, x) - A_\nu(u_j, -y)) \phi_1(u_j, y) \exp\{iS(x + y)j\} \right| \\ & \leq K |x + y| L_M(S(x + y)) \\ & \leq K |x + \lambda| L_M(S(x + y)) + K |y - \lambda| L_M(S(x + y)) \\ & \left| \sum_{k=1}^M (A_\nu(u_k, -x) - A_\nu(u_k, y)) \phi_2(u_k, y) \exp\{iS(x + y)k\} \right| \\ & \leq K |x + y| L_M(S(x + y)) \\ & \leq K |x + \mu| L_M(S(x + y)) + K |y - \mu| L_M(S(x + y)), \end{aligned}$$

ahora nos encontramos con la siguiente cota superior para ambos terminos

$$\begin{aligned} & K \frac{T}{M^2 N^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} L_N(x + \mu) L_N(y - \mu) L_N(-x - \lambda) \\ & \quad \times L_M(S(x + y))^2 dx dy d\lambda d\mu \\ & \leq K \frac{T}{M^2 N^2} \log^2(N) \frac{NM}{S} \log(S) \leq K \frac{\log^2(N)}{N} \log(S) \rightarrow 0. \end{aligned}$$

Por los resultados anteriores, se tiene que el termino $B_N^{(2)}$ es igual a

$$\begin{aligned}
B_N^{(2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) H_N(-x - \lambda) A_\nu(u_k, y) H_N(x + \mu) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dy dx d\mu d\lambda + o(1) \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) H_N(x + \mu) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} \\
&\quad \times \sum_{s,t=0}^{N-1} h\left(\frac{s}{N}\right) h\left(\frac{t}{N}\right) e^{isx + ity} \int_{-\pi}^{\pi} e^{-i\lambda(t-s)} d\lambda dy dx d\mu + o(1) \\
&= \frac{2\pi \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) e^{i(x+y)(t_j - t_k)} \sum_{s=0}^{N-1} h^2\left(\frac{s}{N}\right) e^{is(x+y)} \\
&\quad \times \sum_{m,p=0}^{N-1} h\left(\frac{m}{N}\right) h\left(\frac{p}{N}\right) e^{-imy - ipx} \int_{-\pi}^{\pi} e^{-i\mu(p-m)} d\mu dy dx + o(1) \\
&= \frac{[2\pi]^2 \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) e^{i(x+y)(t_j - t_k)} \sum_{s=0}^{N-1} h^2\left(\frac{s}{N}\right) e^{is(x+y)} \\
&\quad \times \sum_{m=0}^{N-1} h^2\left(\frac{m}{N}\right) e^{-im(y+x)} dy dx + o(1) \\
&= \frac{[2\pi]^2 \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \sum_{s,m=0}^{N-1} h^2\left(\frac{m}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) e^{i(x+y)(t_j - t_k + s - m)} dy dx + o(1),
\end{aligned}$$

utilizando el teorema cambio de variable $z = x + y$ y $w = y$, se tiene que

$$\begin{aligned}
B_N^{(2)} &= \frac{[2\pi]^2 \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \sum_{s,m=0}^{N-1} h^2\left(\frac{m}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, w) \phi_2(u_k, w) \\
&\quad \times A_\nu(u_j, -w) A_\nu(u_k, w) e^{iz(t_j - t_k + s - m)} dz dw + o(1) \\
&= \frac{2\pi \times T}{[M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{s,m=0}^{N-1} \sum_{\substack{j,k=1 \\ S(j-k)=s-m}}^M h^2\left(\frac{m}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} \phi_1(u_j, w) \phi_2(u_k, w) \\
&\quad \times A_\nu(u_j, -w) A_\nu(u_k, w) dw + o(1)
\end{aligned}$$

Para $S < N$ el termino $B_N^{(2)}$ se puede escribir como

$$\begin{aligned}
B_N^{(2)} &= \frac{2\pi \times T}{[M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \\
&\quad \times \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_{j+p}, x) A_\nu(u_{j+p}, x) A_\nu(u_j, -x) dx + o(1)
\end{aligned}$$

Note que por los supuestos de esta proposición, el producto $\phi_1(u, x) f_\nu(u, x)$ y $\phi_2(u, x) f_\nu(u, x)$ son diferenciables con respecto a u , más aún, el límite $\lim_{S,T \rightarrow \infty} \frac{S|p|}{T} = 0$ para todo $|p| \leq \frac{N}{S}$, en consecuencia

$$\begin{aligned}
\frac{S}{T} \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_{j+p}, x) A_\nu(u_j, -x) A_\nu(u_{j+p}, x) \frac{\sigma_\epsilon^2}{2\pi} dx \\
\longrightarrow \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) f_\nu(u, x) \frac{\sigma_\epsilon^2}{2\pi} dx du,
\end{aligned}$$

para todo $|p| < \frac{N}{S}$ cuando $M, N, S, T \rightarrow \infty$. Por otro lado se tiene

$$\begin{aligned}
\frac{2\pi \times T^2 N^2}{S^2 [M H_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \frac{S}{N^2} \\
\longrightarrow 2\pi \int_0^1 \int_{-x}^{1-x} h^2(x) h^2(x+y) dx dy \left(\int_0^1 h^2(x) dx \right)^{-2} = 2\pi,
\end{aligned}$$

cuando $M, N, S, T \rightarrow \infty$. Con ambos resultados tenemos

$$B_N^{(2)} \longrightarrow 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) f_\nu(u, x) \frac{\sigma_\epsilon^2}{2\pi} dx du, \quad \text{cuando } M, N, S, T \rightarrow \infty.$$

Para $B_N^{(1)}$ se usa directamente el Lema A.9 de Dahlhaus (1997), donde

$$\begin{aligned} B_N^{(1)} \longrightarrow & 4\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) f_\nu^2(u, x) dx du \\ & + 2\pi \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) f_\nu(u, \lambda) f_\nu(u, \mu) g_{\nu,4}(\lambda, -\lambda, \mu) d\lambda d\mu du, \end{aligned}$$

cuando $M, N, S, T \rightarrow \infty$. Note que si el proceso $\{\nu_{t,T}\}$ es generado por innovaciones Gaussianas, entonces $g_{\nu,4}(\lambda, -\lambda, \mu) = 0$. Para el caso $B_N^{(6)}$ utilizamos nuevamente el resultado de Dahlhaus (1997), de modo que

$$\begin{aligned} B_N^{(6)} \longrightarrow & 4\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 dx du \\ & + \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) d\lambda d\mu du, \end{aligned}$$

Proposición B.3. *Sea $\text{cum}_\ell(\cdot)$ el ℓ -ésimo cumulante con $\ell \geq 3$. Suponga que se tienen los supuestos 2.1, entonces, $T^{\ell/2} \text{cum}_\ell(J_T(\phi)) \rightarrow 0$, cuando $T \rightarrow \infty$.*

DEMOSTRACIÓN. Recuerde que $J_T(\phi)$ es

$$J_T(\phi) = \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) I_N(u_j, \lambda) d\lambda, \quad (\text{B.4})$$

y defina $\Pi = (-\pi, \pi]$, entonces

$$\begin{aligned}
& T^{\ell/2} \mathbf{cum}_\ell(J_T(\phi)) \\
&= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \int_{\Pi^\ell} \left[\prod_{r=1}^\ell \phi(u_{j_r}, \lambda_r) \right] \mathbf{cum} \left(D_N(u_{j_1}, \lambda_1) D_N(u_{j_1}, -\lambda_1), \right. \\
&\quad \left. \dots, D_N(u_{j_\ell}, \lambda_\ell) D_N(u_{j_\ell}, -\lambda_\ell) \right) d\lambda_1 \cdots d\lambda_\ell \\
&= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \int_{\Pi^\ell} \left[\prod_{r=1}^\ell \phi(u_{j_r}, \lambda_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \exp\{-i\lambda_1(k_1 - n_1)\} \times \cdots \times \exp\{-i\lambda_\ell(k_\ell - n_\ell)\} \\
&\quad \times \mathbf{cum} \left(Y_{[u_{j_1} T] - N/2 + k_1 + 1, T} Y_{[u_{j_1} T] - N/2 + n_1 + 1, T}, \dots, \right. \\
&\quad \quad \left. Y_{[u_{j_\ell} T] - N/2 + k_\ell + 1, T} Y_{[u_{j_\ell} T] - N/2 + n_\ell + 1, T} \right) d\lambda_1 \cdots d\lambda_\ell \\
&= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^\ell \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \mathbf{cum} \left(Y_{[u_{j_1} T] - N/2 + k_1 + 1, T} Y_{[u_{j_1} T] - N/2 + n_1 + 1, T}, \dots, \right. \\
&\quad \quad \left. Y_{[u_{j_\ell} T] - N/2 + k_\ell + 1, T} Y_{[u_{j_\ell} T] - N/2 + n_\ell + 1, T} \right),
\end{aligned}$$

donde $\widehat{\phi}(u, k)$ corresponde al k éximo coeficiente de Fourier de $\phi(u, \lambda)$, es decir,

$$\widehat{\phi}(u, k) = \int_{-\pi}^{\pi} \phi(u, \lambda) \exp(-i\lambda k) d\lambda.$$

Para realizar este cálculo se utilizará el teorema de los cumulantes para productos de variables aleatorias de Brillinger (1981) (teorema 2.3.2). Considere el esquema

$$\begin{array}{cc}
(1, 1) & (1, 2) \\
\vdots & \vdots \\
(\ell, 1) & (\ell, 2)
\end{array}$$

aquí (m, n) denota la posición donde se ubica la variable aleatoria, es decir, será la variable $Y_{[u_{j_m}]T-N/2+k_m+1, T}$ cuando $m = 1, \dots, \ell$ y $n = 1$, y $Y_{[u_{j_m}]T-N/2+n_m+1, T}$ cuando $m = 1, \dots, \ell$ y $n = 2$. Por el teorema de los cumulantes se tiene que la suma será realizada sobre todas las particiones indescompensable $\{P_1, \dots, P_m\}$ del conjunto $\mathcal{S} = \{(1, 1), \dots, (\ell, 2)\}$ con $|P_t| \geq 2$ (La secuencia se considera centrada en cero).

$$= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \sum_{\{P_1, \dots, P_m\}} \prod_{ip} \prod_{t=1}^m \text{cum}(Y_{[u_{j_{\omega_1} T]-N/2+\omega_2+1, T} | (\omega_1, \omega_2) \in P_t),$$

como los procesos $\{\nu_{t,T}\}$ y $\{\epsilon_t\}$ son ortogonales, se puede utilizar la descomposición de cumulantes

$$= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \sum_{\{P_1, \dots, P_m\}} \prod_{ip} \left[\text{cum}(\nu_{[u_{j_{\omega_1} T]-N/2+\omega_2+1, T} | (\omega_1, \omega_2) \in P_t) \right. \\ \left. + \text{cum}(\epsilon_{[u_{j_{\omega_1} T]-N/2+\omega_2+1} | (\omega_1, \omega_2) \in P_t) \right] \\ = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \sum_{\{P_1, \dots, P_m\}} \left(\prod_{t=1}^m \text{cum}(\nu_{[u_{j_{\omega_1} T]-N/2+\omega_2+1, T} | (\omega_1, \omega_2) \in P_t) \right. \\ \left. + \sum_{I \in \mathcal{I}} \prod_{t_1 \in I} \prod_{t_2 \in I^c} \text{cum}(\nu_{[u_{j_{\omega_1} T]-N/2+\omega_2+1, T} | (\omega_1, \omega_2) \in P_{t_1}) \right. \\ \left. \times \text{cum}(\epsilon_{[u_{j_{\omega_1} T]-N/2+\omega_2+1} | (\omega_1, \omega_2) \in P_{t_2}) \right. \\ \left. + \prod_{t=1}^m \text{cum}(\epsilon_{[u_{j_{\omega_1} T]-N/2+\omega_2+1} | (\omega_1, \omega_2) \in P_t) \right),$$

donde el conjunto \mathcal{I} se define como $\mathcal{I} = \{I : I \subset \{1, \dots, m\} \wedge \{1, \dots, m\} \not\subset I\}$, si $I \in \mathcal{I}$, entonces se define I^c como $I^c = \{1, \dots, m\} - I$. Note que la cardinalidad del

conjunto \mathcal{I} es $\#\mathcal{I} = 2^m - 2$ y para todo $I \in \mathcal{I}$ se tiene que $1 \leq |I| < m$.

Defina los siguientes cumulantes

$$\begin{aligned}
A_T^{(1)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}_{ip}} \prod_{t=1}^m \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} | (\omega_1, \omega_2) \in P_t), \\
A_T^{(2)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}_{ip}} \sum_{I \in \mathcal{I}} \prod_{t_1 \in I} \prod_{t_2 \in I^c} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} | (\omega_1, \omega_2) \in P_{t_1}) \\
&\quad \quad \times \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_{t_2}), \\
A_T^{(3)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}_{ip}} \prod_{t=1}^m \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_t).
\end{aligned}$$

Por Lema A.10 de Dahlhaus (1997) se tiene que $A_T^{(1)}, A_T^{(3)} \rightarrow 0$, cuando $T \rightarrow \infty$. Recuerde que $\{\epsilon_t\}$ es una secuencia independiente idénticamente distribuida con todos los momentos finitos, entonces, podemos elegir una constante C_k , con $0 < C_k < \infty$ para acotar el k ésimo cumulante, es decir, $|\text{cum}(\epsilon_{t_1}, \dots, \epsilon_{t_k})| < C_k$. Considerando lo anterior, se puede operar con $A_T^{(2)}$ de la siguiente manera

$$\begin{aligned}
A_T^{(2)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}} \sum_{ip} \prod_{I \in \mathcal{I}} \prod_{t_1 \in I} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} | (\omega_1, \omega_2) \in P_{t_1}) \\
&\quad \quad \times \prod_{t_2 \in I^c} \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_{t_2}) \\
A_T^{(2)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{\{P_1, \dots, P_m\}} \sum_{ip} \prod_{I \in \mathcal{I}} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \\
&\quad \times \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \quad \prod_{t_1 \in I} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} | (\omega_1, \omega_2) \in P_{t_1}) \\
&\quad \quad \times \prod_{t_2 \in I^c} \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_{t_2}) \\
&= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{\{P_1, \dots, P_m\}} \sum_{ip} \prod_{I \in \mathcal{I}} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \\
&\quad \times \sum_{n_1, \dots, n_\ell=0}^{N-1} \int_{\Pi^\ell} \left[\prod_{r=1}^{\ell} \phi(u_{j_r}, \lambda_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \quad \times \exp\{-i\lambda_1(k_1 - n_1)\} \times \dots \times \exp\{-i\lambda_\ell(k_\ell - n_\ell)\} \\
&\quad \quad \prod_{t_1 \in I} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} | (\omega_1, \omega_2) \in P_{t_1}) \\
&\quad \quad \times \prod_{t_2 \in I^c} \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_{t_2}) d\lambda_1 \dots d\lambda_\ell.
\end{aligned}$$

Considere que los elementos de partici3n P_t son fijos, as3 utilizaremos el mismo razonamiento que Dahlhaus (1997). Sea $P_t := \{c_1, \dots, c_k\}$ el conjunto de elementos en P_t y su versi3n de dimensi3n reducida ser3 $\bar{P}_t = \{c_1, \dots, c_{k-1}\}$, tambi3n tenemos $\mu_{\bar{P}_t} := (\mu_{c_1}, \dots, \mu_{c_{k-1}})$ y $\mu_{c_k} = -\sum_{j=1}^{k-1} \mu_{c_j}$. Sea m el n3mero de la particiones y $\mu = (\mu_{\bar{P}_1}, \dots, \mu_{\bar{P}_m})$. Denotemos a $a_r = (r, 1)$ y $b_r = (r, 2)$ en la partici3n indescompensable, entonces utilizando la notaci3n se tiene

$$\begin{aligned}
&= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{\{P_1, \dots, P_m\}ip} \sum_{I \in \mathcal{I}} \sum_{j_1, \dots, j_\ell=1}^M \int_{\Pi^\ell} \left[\prod_{r=1}^{\ell} \phi(u_{j_r}, \lambda_r) \right] \\
&\times \int_{\Pi^{2\ell-m}} \prod_{z \in I} \left\{ \prod_{\substack{r=1 \\ a_r \in P_z}}^{2\ell} \left[H_N \left(A_{t_{j_r} - N/2 + 1 + \cdot, T}(\mu_{a_r}) h \left(\frac{\cdot}{N} \right), \lambda_r - \mu_{a_r} \right) \right. \right. \\
&\times \left. \left. \prod_{\substack{r=1 \\ b_r \in P_z}}^{2\ell} H_N \left(A_{t_{j_r} - N/2 + 1 + \cdot, T}(\mu_{b_r}) h \left(\frac{\cdot}{N} \right), -\lambda_r - \mu_{b_r} \right) \right] \right\} \\
&\times \prod_{z \in I^c} \left\{ \prod_{\substack{r=1 \\ a_r \in P_z}}^{2\ell} \left[H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h \left(\frac{\cdot}{N} \right), \lambda_r - \mu_{a_r} \right) \right] \prod_{\substack{r=1 \\ b_r \in P_z}}^{2\ell} \left[H_N \left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h \left(\frac{\cdot}{N} \right), -\lambda_r - \mu_{b_r} \right) \right] \right\} \\
&\times \left[\prod_{\substack{r=1 \\ P_r \in I}}^m g_{|P_r|}(\mu_{\overline{P_r}}) \right] \left[\prod_{\substack{r=1 \\ P_r \in I^c}}^m g_{|P_r|}(\epsilon) \right] \\
&\times \exp \left(i \sum_{r=1}^{\ell} t_{j_r}(\mu_{a_r} + \mu_{b_r}) \right) d\mu d\lambda_1 \cdots d\lambda_\ell
\end{aligned}$$

Primero se utiliza el Lema A.2 de manera recursiva, es decir, se reemplaza sucesivamente $H_N(A_{t_{j_r} - N/2 + 1 + \cdot, T}(\mu)h(\cdot/N), \lambda - \mu)$ por $A(u_{j_r}, \mu)H_N(\lambda - \mu)$ y recordar que $|H_N(h(\cdot/N), \lambda)| \leq KL_N(\lambda)$. Así se puede acotar los restos por

$$\begin{aligned}
&K \frac{T^{\ell/2}}{M^\ell N^\ell} \sum_{\{P_1, \dots, P_m\}ip} \sum_{I \in \mathcal{I}} \int_{\Pi^\ell} \int_{\Pi^{2\ell-m}} M \frac{N}{T} \prod_{z \in I} \left\{ \left[\prod_{\substack{r=1 \\ a_r \in P_z}}^{2\ell} L_N(\lambda_r - \mu_{a_r}) \right] \left[\prod_{\substack{r=1 \\ b_r \in P_z}}^{2\ell} L_N(-\lambda_r - \mu_{b_r}) \right] \right\} \\
&\times \prod_{z \in I^c} \left\{ \left[\prod_{\substack{r=1 \\ a_r \in P_z}}^{2\ell} L_N(\lambda_r - \mu_{a_r}) \right] \left[\prod_{\substack{r=1 \\ b_r \in P_z}}^{2\ell} L_N(-\lambda_r - \mu_{b_r}) \right] \right\} \\
&\times \left[\prod_{r=2}^{\ell} L_M(S(\mu_{a_r} + \mu_{b_r})) \right] d\mu d\lambda_1 \cdots d\lambda_\ell
\end{aligned}$$

Note también que

$$\prod_{z \in I} \left[\prod_{\substack{r=1 \\ a_r \in P_z}}^{2\ell} L_N(\lambda_r - \mu_{a_r}) \prod_{\substack{r=1 \\ a_r \in P_z}}^{2\ell} L_N(\lambda_r - \mu_{a_r}) \right] = \prod_{r=1}^{\ell} L_N(\lambda_r - \mu_{a_r})$$

$$\prod_{z \in I^c} \left[\prod_{\substack{r=1 \\ b_r \in P_z}}^{2\ell} L_N(-\lambda_r - \mu_{b_r}) \prod_{\substack{r=1 \\ b_r \in P_z}}^{2\ell} L_N(-\lambda_r - \mu_{b_r}) \right] = \prod_{r=1}^{\ell} L_N(-\lambda_r - \mu_{b_r})$$

Entonces el resto se encuentra acotado por

$$K \frac{T^{\ell/2}}{M^{\ell} N^{\ell}} \sum_{\{P_1, \dots, P_m\} \text{ ip}} \sum_{I \in \mathcal{I}} \int_{\Pi^{\ell}} \int_{\Pi^{2\ell-m}} M \frac{N}{T} \left[\prod_{r=1}^{\ell} L_N(\lambda_r - \mu_{a_r}) L_N(-\lambda_r - \mu_{b_r}) \right]$$

$$\times \left[\prod_{r=2}^{\ell} L_M(S(\mu_{a_r} + \mu_{b_r})) \right] d\mu d\lambda_1 \cdots d\lambda_{\ell}$$

Aquí se utiliza el mismo resultado de Dahlhaus (1997) donde se integra sobre todos los λ_r y luego sobre todos los μ , de esta forma el error es acotado por

$$\leq K \frac{T^{\ell/2}}{T^{\ell-1}} \frac{N}{T} (\log(N) \log(M) \log(S))^{\ell} \rightarrow 0.$$

Y el termino principal es acotado por

$$K \frac{T^{\ell/2}}{M^{\ell} N^{\ell}} \sum_{\{P_1, \dots, P_m\} \text{ ip}} \sum_{I \in \mathcal{I}} \int_{\Pi^{\ell}} \int_{\Pi^{2\ell-m}} \left[\prod_{r=1}^{\ell} L_N(\lambda_r - \mu_{a_r}) L_N(-\lambda_r - \mu_{b_r}) \right]$$

$$\times \prod_{r=1}^{\ell} L_M(S(\mu_{a_r} + \mu_{b_r})) d\mu d\lambda_1 \cdots d\lambda_{\ell}$$

$$\leq K \frac{T^{\ell/2}}{T^{\ell-1}} (\log(N) \log(M) \log(S))^{\ell} \rightarrow 0.$$

Finalmente, como $A_T^{(1)}, A_T^{(2)}, A_T^{(3)} \rightarrow 0$, se tiene que

$$T^{\ell/2} \text{cum}_{\ell}(J_T(\phi)) \rightarrow 0, \quad \text{cuando } T \rightarrow \infty.$$

B.2. Demostración de Teoremas

DEMOSTRACIÓN DEL TEOREMA 3.1 Vamos a utilizar un argumento estándar para mostrar la consistencia del estimador de Whittle, es decir, es suficiente mostrar que

$$\sup_{\theta \in \Theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \rightarrow 0,$$

en probabilidad, cuando $T \rightarrow \infty$, donde $\mathcal{L}(\theta) := \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left[\log f_{\theta}(u, \lambda) + \frac{f_{\theta_0}(u, \lambda)}{f_{\theta}(u, \lambda)} \right] d\lambda du$. Recuerde que para nuestro caso $f_{\theta}(u, \lambda) = f_{\nu}(u, \lambda) + \frac{\sigma_{\epsilon}^2}{2\pi}$. Defina $g_{\theta}(u, \lambda) = f_{\theta}(u, \lambda)^{-1}$, que por el supuesto 2.1 se tiene que $g_{\theta}(u, \lambda)$ es continua en θ , λ y u , de esta forma g_{θ} puede ser aproximada por la suma de Cesaro

$$g_{\theta}^{(L)}(u, \lambda) = \frac{1}{4\pi^2} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \widehat{g}_{\theta}(\ell, m) \exp(-i2\pi u_j \ell - i\lambda m),$$

donde $\widehat{g}_{\theta}(\ell, m) = \int_0^1 \int_{-\pi}^{\pi} g_{\theta}(u, \lambda) \exp(i2\pi u \ell + i\lambda m) du d\lambda$, note que $\sup_{\theta} |g_{\theta}(u, \lambda) - g_{\theta}^{(L)}(u, \lambda)| < \varepsilon$. Del Teorema 3.2 de Dahlhaus (1997), podemos escribir

$$\begin{aligned} & \sup_{\theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \\ & \leq \mathcal{O}\left(\frac{1}{M}\right) + \frac{\varepsilon}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) + f(u_j, \lambda)] d\lambda \\ & \quad + \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\ & \quad \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \exp(-i2\pi u_j \ell - i\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \end{aligned}$$

aquí se puede notar que $|\widehat{g}_{\theta}(\ell, m)| \leq 2\pi \sup_{(\theta, u, \lambda)} |g_{\theta}(u, \lambda)|$, sin embargo, por el supuesto 2.1 $|g_{\theta}(u, \lambda)|$ es continuo en θ , u y λ , de esta forma, y dado que el espacio paramétrico es compacto, tenemos que $|\widehat{g}_{\theta}(\ell, m)| \leq K$, para alguna constante K

positiva. Ahora, por definición, para valores fijos de $\ell, m = 1, \dots, L$, se puede elegir $\phi(u, \lambda) = \cos(2\pi u\ell) \cos(\lambda m)$ o $\phi(u, \lambda) = \sin(2\pi u\ell) \cos(\lambda m)$ en la proposición B.1 y $\phi_1(u, \lambda) = \cos(2\pi u\ell) \cos(\lambda m)$ o $\phi_1(u, \lambda) = \phi_2(u, \lambda) = \sin(2\pi u\ell) \cos(\lambda m)$ en la proposición B.2, entonces se deduce que

$$\begin{aligned} & \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\ & \quad \times \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \exp(-i2\pi u_j \ell - i\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \\ & \leq \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\ & \quad \times \left\{ \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \cos(-i2\pi u_j \ell) \cos(\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \right. \\ & \quad \left. + \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \sin(-i2\pi u_j \ell) \cos(\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \right\} \xrightarrow{\mathbb{P}} 0 \end{aligned}$$

y demás

$$\frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \{I_N(u_j, \lambda) + f(u_j, \lambda)\} d\lambda \rightarrow 2 \int_0^1 \int_{-\pi}^{\pi} f(u, \lambda) d\lambda du,$$

en probabilidad, cuando $M \rightarrow \infty$, de esta forma

$$\sup_{\theta \in \Theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \xrightarrow{\mathbb{P}} 0, \quad \text{cuando } T \rightarrow \infty$$

DEMOSTRACIÓN DEL TEOREMA 3.2 Sea θ_0 el verdadero valor del parámetro θ . Si se tienen los supuestos 2.1, entonces el estimador de Whittle satisface el teorema del límite central

$$\sqrt{T} (\hat{\theta}_T - \theta_0) \rightarrow N [0, \Gamma(\theta_0)^{-1} [\Gamma(\theta_0) + \Lambda(\theta_0) + \Omega(\theta_0)] \Gamma(\theta_0)^{-1}]$$

en distribución, cuando $T \rightarrow \infty$. Donde las matrices están definidos por

$$\begin{aligned} \Gamma(\theta_0) &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} [\nabla \log f_{\theta_0}(u, \lambda)] [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du \\ \Lambda(\theta_0) &= \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu du \\ \Omega(\theta_0) &= \frac{1}{8\pi} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f_{\nu, \theta_0}(u, \lambda) f_{\nu, \theta_0}(u, \mu) \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' g_{\nu, 4}(\lambda, -\lambda, \mu) d\lambda d\mu du. \end{aligned}$$

Demostración. Sea $\hat{\theta}_T$ el valor del parámetro que minimiza la función de log verosimilitud de Whittle $\mathcal{L}_T(\theta)$ dada en (2.9) y sea θ_0 el verdadero valor del parámetro. Por el teorema del valor medio, existe un vector $\bar{\theta}_T$ que satisface $\|\bar{\theta}_T - \theta_0\| \leq \|\hat{\theta}_T - \theta_0\|$, tal que

$$\nabla \mathcal{L}_T(\hat{\theta}_T) - \nabla \mathcal{L}_T(\theta_0) = [\nabla^2 \mathcal{L}_T(\bar{\theta}_T)] (\hat{\theta}_T - \theta_0).$$

Es suficiente mostrar que

- (a) $\nabla^2 \mathcal{L}_T(\theta_0) \rightarrow \Gamma(\theta_0)$, en probabilidad cuando $T \rightarrow \infty$.
- (b) $\nabla^2 \mathcal{L}_T(\bar{\theta}_T) - \nabla^2 \mathcal{L}_T(\theta_0) \rightarrow 0$, en probabilidad, cuando $T \rightarrow \infty$.
- (c) $\sqrt{T} \nabla \mathcal{L}_T(\theta_0) \rightarrow N [0, \Gamma(\theta_0) + \Lambda(\theta_0) + \Omega(\theta_0)]$, cuando $T \rightarrow \infty$.

Observe que

$$\begin{aligned}
\nabla^2 \mathcal{L}_T(\theta) &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) - f_{\theta}(u_j, \lambda)] \nabla^2 f_{\theta}(u_j, \lambda)^{-1} - \nabla f_{\theta}(u_j, \lambda) [f_{\theta}(u_j, \lambda)^{-1}]' d\lambda \\
&= \frac{1}{4\pi} \frac{1}{M} \left\{ \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) [I_N(u_j, \lambda) - f_{\theta}(u_j, \lambda)] d\lambda \right. \\
&\quad \left. + \sum_{j=1}^M \int_{-\pi}^{\pi} \nabla \log f_{\theta}(u_j, \lambda) \nabla \log f_{\theta}(u_j, \lambda)' d\lambda \right\} \\
&= \frac{1}{4\pi} [J_T(\phi) - J(\phi)] + \Gamma(\theta) + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

donde $\phi(u, \lambda) = \nabla^2 f_{\theta}(u, \lambda)^{-1}$. Aplicando la proposición 1 y la proposición 2 se obtiene la parte (a) y (b). Para la parte (c) se puede mostrar por medio de los cumulantes. Mostraremos que los cumulantes mayores o iguales a tres de $\sqrt{T} \nabla \mathcal{L}_T(\theta_0)$ converge a cero. Note que

$$\begin{aligned}
\nabla \mathcal{L}_T(\theta_0) &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) - f_{\theta_0}(u_j, \lambda)] \nabla f_{\theta_0}(u_j, \lambda)^{-1} d\lambda \\
&= \frac{1}{4\pi} J_T(\phi) - \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} f_{\theta_0}(u_j, \lambda) \nabla f_{\theta_0}(u_j, \lambda)^{-1} d\lambda \\
&= \frac{1}{4\pi} [J_T(\phi) - J(\phi)] + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

donde $\phi(u, \lambda) = \nabla f_{\theta_0}(u, \lambda)^{-1}$. Por la proposición B.1 y el supuesto 2.1 (iv), el primer cumulante de $\sqrt{T} \nabla \mathcal{L}_T(\theta_0)$ satisface

$$\begin{aligned}
\sqrt{T} \mathbb{E}(\nabla \mathcal{L}_T(\theta_0)) &= \mathcal{O}\left(\frac{\sqrt{T}}{M}\right) \\
&\rightarrow 0,
\end{aligned}$$

cuando $T \rightarrow \infty$. Por el resultado anterior se tiene que el cumulante de segundo orden de $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)$ puede ser escrito como

$$T \text{Cov}[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)] = \frac{1}{16\pi^2} T \text{Cov}[J_T(\phi), J_T(\phi)].$$

Y por la proposición B.2 se tiene que

$$\begin{aligned} & \lim_{T \rightarrow \infty} T \text{Cov}[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)] \\ &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \lambda)^{-1}]' f_{\theta_0}(u, \lambda)^2 d\lambda du \\ &+ \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu \\ &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla \log f_{\theta_0}(u, \lambda) [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du \\ &+ \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu \\ &+ \frac{1}{8\pi} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f_{\nu, \theta_0}(u, \lambda) f_{\nu, \theta_0}(u, \mu) \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' g_{\nu, 4}(\lambda, -\lambda, \mu) d\lambda d\mu \\ &= \Gamma(\theta_0) + \Lambda(\theta_0) + \Omega(\theta_0) \end{aligned}$$

donde las matrices son

$$\begin{aligned} \Gamma(\theta_0) &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla \log f_{\theta_0}(u, \lambda) [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du \\ \Lambda(\theta_0) &= \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu \\ \Omega(\theta_0) &= \frac{1}{8\pi} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f_{\nu, \theta_0}(u, \lambda) f_{\nu, \theta_0}(u, \mu) \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' g_{\nu, 4}(\lambda, -\lambda, \mu) d\lambda d\mu. \end{aligned}$$

Para finalizar, para $p > 2$, por la Proposición B.3 tenemos que $T^{p/2} \text{Cum}_p[\nabla\mathcal{L}_T(\theta_0)] \rightarrow 0$, cuando $T \rightarrow \infty$, así mostramos la parte (c).

Apéndice C

Volatilidad Estocástica Larga Memoria

C.1. Proposiciones

En esta sección se desarrollan las demostraciones para los dos teoremas del capítulo 3. Primero se deben revisar algunas definiciones.

Considere la función $\phi : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{R}$ y defina el siguiente operador funcional.

$$J(\phi) = \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) f(u, \lambda) d\lambda du, \quad (\text{C.1})$$

donde $f(u, \lambda)$ corresponde a la densidad espectral límite de variación en el tiempo de la secuencia (3.9), ésta es,

$$f_Y(u, \lambda) = f_\nu(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi} \quad (\text{C.2})$$

donde $f_\nu(u, \lambda)$ es la densidad espectral de variación en tiempo del proceso Gaussiano localmente estacionario de larga memoria $\{\nu_{t,T}\}$. Se define la versión muestral de $J(\cdot)$ como:

$$J_T(\phi) = \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) I_N(u_j, \lambda) d\lambda, \quad (\text{C.3})$$

donde M y u_j , $j = 1, \dots, M$, están definidas en la sección de estimación 3.2.2. Sin pérdida de generalidad, vamos a suponer que el proceso $Y_{t,T}$ esta centrado en cero (al igual que en el apéndice B). Debido al no acotamiento de la densidad espectral $f_\nu(u, \lambda)$ del proceso $\{\nu_{t,T}\}$, las demostraciones utilizaran las propiedades de las transformaciones de Fourier

$$\begin{aligned} \widehat{f}_\nu(u, \cdot) &:= \int_{-\pi}^{\pi} f_\nu(u, \lambda) e^{i\lambda \cdot} d\lambda \\ \widehat{f}_\nu(u, v, \cdot) &:= \int_{-\pi}^{\pi} A_\nu(u, \lambda) A_\nu(v, -\lambda) e^{i\lambda \cdot} d\lambda \end{aligned}$$

Proposición C.1. *Sea $f_Y(u, \lambda) = f_\nu(u, \lambda) + \sigma_\epsilon^2/2\pi$ la densidad espectral del proceso $\{Y_{t,T}\}$ definido en (3.9) y considere a $f_\nu(u, \lambda)$ como la densidad espectral del proceso $\{\nu_{t,T}\}$ que satisface el supuesto 3.1 A1 del capítulo 3, además, asuma que la función $\phi(u, \lambda)$ en la ecuación (C.1) es simétrica en torno a λ y dos veces diferenciables con respecto a u . Sea $\widehat{f}_\nu(u, \lambda)$ y $\widehat{\phi}(u, \lambda)$ los coeficientes de Fourier, respectivamente. Si existe una constante positiva K tal que*

$$\left| \widehat{f}_\nu(u, k) \widehat{\phi}(u, k) \right| \leq K \left(\frac{\log(k)}{k^2} \right), \quad (\text{C.4})$$

para todo $u \in [0, 1]$ y $k \geq 1$, entonces bajo el supuesto 3.1 A2 y A3 del capítulo 3, se tiene que

$$\mathbb{E}[J_T(\phi)] = J(\phi) + \mathcal{O}\left(\frac{\log^2(N)}{N}\right) + \mathcal{O}\left(\frac{1}{M}\right). \quad (\text{C.5})$$

DEMOSTRACIÓN. Aplicando $\mathbb{E}(\cdot)$ sobre el funcional muestral, se tiene que

$$\begin{aligned}
\mathbb{E}[J_T(\phi)] &= \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \mathbb{E}[I_N(u_j, \lambda)] d\lambda \\
&= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \mathbb{E}|D_N(u_j, \lambda)|^2 d\lambda \\
&= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s,t=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \mathbb{C}(u_j, t, s) e^{i\lambda(s-t)} d\lambda,
\end{aligned}$$

donde

$$\mathbb{C}_Y(u, t, s) = \mathbf{Cov}(Y_{[uT]-N/2+t+1,T}; Y_{[uT]-N/2+s+1,T}),$$

y por la ortogonalidad de $\{\nu_{t,T}\}$ y $\{\epsilon_t\}$ se tiene que

$$\begin{aligned}
\mathbb{C}_Y(u, t, s) &= \mathbf{Cov}(\nu_{[uT]-N/2+t+1,T}; \nu_{[uT]-N/2+s+1,T}) + \mathbf{Cov}(\epsilon_{[uT]-N/2+t+1}; \epsilon_{[uT]-N/2+s+1}) \\
&= \mathbb{C}_\nu(u_j, t, s) + \mathbf{Cov}(\epsilon_{[uT]-N/2+t+1}; \epsilon_{[uT]-N/2+s+1}),
\end{aligned}$$

note que para el ruido blanco $\{\epsilon_t\}$ se tiene

$$\mathbf{Cov}(\epsilon_{[uT]-N/2+t+1}; \epsilon_{[uT]-N/2+s+1}) = \begin{cases} \sigma_\epsilon^2, & \text{si } s = t, \\ 0, & \text{en otro caso.} \end{cases}$$

Utilizando la descomposición de $\mathbb{C}_Y(u, t, s)$ se obtiene

$$\begin{aligned}
\mathbb{E}[J_T(\phi)] &= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s,t=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \mathbb{C}_\nu(u_j, t, s) e^{i\lambda(s-t)} d\lambda \\
&+ \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right)^2 \sigma_\epsilon^2 d\lambda
\end{aligned}$$

Para el primer término de la suma

$$\frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s,t=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \mathbb{C}_{\nu}(u_j, t, s) e^{i\lambda(s-t)} d\lambda,$$

Palma and Olea (2010) mostraron que si $\{\nu_{t,T}\}$ es un proceso Gaussiano localmente estacionario de larga memoria, que satisface los supuestos 3.1 A1, A2 y A3, la suma es igual a

$$\int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) f_{\nu}(u, \lambda) d\lambda du + \mathcal{O}\left(\frac{\log^2(N)}{N}\right) + \mathcal{O}\left(\frac{1}{M}\right).$$

Para el segundo termino se tiene que

$$\begin{aligned} \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right)^2 \sigma_{\epsilon}^2 d\lambda &= \frac{\sigma_{\epsilon}^2}{2\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) d\lambda \\ &= \frac{1}{M} \sum_{j=1}^M A(u_j), \end{aligned}$$

donde $A(u) = \int_{-\pi}^{\pi} \phi(u, \lambda) \frac{\sigma_{\epsilon}^2}{2\pi} d\lambda$, así entonces

$$\begin{aligned} \frac{1}{M} \sum_{j=1}^M A(u_j) &= \int_0^1 A(u) du + \mathcal{O}\left(\frac{1}{M}\right) \\ &= \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) \frac{\sigma_{\epsilon}^2}{2\pi} d\lambda du + \mathcal{O}\left(\frac{1}{M}\right). \end{aligned}$$

Finalmente, sumando ambos resultados, se tiene que:

$$\mathbb{E}[J_T(\phi)] = \int_0^1 \int_{-\pi}^{\pi} \phi(u, \lambda) \left\{ f_{\nu}(u, \lambda) + \frac{\sigma_{\epsilon}^2}{2\pi} \right\} d\lambda du + \mathcal{O}\left(\frac{\log^2(N)}{N}\right) + \mathcal{O}\left(\frac{1}{M}\right).$$

Proposición C.2. Sea $f_y(u, \lambda) = f_\nu(u, \lambda) + \sigma_\epsilon^2/2\pi$ la densidad espectral del proceso $\{Y_{t,T}\}$ definido en (3.9) donde $f_\nu(u, \lambda)$ corresponde a la densidad espectral de variación en el tiempo del proceso $\{\nu_{t,T}\}$, que satisface el supuesto A1 del capítulo 3. Sea $\phi_1, \phi_2 : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{R}$ dos funciones tales que $\phi_1(u, \lambda)$ y $\phi_2(u, \lambda)$ son simétricas en torno a λ , dos veces diferenciables con respecto a u y sus coeficientes de Fourier satisfacen $|\widehat{\phi}_1(u, k)|, |\widehat{\phi}_2(u, k)| \leq K|k|^{-2d(u)-1}$ para $u \in [0, 1]$ y $|k| > 1$. Si se tiene el supuesto A2 y A3, entonces

$$\begin{aligned} \lim_{T \rightarrow \infty} T \mathbf{Cov} [J_T(\phi_1), J_T(\phi_2)] = & \\ & 4\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \lambda) \left(f_\nu(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi} \right)^2 d\lambda du \\ & + \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) d\lambda d\mu du. \end{aligned}$$

DEMOSTRACIÓN. Aplicando directamente la covarianza sobre el funcional muestral, se tiene que

$$T \mathbf{Cov} [J_T(\phi_1), J_T(\phi_2)] = \frac{T}{M^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \mathbf{Cov} [I_N(u_j, \lambda), I_N(u_k, \mu)] d\lambda d\mu.$$

Donde

$$\begin{aligned} \mathbf{Cov} [I_N(u_j, \lambda), I_N(u_k, \lambda)] & \\ &= \frac{1}{[2\pi H_{2,N}(0)]^2} \mathbf{Cov} (|D_N(u_j, \lambda)|^2, |D_N(u_k, \mu)|^2) \\ &= \frac{1}{[2\pi H_{2,N}(0)]^2} \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\ &\quad \times e^{i\lambda(s-t)+i\mu(m-p)} \\ &\quad \times \mathbf{Cov} \left(Y_{[u_j T]-N/2+s+1, T} Y_{[u_j T]-N/2+t+1, T}; Y_{[u_k T]-N/2+p+1, T} Y_{[u_k T]-N/2+m+1, T} \right). \end{aligned}$$

Como se asume que $\mathbb{E}(Y_{t,T}) = 0$ se tiene que

$$\begin{aligned}
& \text{Cov} \left(Y_{[u_j T] - N/2 + s + 1, T} Y_{[u_j T] - N/2 + t + 1, T}; Y_{[u_k T] - N/2 + p + 1, T} Y_{[u_k T] - N/2 + m + 1, T} \right) \\
&= \text{Cov} \left(\nu_{[u_j T] - N/2 + s + 1, T} \nu_{[u_j T] - N/2 + t + 1, T}; \nu_{[u_k T] - N/2 + p + 1, T} \nu_{[u_k T] - N/2 + m + 1, T} \right) \\
&+ \mathbb{E} \left(\nu_{[u_j T] - N/2 + s + 1, T} \nu_{[u_k T] - N/2 + p + 1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T] - N/2 + t + 1} \epsilon_{[u_k T] - N/2 + m + 1} \right) \\
&+ \mathbb{E} \left(\nu_{[u_j T] - N/2 + s + 1, T} \nu_{[u_k T] - N/2 + m + 1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T] - N/2 + t + 1} \epsilon_{[u_k T] - N/2 + p + 1} \right) \\
&+ \mathbb{E} \left(\nu_{[u_j T] - N/2 + t + 1, T} \nu_{[u_k T] - N/2 + p + 1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T] - N/2 + s + 1} \epsilon_{[u_k T] - N/2 + m + 1} \right) \\
&+ \mathbb{E} \left(\nu_{[u_j T] - N/2 + t + 1, T} \nu_{[u_k T] - N/2 + m + 1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T] - N/2 + s + 1} \epsilon_{[u_k T] - N/2 + p + 1} \right) \\
&+ \text{Cov} \left(\epsilon_{[u_j T] - N/2 + s + 1} \epsilon_{[u_j T] - N/2 + t + 1}; \epsilon_{[u_k T] - N/2 + p + 1} \epsilon_{[u_k T] - N/2 + m + 1} \right).
\end{aligned}$$

Se definen las siguientes cantidades

$$\begin{aligned}
B_N^{(1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t) + i\mu(m-p)} \\
&\times \text{Cov} \left(\nu_{[u_j T] - N/2 + s + 1, T} \nu_{[u_j T] - N/2 + t + 1, T}; \nu_{[u_k T] - N/2 + p + 1, T} \nu_{[u_k T] - N/2 + m + 1, T} \right) d\lambda d\mu, \\
B_N^{(2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t) + i\mu(m-p)} \\
&\times \mathbb{E} \left(\nu_{[u_j T] - N/2 + s + 1, T} \nu_{[u_k T] - N/2 + p + 1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T] - N/2 + t + 1} \epsilon_{[u_k T] - N/2 + m + 1} \right) d\lambda d\mu \\
B_N^{(3)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu), \\
&\times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t) + i\mu(m-p)} \\
&\times \mathbb{E} \left(\nu_{[u_j T] - N/2 + s + 1, T} \nu_{[u_k T] - N/2 + m + 1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T] - N/2 + t + 1} \epsilon_{[u_k T] - N/2 + p + 1} \right) d\lambda d\mu,
\end{aligned}$$

$$\begin{aligned}
B_N^{(4)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbb{E} \left(\nu_{[u_j T]-N/2+t+1, T} \nu_{[u_k T]-N/2+p+1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T]-N/2+s+1} \epsilon_{[u_k T]-N/2+m+1} \right) d\lambda d\mu, \\
B_N^{(5)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbb{E} \left(\nu_{[u_j T]-N/2+t+1, T} \nu_{[u_k T]-N/2+m+1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T]-N/2+s+1} \epsilon_{[u_k T]-N/2+p+1} \right) d\lambda d\mu, \\
B_N^{(6)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbf{Cov} \left(\epsilon_{[u_j T]-N/2+s+1} \epsilon_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+p+1} \epsilon_{[u_k T]-N/2+m+1} \right) d\lambda d\mu.
\end{aligned}$$

Note que $B_N^{(2)} = B_N^{(3)} = B_N^{(4)} = B_N^{(5)}$, entonces

$$T \mathbf{Cov} [J_T(\phi_1), J_T(\phi_2)] = B_N^{(1)} + 4B_N^{(2)} + B_N^{(6)},$$

La primera covarianza $B_N^{(1)}$ fue desarrollado por Palma and Olea (2010) (ver proposición 2). Entonces, primero nos enfocaremos en desarrollar el término cruzado $B_N^{(2)}$. Considere

$$\begin{aligned}
&\sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbb{E} \left(\nu_{[u_j T]-N/2+s+1, T} \nu_{[u_k T]-N/2+p+1, T} \right) \mathbb{E} \left(\epsilon_{[u_j T]-N/2+t+1} \epsilon_{[u_k T]-N/2+m+1} \right) d\lambda d\mu
\end{aligned}$$

desarrollando se tiene que

$$\begin{aligned}
&= \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \int_{-\pi}^{\pi} A_{\nu; t_j-N/2+s+1, T}^0(x) \overline{A_{\nu; t_k-N/2+p+1, T}^0(x)} e^{ix(t_j-t_k+s-p)} dx \times \int_{-\pi}^{\pi} \frac{\sigma_\epsilon^2}{2\pi} e^{iy(t_j-t_k+t-m)} dy \\
&= \int_{-\pi}^{\pi} \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) A_{\nu; t_j-N/2+s+1, T}^0(x) e^{is(x+\lambda)} \sum_{p=0}^{N-1} h\left(\frac{p}{N}\right) \overline{A_{\nu; t_k-N/2+p+1, T}^0(x)} e^{-ip(x+\mu)} \\
&\quad \times e^{ix(t_j-t_k)} dx \int_{-\pi}^{\pi} \sum_{t=0}^{N-1} h\left(\frac{t}{N}\right) \frac{\sigma_\epsilon}{\sqrt{2\pi}} e^{-it(\lambda-y)} \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right) \frac{\sigma_\epsilon}{\sqrt{2\pi}} e^{-im(y-\mu)} e^{iy(t_j-t_k)} dy \\
&= \int_{-\pi}^{\pi} H_N\left(A_{\nu; t_j-N/2+1+, T}^0(x) h\left(\frac{\cdot}{N}\right), -x-\lambda\right) H_N\left(\overline{A_{\nu; t_k-N/2+1+, T}^0(x)} h\left(\frac{\cdot}{N}\right), x+\mu\right) \\
&\quad \times e^{ix(t_j-t_k)} dx \int_{-\pi}^{\pi} H_N\left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), \lambda-y\right) H_N\left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), y-\mu\right) e^{iy(t_j-t_k)} dy \\
&= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} H_N\left(A_{\nu; t_j-N/2+1+, T}^0(x) h\left(\frac{\cdot}{N}\right), -x-\lambda\right) H_N\left(\overline{A_{\nu; t_k-N/2+1+, T}^0(x)} h\left(\frac{\cdot}{N}\right), x+\mu\right) \\
&\quad \times H_N\left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), \lambda-y\right) H_N\left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), y-\mu\right) e^{i(x+y)(t_j-t_k)} dx dy
\end{aligned}$$

Con este resultado, la cantidad $B_N^{(2)}$ queda

$$\begin{aligned}
B_N^{(2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times H_N\left(A_{\nu; t_j-N/2+1+, T}^0(x) h\left(\frac{\cdot}{N}\right), -x-\lambda\right) H_N\left(\overline{A_{\nu; t_k-N/2+1+, T}^0(x)} h\left(\frac{\cdot}{N}\right), x+\mu\right) \\
&\quad \times H_N\left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), \lambda-y\right) H_N\left(\frac{\sigma_\epsilon}{\sqrt{2\pi}} h\left(\frac{\cdot}{N}\right), y-\mu\right) e^{i(x+y)(t_j-t_k)} dx dy d\mu d\lambda \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times H_N\left(A_{\nu; t_j-N/2+1+, T}^0(x) h\left(\frac{\cdot}{N}\right), -x-\lambda\right) H_N\left(\overline{A_{\nu; t_k-N/2+1+, T}^0(x)} h\left(\frac{\cdot}{N}\right), x+\mu\right) \\
&\quad \times H_N(\lambda-y) H_N(y-\mu) e^{i(x+y)(t_j-t_k)} dx dy d\mu d\lambda
\end{aligned}$$

aplicando el lema A.5 el termino anterior, se tiene que

$$\begin{aligned}
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \left\{ \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \right. \\
&\quad \times A_\nu(u_j, x) H_N(-x - \lambda) A_\nu(u_k, -x) H_N(x + \mu) \\
&\quad \left. \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dy dx d\mu d\lambda + \tilde{R}_N \right\} \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \left\{ \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, y) \right. \\
&\quad \times A_\nu(u_j, x) H_N(-x - \lambda) A_\nu(u_k, -x) H_N(x + \mu) \\
&\quad \left. \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dy dx d\mu d\lambda + \tilde{\Delta}_N^{(1)} + \tilde{\Delta}_N^{(2)} + \tilde{R}_N \right\} \tag{C.6}
\end{aligned}$$

donde las expresiones $\tilde{\Delta}_N^{(1)}$ y $\tilde{\Delta}_N^{(2)}$ son

$$\begin{aligned}
\tilde{\Delta}_N^{(1)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \phi_2(u_k, \mu) \\
&\quad \times A_\nu(u_j, x) H_N(-x - \lambda) A_\nu(u_k, -x) H_N(x + \mu) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dx dy d\mu d\lambda \\
\tilde{\Delta}_N^{(2)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) [\phi_2(u_k, \mu) - \phi_2(u_k, y)] \\
&\quad \times A_\nu(u_j, x) H_N(-x - \lambda) A_\nu(u_k, -x) H_N(x + \mu) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dy dx d\mu d\lambda
\end{aligned}$$

Note que si se integra el término $\tilde{\Delta}_N^{(1)}$ con respecto a μ se tiene que

$$\begin{aligned}
\tilde{\Delta}_N^{(1)} &= \sum_{j,k=1}^M \sum_{p,m=0}^{N-1} h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \hat{\phi}_2(u_k, m - p) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \\
&\quad \times A_\nu(u_j, x) A_\nu(u_k, -x) H_N(-x - \lambda) H_N(\lambda - y) \\
&\quad \times e^{i(x+y)(t_j - t_k) - i x p - i y m} dx dy d\lambda
\end{aligned}$$

y posteriormente se integra con respecto a y obtenemos

$$\begin{aligned}\tilde{\Delta}_N^{(1)} &= \sum_{j,k=1}^M \sum_{t,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\ &\quad \times \hat{\phi}_2(u_k, m-p) \delta_0(t_j - t_k + t - m) \\ &\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \\ &\quad \times A_\nu(u_j, x) A_\nu(u_k, -x) H_N(-x - \lambda) e^{ix(t_j - t_k - p) - it\lambda} d\lambda dx\end{aligned}$$

Por lo simetría de $\phi_1(u, \lambda)$ en torno a λ , se tiene que es $\tilde{\Delta}_N^{(1)}$ es igual a

$$\begin{aligned}\tilde{\Delta}_N^{(1)} &= \sum_{j,k=1}^M \sum_{t,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\ &\quad \times \hat{\phi}_2(u_k, m-p) \delta_0(t_j - t_k + t - m) \\ &\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \\ &\quad \times A_\nu(u_j, x) A_\nu(u_k, -x) H_N(\lambda - x) e^{ix(t_j - t_k - p) + it\lambda} d\lambda dx.\end{aligned}$$

Note que $\tilde{\Delta}_N^{(1)}$ es similar al termino $\Delta_N^{(1)}$ definido en la proposición 2 de Palma and Olea (2010), con la única diferencia en el término δ_0 , que para el caso de $\Delta_N^{(1)}$ es \hat{f}_ν , es decir,

$$\begin{aligned}\Delta_N^{(1)} &= \sum_{j,k=1}^M \sum_{t,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\ &\quad \times \hat{\phi}_2(u_k, m-p) \hat{f}(u_j, u_k, t_j - t_k + t - m) \\ &\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \\ &\quad \times A_\nu(u_j, x) A_\nu(u_k, -x) H_N(\lambda - x) e^{ix(t_j - t_k - p) + it\lambda} d\lambda dx,\end{aligned}$$

donde Palma and Olea (2010) muestran que es acotado por

$$\begin{aligned}
|\Delta_N^{(1)}| &\leq \sum_{j,k=1}^M \sum_{t,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\
&\quad \times |\widehat{\phi}_2(u_k, m-p)| |\widehat{f}_\nu(u_j, u_k, t_j - t_k + t - m)| \\
&\quad \times \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \right. \\
&\quad \quad \times A_\nu(u_j, x) A_\nu(u_k, -x) H_N(\lambda - x) e^{ix(t_j - t_k - p) + it\lambda} d\lambda dx \left. \right| \\
&\leq K \left(\frac{N}{S}\right)^2 M \log M + M^2.
\end{aligned}$$

Entonces se puede elegir un valor K finito, tal que $|\delta_0(s)| \leq K |\widehat{f}_\nu(u_j, u_k, s)| \forall s \in \mathbb{Z}$,
y así utilizar la cota de Palma and Olea (2010) para acotar $\widetilde{\Delta}_N^{(1)}$,

$$\begin{aligned}
|\widetilde{\Delta}_N^{(1)}| &\leq \sum_{j,k=1}^M \sum_{t,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\
&\quad \times |\widehat{\phi}_2(u_k, m-p)| |\delta_0(t_j - t_k + t - m)| \\
&\quad \times \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \right. \\
&\quad \quad \times A_\nu(u_j, x) A_\nu(u_k, -x) H_N(\lambda - x) e^{ix(t_j - t_k - p) + it\lambda} d\lambda dx \left. \right| \\
&\leq K \sum_{j,k=1}^M \sum_{t,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) \\
&\quad \times |\widehat{\phi}_2(u_k, m-p)| |\widehat{f}_\nu(u_j, u_k, t_j - t_k + t - m)| \\
&\quad \times \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \right. \\
&\quad \quad \times A_\nu(u_j, x) A_\nu(u_k, -x) H_N(\lambda - x) e^{ix(t_j - t_k - p) + it\lambda} d\lambda dx \left. \right| \\
&\leq K \left(\frac{N}{S}\right)^2 M \log M + M^2.
\end{aligned}$$

De manera similar se puede acotar $\widetilde{\Delta}_N^{(2)}$. Utilizando el lema A.5, el resto \widetilde{R}_N puede ser mayorado por

$$\begin{aligned} \left| \tilde{R}_N \right| &\leq \frac{N}{T} \frac{\sigma_\epsilon^2}{2\pi} \left| \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \right. \\ &\quad \times A_\nu(u_j, x) H_N(-x - \lambda) |x|^{-d(u_k)} L_N(x + \mu) \\ &\quad \left. \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dy dx d\mu d\lambda \right|. \end{aligned}$$

Integramos con respecto a λ se tiene

$$\begin{aligned} \left| \tilde{R}_N \right| &\leq \frac{N}{T} \frac{\sigma_\epsilon^2}{2\pi} \left| \sum_{j,k=1}^M \sum_{t,s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \hat{\phi}_1(u_j, s - t) \right. \\ &\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_2(u_k, \mu) A_\nu(u_j, x) |x|^{-d(u_k)} L_N(x + \mu) \\ &\quad \left. \times H_N(y - \mu) e^{ix(t_j - t_k + s) + iy(t_j - t_k + t)} dy dx d\mu \right| \\ &= \frac{N}{T} \frac{\sigma_\epsilon^2}{2\pi} \left| \sum_{j,k=1}^M \sum_{t,s,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{m}{N}\right) \right. \\ &\quad \times \hat{\phi}_1(u_j, s - t) \delta_0(t_j - t_k + t - m) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_2(u_k, \mu) A_\nu(u_j, x) \\ &\quad \left. \times |x|^{-d(u_k)} L_N(x + \mu) e^{ix(t_j - t_k + s) + im\mu} dx d\mu \right| \end{aligned}$$

aquí se puede recurrir a un argumento similar al utilizado en $\tilde{\Delta}_N^{(1)}$, ya que podemos utilizar la cota de R_N de Palma and Olea (2010)

$$\begin{aligned} \left| \tilde{R}_N \right| &\leq \frac{N}{T} \frac{\sigma_\epsilon^2}{2\pi} \left| \sum_{j,k=1}^M \sum_{t,s,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{m}{N}\right) \right. \\ &\quad \times \hat{\phi}_1(u_j, s - t) \delta_0(t_j - t_k + t - m) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_2(u_k, \mu) A_\nu(u_j, x) \\ &\quad \left. \times |x|^{-d(u_k)} L_N(x + \mu) e^{ix(t_j - t_k + s) + im\mu} dx d\mu \right| \end{aligned}$$

$$\begin{aligned}
&\leq \frac{N}{T} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \sum_{t,s,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{m}{N}\right) \\
&\quad \times |\widehat{\phi}_1(u_j, s-t)| |\delta_0(t_j - t_k + t - m)| \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_2(u_k, \mu) A_\nu(u_j, x) \right. \\
&\quad \left. \times |x|^{-d(u_k)} L_N(x + \mu) e^{ix(t_j - t_k + s) + im\mu} dx d\mu \right| \\
&\leq K \frac{N}{T} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \sum_{t,s,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{m}{N}\right) \\
&\quad \times |\widehat{\phi}_1(u_j, s-t)| |\widehat{f}_\nu(u_j, u_k, t_j - t_k + t - m)| \left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_2(u_k, \mu) A_\nu(u_j, x) \right. \\
&\quad \left. \times |x|^{-d(u_k)} L_N(x + \mu) e^{ix(t_j - t_k + s) + im\mu} dx d\mu \right| \\
&\leq KN^3 M^2 T^{d-2}.
\end{aligned}$$

Entonces se concluye que

$$\begin{aligned}
B_N^{(2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, x) H_N(-x - \lambda) A_\nu(u_k, -x) H_N(x + \mu) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dy dx d\mu d\lambda + \mathcal{O}\left(\frac{T}{N^2} + \frac{T \log M}{M S^2} + \frac{N}{T^{1-d}}\right).
\end{aligned}$$

Siguiendo con el desarrollo de $B_N^{(2)}$, vamos a reemplazar las $\phi_1(u_j, x)$ por $\phi_1(u_j, y)$, $A_\nu(u_j, x)$ por $A_\nu(u_j, -y)$ y $A_\nu(u_k, -x)$ por $A_\nu(u_k, y)$ respectivamente, así

$$\begin{aligned}
B_N^{(2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) H_N(-x - \lambda) A_\nu(u_k, y) H_N(x + \mu) \\
&\quad \times H_N(\lambda - y) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} dy dx d\mu d\lambda + o(1) \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) H_N(x + \mu) H_N(y - \mu) e^{i(x+y)(t_j - t_k)} \\
&\quad \times \sum_{s,t=0}^{N-1} h\left(\frac{s}{N}\right) h\left(\frac{t}{N}\right) e^{isx + ity} \int_{-\pi}^{\pi} e^{-i\lambda(t-s)} d\lambda dy dx d\mu + o(1) \\
&= \frac{2\pi \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) e^{i(x+y)(t_j - t_k)} \sum_{s=0}^{N-1} h^2\left(\frac{s}{N}\right) e^{is(x+y)} \\
&\quad \times \sum_{m,p=0}^{N-1} h\left(\frac{m}{N}\right) h\left(\frac{p}{N}\right) e^{-imy - ipx} \int_{-\pi}^{\pi} e^{-i\mu(p-m)} d\mu dy dx + o(1) \\
&= \frac{[2\pi]^2 \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) e^{i(x+y)(t_j - t_k)} \sum_{s=0}^{N-1} h^2\left(\frac{s}{N}\right) e^{is(x+y)} \\
&\quad \times \sum_{m=0}^{N-1} h^2\left(\frac{m}{N}\right) e^{-im(y+x)} dy dx + o(1) \\
&= \frac{[2\pi]^2 \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \sum_{s,m=0}^{N-1} h^2\left(\frac{m}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, y) \phi_2(u_k, y) \\
&\quad \times A_\nu(u_j, -y) A_\nu(u_k, y) e^{i(x+y)(t_j - t_k + s - m)} dy dx + o(1),
\end{aligned}$$

utilizando el teorema cambio de variable $z = x + y$ y $w = y$, tenemos

$$\begin{aligned}
B_N^{(2)} &= \frac{[2\pi]^2 \times T}{[2\pi M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{j,k=1}^M \sum_{s,m=0}^{N-1} h^2\left(\frac{m}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, w) \phi_2(u_k, w) \\
&\quad \times A_\nu(u_j, -w) A_\nu(u_k, w) e^{iz(t_j - t_k + s - m)} dz dw + o(1) \\
&= \frac{2\pi \times T}{[M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{s,m=0}^{N-1} \sum_{\substack{j,k=1 \\ S(j-k) = s-m}}^M h^2\left(\frac{m}{N}\right) h^2\left(\frac{s}{N}\right) \int_{-\pi}^{\pi} \phi_1(u_j, w) \phi_2(u_k, w) \\
&\quad \times A_\nu(u_j, -w) A_\nu(u_k, w) dw + o(1).
\end{aligned}$$

Para $S < N$ se puede escribir

$$\begin{aligned}
B_N^{(2)} &= \frac{2\pi \times T}{[M H_{2,N}(0)]^2} \frac{\sigma_\epsilon^2}{2\pi} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \\
&\quad \times \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_{j+p}, x) A_\nu(u_{j+p}, x) A_\nu(u_j, -x) dx + o(1).
\end{aligned}$$

Por los supuestos de esta proposición, el producto $\phi_1(u, x) f_\nu(u, x)$ y $\phi_2(u, x) f_\nu(u, x)$ son diferenciables con respecto a u , más aún, el límite $\lim_{S,T \rightarrow \infty} \frac{S|p|}{T} = 0$ para todo $|p| \leq \frac{N}{S}$, en consecuencia

$$\begin{aligned}
\frac{S}{T} \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_{j+p}, x) A_\nu(u_j, -x) A_\nu(u_{j+p}, x) \frac{\sigma_\epsilon^2}{2\pi} dx \\
\longrightarrow \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) f_\nu(u, x) \frac{\sigma_\epsilon^2}{2\pi} dx du,
\end{aligned}$$

para todo $|p| < \frac{N}{S}$ cuando $M, N, S, T \rightarrow \infty$. Por otro lado se tiene

$$\begin{aligned}
\frac{2\pi \times T^2 N^2}{S^2 [M H_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \frac{S}{N^2} \\
\longrightarrow 2\pi \int_0^1 \int_{-x}^{1-x} h^2(x) h^2(x+y) dx dy \left(\int_0^1 h^2(x) dx \right)^{-2} = 2\pi,
\end{aligned}$$

cuando $M, N, S, T \rightarrow \infty$. Con ambos resultados se tienen

$$B_N^{(2)} \longrightarrow 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) f_\nu(u, x) \frac{\sigma_\epsilon^2}{2\pi} dx du.$$

cuando $M, N, S, T \rightarrow \infty$. Para la covarianza $B_N^{(1)}$ utilizamos directamente la proposición 2) de Palma and Olea (2010), donde tenemos que

$$B_N^{(1)} \longrightarrow 4\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) f_\nu^2(u, x) dx du,$$

cuando $M, N, S, T \rightarrow \infty$.

Ahora vamos a estudiar el límite de $B_N^{(6)}$, donde

$$\begin{aligned} B_N^{(6)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\ &\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\ &\quad \times \mathbf{Cov} \left(\epsilon_{[u_j T]-N/2+s+1} \epsilon_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+p+1} \epsilon_{[u_k T]-N/2+m+1} \right) d\lambda d\mu, \end{aligned}$$

recuerde que $\{\epsilon_t, t \in \mathbb{Z}\}$ es una secuencia independiente idénticamente distribuida, centrada en cero, usando el teorema de los cumulantes para polinomios 2.3.2 de Brillinger (1981) se tiene que:

$$\begin{aligned} B_N^{(6)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\ &\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\ &\quad \times \left\{ \mathbf{Cov} \left(\epsilon_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+m+1} \right) \times \mathbf{Cov} \left(\epsilon_{[u_j T]-N/2+s+1}; \epsilon_{[u_k T]-N/2+p+1} \right) \right. \\ &\quad + \mathbf{Cov} \left(\epsilon_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+p+1} \right) \times \mathbf{Cov} \left(\epsilon_{[u_j T]-N/2+s+1}; \epsilon_{[u_k T]-N/2+m+1} \right) \\ &\quad \left. + \mathbf{Cum} \left(\epsilon_{[u_j T]-N/2+s+1}, \epsilon_{[u_j T]-N/2+t+1}, \epsilon_{[u_k T]-N/2+p+1}, \epsilon_{[u_k T]-N/2+m+1} \right) \right\} d\lambda d\mu, \end{aligned}$$

de esta forma $B_N^{(6)}$ se puede descomponer en los siguientes terminos

$$B_N^{(6)} = B_N^{(6.1)} + B_N^{(6.2)} + B_N^{(6.3)},$$

donde

$$\begin{aligned}
B_N^{(6.1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbf{Cov}\left(\epsilon_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+m+1}\right) \times \mathbf{Cov}\left(\epsilon_{[u_j T]-N/2+s+1}; \epsilon_{[u_k T]-N/2+p+1}\right) d\lambda d\mu, \\
B_N^{(6.2)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbf{Cov}\left(\epsilon_{[u_j T]-N/2+t+1}; \epsilon_{[u_k T]-N/2+p+1}\right) \times \mathbf{Cov}\left(\epsilon_{[u_j T]-N/2+s+1}; \epsilon_{[u_k T]-N/2+m+1}\right) d\lambda d\mu, \\
B_N^{(6.3)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \mathbf{Cum}\left(\epsilon_{[u_j T]-N/2+s+1}, \epsilon_{[u_j T]-N/2+t+1}, \epsilon_{[u_k T]-N/2+p+1}, \epsilon_{[u_k T]-N/2+m+1}\right) d\lambda d\mu.
\end{aligned}$$

El termino $B_N^{(6.1)}$ se puede expandir de la siguiente manera

$$\begin{aligned}
B_N^{(6.1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{ix(t_j-t_k+t-m)+iy(t_j-t_k+s-p)} dx dy d\lambda d\mu \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \left\{ \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, y) \right. \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \left. \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu + \Delta_N^{(6.1.1)} + \Delta_N^{(6.1.2)} \right\}
\end{aligned}$$

con

$$\begin{aligned}
\Delta_N^{(6.1.1)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \phi_2(u_k, \mu) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu \\
\Delta_N^{(6.1.2)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) [\phi_2(u_k, \mu) - \phi_2(u_k, y)] \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j-t_k)} dx dy d\lambda d\mu
\end{aligned}$$

desarrollando $\Delta_N^{(6.1.1)}$ se tiene

$$\begin{aligned}
\Delta_N^{(6.1.1)} &= \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \phi_2(u_k, \mu) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j - t_k)} dx dy d\lambda d\mu \\
&= \sum_{j,k=1}^M \sum_{t,s=0}^{N-1} h\left(\frac{s}{N}\right) h\left(\frac{t}{N}\right) \widehat{\phi}_2(u_k, s - t) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] \\
&\quad \times e^{-isx - ity} H_N(\lambda - x) H_N(-y - \lambda) e^{i(x+y)(t_j - t_k)} dx dy d\lambda \\
&= 2\pi \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} h\left(\frac{s}{N}\right) h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) \widehat{\phi}_2(u_k, s - t) \delta_0(t_j - t_k + p - t) \\
&\quad \times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] H_N(\lambda - x) e^{ix(t_j - t_k - s) + ip\lambda} dx d\lambda \\
&= 2\pi \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} h\left(\frac{s}{N}\right) h\left(\frac{t}{N}\right) h\left(\frac{p}{N}\right) \widehat{\phi}_2(u_k, s - t) \delta_0(t_j - t_k + p - t) \\
&\quad \times \widetilde{\varepsilon}_N(u_j, u_k, p, t_j - t_k - s)
\end{aligned}$$

donde $\widetilde{\varepsilon}_N(u_j, u_k, p, t_j - t_k - s)$ está dado por

$$\begin{aligned}
&\widetilde{\varepsilon}_N(u_j, u_k, p, t_j - t_k - s) \\
&\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] H_N(\lambda - x) e^{ix(t_j - t_k - s) + ip\lambda} dx d\lambda \\
&= \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right) \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [\phi_1(u_j, \lambda) - \phi_1(u_j, x)] e^{ix(t_j - t_k + m - s)} e^{i\lambda(p - m)} dx d\lambda \\
&= 2\pi \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right) \left[\widehat{\phi}_1(u_j, p - m) \delta_0(t_j - t_k + m - s) \right. \\
&\quad \left. - \widehat{\phi}_1(u_j, t_j - t_k + m - s) \delta_0(p - m) \right] \\
&= 2\pi \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right) \widehat{\phi}_1(u_j, p - m) \delta_0(t_j - t_k + m - s) \\
&\quad - 2\pi h\left(\frac{p}{N}\right) \widehat{\phi}_1(u_j, t_j - t_k + p - s)
\end{aligned}$$

utilizando el teorema del valor medio $h\left(\frac{m}{N}\right) = h\left(\frac{p}{N}\right) + h'(\xi_{p,m}) \frac{m-p}{N}$ para algún

$\xi_{p,m} \in [0, 1]$. Así

$$\begin{aligned}
& \tilde{\varepsilon}_N(u_j, u_k, p, t_j - t_k - s) \\
&= h \left(\frac{p}{N} \right) \left\{ \sum_{m=0}^{N-1} \hat{\phi}_1(u_j, p - m) 2\pi \delta_0(t_j - t_k + m - s) - 2\pi \hat{\phi}_1(u_j, t_j - t_k + m - s) \right\} \\
&\quad + \sum_{m=0}^{N-1} h'(\xi_{p,m}) \frac{m-p}{N} \hat{\phi}_1(u_j, p - m) 2\pi \delta_0(t_j - t_k + m - s) \\
&= h \left(\frac{p}{N} \right) \tilde{\varepsilon}_N^{(1)}(u_j, u_k, p, t_j - t_k - s) + \tilde{\varepsilon}_N^{(2)}(u_j, u_k, p, t_j - t_k - s),
\end{aligned}$$

el término $\tilde{\varepsilon}_N(u_j, u_k, p, t_j - t_k - s)$ es acotado por $\varepsilon_N(u_j, u_k, p, t_j - t_k - s)$ definido en la proposición 2 de Palma and Olea (2010), dado que es el mismo caso, solo que aquí se considera la función de transferencia constante. De esta forma podemos utilizar la misma cota de $\tilde{\Delta}_N^{(1)}$

$$\begin{aligned}
|\Delta_N^{(6.1.1)}| &\leq \sum_{j,k=1}^M \sum_{t,s,p=0}^{N-1} h \left(\frac{s}{N} \right) h \left(\frac{t}{N} \right) h \left(\frac{p}{N} \right) |\hat{\phi}_2(u_k, s - t)| |2\pi \delta_0(t_j - t_k + p - t)| \\
&\quad \times |\tilde{\varepsilon}_N(u_j, u_k, p, t_j - t_k - s)| \\
&\leq K \left(\frac{N}{S} \right)^2 M \log M + M^2.
\end{aligned}$$

El mismo desarrollo se utiliza para acotar $\Delta_N^{(6.1.2)}$. Con esto se tiene que

$$\begin{aligned}
B_N^{(6.1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, y) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j - t_k)} dx dy d\lambda d\mu + \mathcal{O} \left(\frac{\log M}{S} + \frac{T}{N^2} \right).
\end{aligned}$$

Similar al calculo de la covarianza cruzada, reemplazamos $\phi_2(\mu_k, y)$ por $\phi_2(\mu_k, x)$,

$$\begin{aligned}
B_N^{(6.1)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, x) \\
&\quad \times H_N(\lambda - x) H_N(x - \mu) H_N(\mu + y) H_N(-y - \lambda) \\
&\quad \times e^{i(x+y)(t_j - t_k)} dx dy d\lambda d\mu + o(1).
\end{aligned}$$

Integrando sobre μ y λ se tiene que

$$\begin{aligned}
B_N^{(6.1)} &= \frac{T}{[M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, x) \\
&\quad \times |H_{2,N}(x+y)|^2 \times e^{i(x+y)(t_j - t_k)} dx dy + o(1) \\
&= \frac{T}{[M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, x) \\
&\quad \times |H_{2,N}(x+y)|^2 \times e^{i(x+y)[S(j-k)]} dx dy + o(1) \\
&= \frac{T}{[M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, x) \\
&\quad \times |H_{2,N}(z)|^2 \times e^{iz[S(j-k)]} dx dz + o(1) \\
&= \frac{T}{[M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{j,k=1}^M \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, x) \\
&\quad \times \sum_{t,s=0}^{N-1} h^2 \left(\frac{t}{N} \right) h^2 \left(\frac{s}{N} \right) \int_{-\pi}^{\pi} e^{iz[S(j-k)+t-s]} dz dx + o(1) \\
&= \frac{2\pi \times T}{[M H_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{t,s=0}^{N-1} \sum_{\substack{j,k=1 \\ S(j-k)=s-t}}^M h^2 \left(\frac{t}{N} \right) h^2 \left(\frac{s}{N} \right) \\
&\quad \times \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_k, x) dx + o(1).
\end{aligned}$$

Note que se tiene como supuesto que $S < N$, $B_N^{(6.1)}$ se puede escribir

$$B_N^{(6.1)} = \frac{2\pi \times T}{[MH_{2,N}(0)]^2} \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2 \left(\frac{t}{N} \right) h^2 \left(\frac{t}{N} + \frac{pS}{N} \right) \\ \times \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_{j+p}, x) dx + o(1)$$

Por el supuesto 3.1 A.3, se tiene que $\lim_{S,T \rightarrow \infty} \frac{S|p|}{T} = 0$ para todo $|p| \leq \frac{N}{S}$, así

$$\frac{S}{T} \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \phi_1(u_j, x) \phi_2(u_{j+p}, x) dx \longrightarrow \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) dx du,$$

y como lo vimos anteriormente, se tiene

$$\frac{2\pi \times T^2 N^2}{S^2 [MH_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2 \left(\frac{t}{N} \right) h^2 \left(\frac{t}{N} + \frac{pS}{N} \right) \frac{S}{N^2} \\ \longrightarrow 2\pi \int_0^1 \int_{-x}^{1-x} h^2(x) h^2(x+y) dx dy \left(\int_0^1 h^2(x) dx \right)^{-2} = 2\pi,$$

Así finalmente se tiene que

$$B_N^{(6.1)} \longrightarrow 2\pi \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) dx du,$$

cuando $M, N, S, T \rightarrow \infty$. Si se procede de la misma manera, se puede mostrar que

$$B_N^{(6.2)} \longrightarrow 2\pi \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) dx du,$$

cuando $M, N, S, T \rightarrow \infty$. Para el caso de $B_N^{(6.3)}$, recuerde que la secuencia $\{\epsilon_t, t \in \mathbb{Z}\}$

es independiente, idénticamente distribuida centrada en cero, entonces

$$\text{Cum}(\epsilon_s, \epsilon_t, \epsilon_p, \epsilon_m) = \begin{cases} \sigma_\epsilon^4 \kappa_4, & \text{si } s = t = p = m, \\ 0 & \text{en otro caso.} \end{cases}$$

donde $\kappa_4 = \mathbb{E}(z_1^4) - 3\mathbb{E}^2(z_1)$, con $\epsilon_t = \sigma_\epsilon z_t$ y $\{z_t\}$ una secuencia centrada en cero y de varianza uno (usualmente ϵ_t es un secuencia $\log \chi_1^2$). Para $B_N^{(6.3)}$ se tiene

$$\begin{aligned}
B_N^{(6.3)} &= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \text{Cum} \left(\epsilon_{[u_j T]-N/2+s+1}, \epsilon_{[u_j T]-N/2+t+1}, \epsilon_{[u_k T]-N/2+p+1}, \epsilon_{[u_k T]-N/2+m+1} \right) d\lambda d\mu \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right) h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\quad \times \frac{\kappa_4 \times \sigma_\epsilon^4}{(2\pi)^3} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{ix(t_j-t_k+s-m)} e^{iy(t_j-t_k+t-m)} e^{iz(p-m)} dx dy dz d\lambda d\mu \\
&= \frac{T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h^2\left(\frac{m}{N}\right) e^{i\lambda(s-t)} \\
&\quad \times \frac{\kappa_4 \times \sigma_\epsilon^4}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{ix(t_j-t_k+s-m)} e^{iy(t_j-t_k+t-m)} dx dy d\lambda d\mu \\
&= \frac{\kappa_4 \times \sigma_\epsilon^4 \times T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,s,m=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h^2\left(\frac{m}{N}\right) e^{i\lambda(s-t)} \\
&\quad \times \delta_0(t_j - t_k + s - m) \delta_0(t_j - t_k + t - m) d\lambda d\mu \\
&= \frac{\kappa_4 \times \sigma_\epsilon^4 \times T}{[2\pi M H_{2,N}(0)]^2} \sum_{j,k=1}^M \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_k, \mu) \\
&\quad \times \sum_{t,m=0}^{N-1} h^2\left(\frac{t}{N}\right) h^2\left(\frac{m}{N}\right) \delta_0(t_j - t_k + t - m) d\lambda d\mu
\end{aligned}$$

Utilizando el argumento anterior para $S < N$, podemos escribir

$$B_N^{(6.3)} = \frac{\kappa_4 \times \sigma_\epsilon^4 \times T}{[2\pi M H_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \\ \times \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_{j+p}, \mu) d\lambda d\mu$$

luego

$$\frac{S}{T} \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u_j, \lambda) \phi_2(u_{j+p}, \mu) d\lambda d\mu \longrightarrow \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) d\lambda d\mu du,$$

cuando $T, M, N, S \rightarrow \infty$. Además

$$\frac{T^2 \times N^2}{S^2 \times M^2 \times [H_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \frac{S}{N^2} \\ \longrightarrow \int_0^1 \int_{-x}^{1-x} h^2(x) h^2(x+y) dx dy \left(\int_0^1 h^2(x) dx \right)^{-2} = 1,$$

cuando $T, M, N, S \rightarrow \infty$. Utilizando ambas expresiones

$$B_N^{(6.3)} \longrightarrow \frac{\kappa_4 \sigma_\epsilon^4}{(2\pi)^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) d\lambda d\mu du$$

Entonces el termino $B_N^{(6)}$ se comporta

$$B_N^{(6)} = B_N^{(6.1)} + B_N^{(6.2)} + B_N^{(6.3)} \longrightarrow 4\pi \left(\frac{\sigma_\epsilon^2}{2\pi} \right)^2 \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, x) \phi_2(u, x) dx du \\ + \frac{\kappa_4 \sigma_\epsilon^4}{(2\pi)^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) d\lambda d\mu du$$

cuando $T, M, N, S \rightarrow \infty$. Finalmente se tiene que la covarianza

$$T \text{Cov} [J_T(\phi_1), J_T(\phi_2)] \longrightarrow 4\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \lambda) \left(f_\nu(u, \lambda) + \frac{\sigma_\epsilon^2}{2\pi} \right)^2 d\lambda du \\ + \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_1(u, \lambda) \phi_2(u, \mu) d\lambda d\mu du$$

cuando $M, N, S, T \rightarrow \infty$.

Proposición C.3. *Sea $\text{cum}_\ell(\cdot)$ el ℓ -ésimo cumulante con $\ell \geq 3$. Entonces, $T^{\ell/2} \text{cum}_\ell(J_T(\phi)) \rightarrow 0$, cuando $T \rightarrow \infty$.*

DEMOSTRACIÓN. El funcional lineal $J_T(\phi)$ está definido como

$$J_T(\phi) = \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) I_N(u_j, \lambda) d\lambda, \quad (\text{C.7})$$

donde vamos a definir $\Pi = (-\pi, \pi]$, entonces

$$T^{\ell/2} \text{cum}_\ell(J_T(\phi)) \\ = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \int_{\Pi^\ell} \left[\prod_{r=1}^{\ell} \phi(u_{j_r}, \lambda_r) \right] \text{cum} \left(D_N(u_{j_1}, \lambda_1) D_N(u_{j_1}, -\lambda_1), \right. \\ \left. \dots, D_N(u_{j_\ell}, \lambda_\ell) D_N(u_{j_\ell}, -\lambda_\ell) \right) d\lambda_1 \cdots d\lambda_\ell \\ = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \int_{\Pi^\ell} \left[\prod_{r=1}^{\ell} \phi(u_{j_r}, \lambda_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \exp\{-i\lambda_1(k_1 - m_1)\} \times \cdots \times \exp\{-i\lambda_\ell(k_\ell - m_\ell)\} \\ \times \text{cum} \left(Y_{[u_{j_1} T] - N/2 + k_1 + 1, T} Y_{[u_{j_1} T] - N/2 + n_1 + 1, T}, \dots, \right. \\ \left. Y_{[u_{j_\ell} T] - N/2 + k_\ell + 1, T} Y_{[u_{j_\ell} T] - N/2 + n_\ell + 1, T} \right) d\lambda_1 \cdots d\lambda_\ell \\ = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \text{cum} \left(Y_{[u_{j_1} T] - N/2 + k_1 + 1, T} Y_{[u_{j_1} T] - N/2 + n_1 + 1, T}, \dots, \right. \\ \left. Y_{[u_{j_\ell} T] - N/2 + k_\ell + 1, T} Y_{[u_{j_\ell} T] - N/2 + n_\ell + 1, T} \right),$$

aquí $\widehat{\phi}(u, k)$ es el k ésimo coeficiente de Fourier de $\phi(u, \lambda)$,

$$\widehat{\phi}(u, k) = \int_{-\pi}^{\pi} \phi(u, \lambda) \exp(-i\lambda k) d\lambda.$$

Al igual que en el apéndice B, se utilizara el teorema cumulantes para productos de variables aleatorias de Brillinger (1981) (teorema 2.3.2), denotamos la matriz

$$\begin{array}{cc} (1, 1) & (1, 2) \\ \vdots & \vdots \\ (\ell, 1) & (\ell, 2) \end{array}$$

donde (m, n) denota la posición, es decir, $Y_{[u_{j_m}]T-N/2+k_m+1, T}$ para $m = 1, \dots, \ell$ y $n = 1$ y $Y_{[u_{j_m}]T-N/2+n_m+1, T}$ para $m = 1, \dots, \ell$ y $n = 2$. Al aplicar el teorema, la suma se realiza sobre todas las particiones indescompensable $\{P_1, \dots, P_m\}$ del conjunto $\mathcal{S} = \{(1, 1), \dots, (\ell, 2)\}$, donde $|P_t| \geq 2$. Entonces

$$\begin{aligned} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^{\ell}} \sum_{j_1, \dots, j_{\ell}=1}^M \sum_{k_1, \dots, k_{\ell}=0}^{N-1} \sum_{n_1, \dots, n_{\ell}=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ &\quad \times \sum_{\{P_1, \dots, P_m\}} \prod_{t=1}^m \text{cum}(Y_{[u_{j_{\omega_1}}]T-N/2+\omega_2+1, T} \mid (\omega_1, \omega_2) \in P_t) \end{aligned}$$

Por la ortogonalidad de los procesos $\{\nu_{t,T}\}$ y $\{\epsilon_t\}$, se utiliza la descomposición de cumulantes,

$$\begin{aligned}
&= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}ip} \prod_{t=1}^m \left[\text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} \mid (\omega_1, \omega_2) \in P_t) \right. \\
&\quad \quad \left. + \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} \mid (\omega_1, \omega_2) \in P_t) \right] \\
&= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}ip} \left(\prod_{t=1}^m \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} \mid (\omega_1, \omega_2) \in P_t) \right. \\
&\quad \quad + \sum_{I \in \mathcal{I}} \prod_{t_1 \in I} \prod_{t_2 \in I^c} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} \mid (\omega_1, \omega_2) \in P_{t_1}) \\
&\quad \quad \quad \times \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} \mid (\omega_1, \omega_2) \in P_{t_2}) \\
&\quad \quad \quad \left. \left. + \prod_{t=1}^m \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} \mid (\omega_1, \omega_2) \in P_t) \right) \right),
\end{aligned}$$

donde $\mathcal{I} = \{I : I \subset \{1, \dots, m\}\}$ con $\{1, \dots, m\} \notin \mathcal{I}$, si $I \in \mathcal{I}$, entonces se define I^c como $I^c = \{1, \dots, m\} - I$. Note que el número de conjuntos de $\#\mathcal{I} = 2^m - 2$ y para todo $I \in \mathcal{I}$, se tiene que $1 < |I| < m$. Defina

$$\begin{aligned}
A_T^{(1)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}ip} \prod_{t=1}^m \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} \mid (\omega_1, \omega_2) \in P_t) \\
A_T^{(2)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\
&\quad \times \sum_{\{P_1, \dots, P_m\}ip} \sum_{I \in \mathcal{I}} \prod_{t_1 \in I} \prod_{t_2 \in I^c} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} \mid (\omega_1, \omega_2) \in P_{t_1}) \\
&\quad \quad \times \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} \mid (\omega_1, \omega_2) \in P_{t_2})
\end{aligned}$$

$$A_T^{(3)} = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \sum_{\{P_1, \dots, P_m\}_{ip}} \prod_{t=1}^m \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} \mid (\omega_1, \omega_2) \in P_t)$$

Por la proposición 3 de Palma and Olea (2010) se tiene que $A_T^{(1)} \rightarrow 0$, cuando $T, N, S, M \rightarrow \infty$. Para demostrar que $A_T^{(2)}$ y $A_T^{(3)}$ convergen a 0, vamos a suponer que el k ésimo cumulante de $\{\epsilon_t\}$ es finito, es decir, $|\text{cum}(\epsilon_{t_1}, \dots, \epsilon_{t_k})| < C_k$. Primero vamos a estudiar el comportamiento asintótico de $A_T^{(3)}$.

$$A_T^{(3)} = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \sum_{\{P_1, \dots, P_m\}_{ip}} \prod_{t=1}^m \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} \mid (\omega_1, \omega_2) \in P_t) \\ = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{\{P_1, \dots, P_m\}_{ip}} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \int_{\Pi^\ell} \left[\prod_{r=1}^{\ell} \phi(u_{j_r}, \lambda_r) \right] \\ \times \left[\prod_{r=1}^{\ell} h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) e^{-i\lambda_r(k_r - n_r)} \right] \\ \times \prod_{t=1}^m \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} \mid (\omega_1, \omega_2) \in P_t) d\lambda_1 \dots d\lambda_\ell$$

Para ser consistente con la notación utilizada en el apéndice B, se utilizará el mismo razonamiento que Dahlhaus (1997). Sea $P_t := \{c_1, \dots, c_k\}$ el conjunto de elementos en P_t y su versión de dimensión reducida será $\bar{P}_t = \{c_1, \dots, c_{k-1}\}$, también tenemos $\mu_{\bar{P}_t} := (\mu_{c_1}, \dots, \mu_{c_{k-1}})$ y $\mu_{c_k} = -\sum_{j=1}^{k-1} \mu_{c_j}$. Sea m el tamaño de la partición y $\mu = (\mu_{\bar{P}_1}, \dots, \mu_{\bar{P}_m})$. Denotemos a $a_r = (r, 1)$ y $b_r = (r, 2)$ en la partición indescomponible, entonces utilizando la notación se tiene

$$\begin{aligned}
A_T^{(3)} &= \frac{T^{\ell/2} \times \sigma_\epsilon^{2\ell}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{\{P_1, \dots, P_m\}ip} \sum_{j_1, \dots, j_\ell=1}^M \int_{\Pi^\ell} \left[\prod_{r=1}^{\ell} \phi(u_{j_r}, \lambda_r) \right] \\
&\times \int_{\Pi^{2\ell-m}} \left[\prod_{r=1}^{\ell} H_N \left(h \left(\frac{\cdot}{N} \right), \lambda_r - \mu_{a_r} \right) \times H_N \left(h \left(\frac{\cdot}{N} \right), -\lambda_r - \mu_{b_r} \right) \right] \\
&\times \left[\prod_{t=1}^m \frac{\kappa_{|P_t|}}{(2\pi)^{|P_t|-1}} \right] \exp \left(i \sum_{r=1}^{\ell} t_{j_r} (\mu_{a_r} + \mu_{b_r}) \right) d\mu d\lambda_1 \cdots d\lambda_\ell
\end{aligned}$$

Observe que $\phi(u, \lambda) \sim C|\lambda|^{2d(u)}$, cuando $|\lambda| \rightarrow 0$ y que $d(u) > 0$ para todo $u \in [0, 1]$.

Entonces la cantidad anterior puede ser acotada de manera natural por

$$\begin{aligned}
|A_T^{(3)}| &\leq K \frac{T^{\ell/2}}{M^\ell N^\ell} \sum_{\{P_1, \dots, P_m\}ip} \sum_{j_1, \dots, j_\ell=1}^M \int_{\Pi^\ell} \int_{\Pi^{2\ell-m}} \left[\prod_{r=1}^{\ell} L_N(\lambda_r - \mu_{a_r}) \times L_N(-\lambda_r - \mu_{b_r}) \right] \\
&\times \exp \left(i \sum_{r=1}^{\ell} t_{j_r} (\mu_{a_r} + \mu_{b_r}) \right) d\mu d\lambda_1 \cdots d\lambda_\ell \\
&\leq K \frac{T^{\ell/2}}{M^\ell N^\ell} \sum_{\{P_1, \dots, P_m\}ip} \int_{\Pi^\ell} \int_{\Pi^{2\ell-m}} \left[\prod_{r=1}^{\ell} L_N(\lambda_r - \mu_{a_r}) \times L_N(-\lambda_r - \mu_{b_r}) \right] \\
&\times \left[\prod_{r=1}^{\ell} L_M(S(\mu_{a_r} + \mu_{b_r})) \right] d\mu d\lambda_1 \cdots d\lambda_\ell
\end{aligned}$$

Utilizando el lemma A.1 (e) se tiene

$$\begin{aligned}
|A_T^{(3)}| &\leq K \frac{T^{\ell/2}}{M^\ell N^\ell} \times (\log N)^\ell \times \sum_{\{P_1, \dots, P_m\}ip} \int_{\Pi^{2\ell-m}} \left[\prod_{r=1}^{\ell} L_N(-\mu_{a_r} - \mu_{b_r}) \right] \\
&\times \left[\prod_{r=1}^{\ell} L_M(S(\mu_{a_r} + \mu_{b_r})) \right] d\mu \\
&\leq K \frac{T^{\ell/2}}{M^\ell N^\ell} \times (\log N)^\ell \times M \times \frac{N^\ell}{S^{\ell-1}} \times [\log(M) \log(S)]^{\ell-1} \rightarrow 0,
\end{aligned}$$

cuando $M, N, S, T \rightarrow \infty$.

Para finalizar, debemos desarrollar la expresión cruzada $A_T^{(2)}$,

$$A_T^{(2)} = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \sum_{\{P_1, \dots, P_m\}_{ip}} \sum_{I \in \mathcal{I}} \prod_{t_1 \in I} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} | (\omega_1, \omega_2) \in P_{t_1}) \\ \times \prod_{t_2 \in I^c} \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_{t_2}).$$

Como $\{\nu_{t,T}\}$ es un proceso Gaussiano, se tiene que $\text{cum}_k(\nu_{t,T}) = 0$ para todo $k \geq 3$, así, todos los $t_1 \in I$, se debe tener que la partición indescompensable asociada $|P_{t_1}| = 2$.

$$A_T^{(2)} = \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ \times \sum_{\{P_1, \dots, P_m\}_{ip}} \sum_{\substack{I \in \mathcal{I} \\ \forall t \in I \\ |P_t|=2}} \prod_{t_1 \in I} \text{cum}(\nu_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1, T} | (\omega_1, \omega_2) \in P_{t_1}) \\ \times \prod_{t_2 \in I^c} \text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_{t_2}),$$

la cardinalidad de la multiplicación es $1 \leq |I| \leq m - 1$ y la cardinalidad del complemento es $|I^c| = m - |I|$. La secuencia $\{\epsilon_t\}$ es independiente idénticamente distribuida, entonces el cumulante es

$$\text{cum}(\epsilon_{t_1}, \dots, \epsilon_{t_p}) = \begin{cases} \kappa_p, & \text{si } t_1 = \dots = t_p, \\ 0, & \text{en otro caso.} \end{cases}$$

Si consideramos a $P_{t_2} = \{(\omega_{1,1}, \omega_{2,1}), \dots, (\omega_{1,|P_{t_2}|}, \omega_{2,|P_{t_2}|})\}$, entonces

$$\text{cum}(\epsilon_{[u_{j_{\omega_1}} T] - N/2 + \omega_2 + 1} | (\omega_1, \omega_2) \in P_{t_2}) = \begin{cases} \kappa_{|P_{t_2}|}, & \text{si } t_{j_{\omega_1,1}} + \omega_{1,2} = \dots = t_{j_{\omega_1,|P_{t_2}|}} + \omega_{1,|P_{t_2}|}, \\ 0, & \text{en otro caso.} \end{cases}$$

denote a $\delta_{P_{t_2}}^*$ como

$$\delta_{P_{t_2}}^* = \begin{cases} 1, & \text{si } t_{j_{\omega_1,1}} + \omega_{1,2} = \dots = t_{j_{\omega_1,|P_{t_2}|}} + \omega_{1,|P_{t_2}|}, \\ 0, & \text{en otro caso.} \end{cases}$$

De esta forma

$$\begin{aligned} A_T^{(2)} &= \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \\ &\times \sum_{\{P_1, \dots, P_m\}ip} \sum_{\substack{I \in \mathcal{I} \\ \forall t \in I \\ |P_t|=2}} \prod_{t_1 \in I} \text{cum}(\nu_{[u_{j_{\omega_1} T] - N/2 + \omega_2 + 1, T]} | (\omega_1, \omega_2) \in P_{t_1}) \\ &\times \prod_{t_2 \in I^c} \kappa_{|P_{t_2}|} \delta_{P_{t_2}}^*, \end{aligned}$$

Como se puede apreciar, entre más grande sea la cardinalidad de I^c y más grande la cardinalidad de $|P_t|$ de los $t \in I^c$, más acotado será $A_T^{(2)}$. Por el argumento anterior se elige un $|I^c| = 1$, donde $I^c = \{t\}$ y la partición es de cardinalidad $|P_t| = 2$, Así se puede acotar $A_T^{(2)}$ por,

$$\begin{aligned} |A_T^{(2)}| &\leq K \frac{T^{\ell/2}}{[2\pi M H_{2,N}(0)]^\ell} \sum_{\{P_1, \dots, P_\ell\}ip} \sum_{j_1, \dots, j_\ell=1}^M \sum_{k_1, \dots, k_\ell=0}^{N-1} \sum_{n_1, \dots, n_\ell=0}^{N-1} \left| \left[\prod_{r=1}^{\ell} \widehat{\phi}(u_{j_r}, k_r - n_r) \right. \right. \\ &\times \left. \left. h\left(\frac{k_r}{N}\right) h\left(\frac{n_r}{N}\right) \right] \right| \times \left| \prod_{t_1=1}^{\ell-1} \text{cum}(\nu_{[u_{j_{\omega_1} T] - N/2 + \omega_2 + 1, T]} | (\omega_1, \omega_2) \in P_{t_1}) \right| \times \left| \delta_{P_\ell}^* \right|, \end{aligned}$$

ahora todas las particiones son de tamaño 2, usando un razonamiento similar a la proposición 2, utilizamos el hecho que $\delta_{P_\ell}^* \leq K |\text{cum}(\nu_{[u_{j_{\omega_1} T] - N/2 + \omega_2 + 1, T]} | (\omega_1, \omega_2) \in P_\ell)|$, entonces se utiliza como cota superior, el orden de convergencia de $A_T^{(3)}$, es decir

$$|A_T^{(2)}| \leq K M^{1-(\ell/2)} \left(\frac{N}{T}\right)^{(1-2d)(p-2)} \left(\frac{\sqrt{T}}{N}\right)^{(p-2)} \rightarrow 0, \quad \text{cuando } T, M, N, S \rightarrow \infty.$$

Finalmente $T^{\ell/2} \text{cum}_\ell(J_T(\phi)) \rightarrow 0$, cuando $T \rightarrow \infty$.

C.2. Demostración de Teoremas

DEMOSTRACIÓN DEL TEOREMA 3.1 Para mostrar la consistencia del estimador de Whittle, es suficiente mostrar que

$$\sup_{\theta \in \Theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \rightarrow 0,$$

en probabilidad, cuando $T \rightarrow \infty$, donde $\mathcal{L}(\theta) := \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left[\log f_{\theta}(u, \lambda) + \frac{f_{\theta_0}(u, \lambda)}{f_{\theta}(u, \lambda)} \right] d\lambda du$. Recuerde que $f_{\theta}(u, \lambda)$ está dado por la ecuación (3.9). Defina $g_{\theta}(u, \lambda) = f_{\theta}(u, \lambda)^{-1}$, que por el supuesto 2.1, $g_{\theta}(u, \lambda)$ es continua en θ , λ y u , de esta forma g_{θ} puede ser aproximada por la suma de Cesaro

$$g_{\theta}^{(L)}(u, \lambda) = \frac{1}{4\pi^2} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \widehat{g}_{\theta}(\ell, m) \exp(-i2\pi u_j \ell - i\lambda m),$$

donde $\widehat{g}_{\theta}(\ell, m) = \int_0^1 \int_{-\pi}^{\pi} g_{\theta}(u, \lambda) \exp(i2\pi u \ell + i\lambda m) du d\lambda$, note que $\sup_{\theta} |g_{\theta}(u, \lambda) - g_{\theta}^{(L)}(u, \lambda)| < \varepsilon$. Del Teorema 3.2 de Dahlhaus (1997), podemos escribir

$$\begin{aligned} & \sup_{\theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \\ & \leq \mathcal{O}\left(\frac{1}{M}\right) + \frac{\varepsilon}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) + f(u_j, \lambda)] d\lambda \\ & \quad + \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\ & \quad \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \exp(-i2\pi u_j \ell - i\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \end{aligned}$$

aquí podemos notar que $|\widehat{g}_{\theta}(\ell, m)| \leq 2\pi \sup_{(\theta, u, \lambda)} |g_{\theta}(u, \lambda)|$, sin embargo, por el supuesto 2.1 $|g_{\theta}(u, \lambda)|$ es continuo en θ , u y λ , de esta forma, y dado que el espacio paramétrico es compacto, tenemos que $|\widehat{g}_{\theta}(\ell, m)| \leq K$, para alguna constante K

positiva. Ahora, por definición, para valores fijos de $\ell, m = 1, \dots, L$, se puede elegir $\phi(u, \lambda) = \cos(2\pi u\ell) \cos(\lambda m)$ o $\phi(u, \lambda) = \sin(2\pi u\ell) \cos(\lambda m)$ en la proposición B.1 y $\phi_1(u, \lambda) = \cos(2\pi u\ell) \cos(\lambda m)$ o $\phi_1(u, \lambda) = \phi_2(u, \lambda) = \sin(2\pi u\ell) \cos(\lambda m)$ en la proposición B.2, entonces deducimos que

$$\begin{aligned} & \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\ & \quad \times \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \exp(-i2\pi u_j \ell - i\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \\ & \leq \frac{1}{16\pi^3} \sum_{\ell=-L}^L \sum_{m=-L}^L \left(1 - \frac{|\ell|}{L}\right) \left(1 - \frac{|m|}{L}\right) \sup_{\theta} |\widehat{g}_{\theta}(\ell, m)| \\ & \quad \times \left\{ \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \cos(-i2\pi u_j \ell) \cos(\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \right. \\ & \quad \left. + \left| \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} \sin(-i2\pi u_j \ell) \cos(\lambda m) \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \right| \right\} \xrightarrow{\mathbb{P}} 0 \end{aligned}$$

y demás

$$\frac{1}{M} \sum_{j=1}^M M \int_{-\pi}^{\pi} \{I_N(u_j, \lambda) - f(u_j, \lambda)\} d\lambda \rightarrow 2 \int_0^1 \int_{-\pi}^{\pi} f(u, \lambda) d\lambda du,$$

en probabilidad, cuando $M \rightarrow \infty$, de esta forma

$$\sup_{\theta \in \Theta} |\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \xrightarrow{\mathbb{P}} 0, \quad \text{cuando } T \rightarrow \infty$$

DEMOSTRACIÓN DEL TEOREMA 3.2 Sea θ_0 el verdadero valor del parámetro θ .

Si se tiene el supuesto 1, entonces el estimador de Whittle satisface el teorema del límite central

$$\sqrt{T} (\hat{\theta}_T - \theta_0) \rightarrow N[0, \Gamma(\theta_0)^{-1} + \Gamma(\theta_0)^{-1} \Lambda(\theta_0) \Gamma(\theta_0)^{-1}],$$

en distribución, cuando $T \rightarrow \infty$. Donde las matrices están definidos por

$$\begin{aligned} \Gamma(\theta_0) &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} [\nabla \log f_{\theta_0}(u, \lambda)] [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du \\ \Lambda(\theta_0) &= \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu du \end{aligned}$$

Demostración. Sea $\hat{\theta}_T$ el valor del parámetro que minimiza la función de log verosimilitud de Whittle $\mathcal{L}_T(\theta)$ dada en (3) y sea θ_0 el verdadero valor del parámetro. Por el teorema del valor medio, existe un vector $\bar{\theta}_T$ que satisface $\|\bar{\theta}_T - \theta_0\| \leq \|\hat{\theta}_T - \theta_0\|$, tal que

$$\nabla \mathcal{L}_T(\hat{\theta}_T) - \nabla \mathcal{L}_T(\theta_0) = [\nabla^2 \mathcal{L}_T(\bar{\theta}_T)] (\hat{\theta}_T - \theta_0).$$

Es suficiente mostrar que

- (a) $\nabla^2 \mathcal{L}_T(\theta_0) \rightarrow \Gamma(\theta_0)$, cuando $T \rightarrow \infty$.
- (b) $\nabla^2 \mathcal{L}_T(\bar{\theta}_T) - \nabla^2 \mathcal{L}_T(\theta_0) \rightarrow 0$, en probabilidad, cuando $T \rightarrow \infty$.
- (c) $\sqrt{T} \nabla \mathcal{L}_T(\theta_0) \rightarrow N[0, \Gamma(\theta_0) + \Lambda(\theta_0)]$, cuando $T \rightarrow \infty$.

Observe que

$$\begin{aligned}
\nabla^2 \mathcal{L}_T(\theta) &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) - f_{\theta}(u_j, \lambda)] \nabla^2 f_{\theta}(u_j, \lambda)^{-1} - \nabla f_{\theta}(u_j, \lambda) [f_{\theta}(u_j, \lambda)^{-1}]' d\lambda \\
&= \frac{1}{4\pi} \frac{1}{M} \left\{ \sum_{j=1}^M \int_{-\pi}^{\pi} \phi(u_j, \lambda) [I_N(u_j, \lambda) - f_{\theta}(u_j, \lambda)] d\lambda \right. \\
&\quad \left. + \sum_{j=1}^M \int_{-\pi}^{\pi} \nabla \log f_{\theta}(u_j, \lambda) \nabla \log f_{\theta}(u_j, \lambda)' d\lambda \right\} \\
&= \frac{1}{4\pi} [J_T(\phi) - J(\phi)] + \Gamma(\theta) + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

donde $\phi(u, \lambda) = \nabla^2 f_{\theta}(u, \lambda)^{-1}$. Aplicando la proposición 1 y la proposición 2 se obtiene la parte (a) y (b). Para la parte (c) se puede mostrar por medio de los cumulantes. Mostraremos que los cumulantes mayores o iguales a tres de $\sqrt{T} \nabla \mathcal{L}_T(\theta_0)$ converge a cero. Note que

$$\begin{aligned}
\nabla \mathcal{L}_T(\theta_0) &= \frac{1}{4\pi} \frac{1}{M} \sum_{j=1}^M \int_{-\pi}^{\pi} [I_N(u_j, \lambda) - f_{\theta_0}(u_j, \lambda)] \nabla f_{\theta_0}(u_j, \lambda)^{-1} d\lambda \\
&= \frac{1}{4\pi} J_T(\phi) - \frac{1}{4\pi} \sum_{j=1}^M \int_{-\pi}^{\pi} f_{\theta_0}(u_j, \lambda) \nabla f_{\theta_0}(u_j, \lambda)^{-1} d\lambda \\
&= \frac{1}{4\pi} [J_T(\phi) - J(\phi)] + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

donde $\phi(u, \lambda) = \nabla f_{\theta_0}(u, \lambda)^{-1}$. Por la suposición 1 y el supuesto A3, el primer cumulante de $\sqrt{T} \nabla \mathcal{L}_T(\theta_0)$ satisface

$$\begin{aligned}
\sqrt{T} \mathbb{E}(\nabla \mathcal{L}_T(\theta_0)) &= \mathcal{O}\left(\frac{\sqrt{T} \log^2 N}{N}\right) + \mathcal{O}\left(\frac{\sqrt{T}}{M}\right) \\
&\rightarrow 0,
\end{aligned}$$

cuando $T \rightarrow \infty$. Note que por lo anterior tenemos que el cumulante de segundo orden de $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)$ puede ser escrito por

$$T \mathbf{Cov}[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)] = \frac{1}{16\pi^2} T \mathbf{Cov}[J_T(\phi), J_T(\phi)].$$

Por la proposición 2 tenemos que

$$\begin{aligned} & \lim_{T \rightarrow \infty} T \mathbf{Cov}[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)] \\ &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \lambda)^{-1}]' f_{\theta_0}(u, \lambda)^2 d\lambda du \\ & \quad + \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu du \\ &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla \log f_{\theta_0}(u, \lambda) [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du \\ & \quad + \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu du \\ &= \Gamma(\theta_0) + \Lambda(\theta_0) \end{aligned}$$

donde las matrices son

$$\begin{aligned} \Gamma(\theta_0) &= \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \nabla \log f_{\theta_0}(u, \lambda) [\nabla \log f_{\theta_0}(u, \lambda)]' d\lambda du \\ \Lambda(\theta_0) &= \frac{1}{16\pi^2} \frac{\kappa_4 \sigma_\epsilon^4}{4\pi^2} \int_0^1 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \nabla f_{\theta_0}(u, \lambda)^{-1} [\nabla f_{\theta_0}(u, \mu)^{-1}]' d\lambda d\mu du \end{aligned}$$

Para finalizar, para $p > 2$, por la Proposición 3 tenemos que $T^{p/2} \mathbf{Cum}_p[\nabla\mathcal{L}_T(\theta_0)] \rightarrow 0$, cuando $T \rightarrow \infty$, así mostramos la parte (c).

Referencias

- Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723.
- Beran, J. (2009). On parameter estimation for locally stationary long-memory processes. *J. Statist. Plann. Inference*, 139(3):900–915.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31(3):307–327.
- Breidt, F. J., Crato, N., and De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of econometrics*, 83(1):325–348.
- Brillinger, D. (1981). *Time Series: Data Analysis and Theory*. Holden Day, San Francisco. Society for Industrial and Applied Mathematics.
- Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods: Theory and Methods*. Springer series in statistics. Springer-Verlag.
- Dahlhaus, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.*, 62(1):139–168.
- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.*, 25(1):1–37.

- Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.*, 28(6):1762–1794.
- Dahlhaus, R. and Giraitis, L. (1998). On the optimal segment length for parameter estimates for locally stationary time series. *Journal of Time Series Analysis*, 19(6):629–655.
- Dahlhaus, R. and Polonik, W. (2006). Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary processes. *Ann. Statist.*, 34(6):2790–2824.
- Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli*, 15(1):1–39.
- Dahlhaus, R., Rao, S. S., et al. (2006). Statistical inference for time-varying arch processes. *The Annals of Statistics*, 34(3):1075–1114.
- Deo, R. S. and Hurvich, C. M. (2001). On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory*, 17(04):686–710.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007.
- Fernández, C. and Steel, M. F. (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371.
- Fryzlewicz, P. (2005). Modelling and forecasting financial log-returns as locally stationary wavelet processes. *J. Appl. Stat.*, 32.
- Fryzlewicz, P., Sapatinas, T., and Subba Rao, S. (2006). A Haar-Fisz technique for locally stationary volatility estimation. *Biometrika*, 93(3):687–704.

- Fryzlewicz, P. and Subba Rao, S. (2011). Mixing properties of ARCH and time-varying ARCH processes. *Bernoulli*, 17(1):320–346.
- Fryzlewicz, P., Van Bellegem, S., and Von Sachs, R. (2003). Forecasting non-stationary time series by wavelet process modelling. *Annals of the Institute of Statistical Mathematics*, 55(4):737–764.
- Harvey, A., Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.
- Hurvich, C. M. and Ray, B. K. (2003). The local whittle estimator of long-memory stochastic volatility. *Journal of Financial Econometrics*, 1(3):445–470.
- Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The Review of Economic Studies*, 65(3):361–393.
- Martin, W. and Flandrin, P. (1985). Wigner-ville spectral analysis of nonstationary processes. *IEEE Transactions on Acoustics, Speech, and Signal Processing*.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.
- Palma, W. and Olea, R. (2010). An efficient estimator for locally stationary Gaussian long-memory processes. *Ann. Statist.*, 38(5):2958–2997.
- Roueff and von Sachs, R. (2011). Locally stationary long memory estimation. *Stochastic Processes and their Applications*, 121(4):813 – 844.
- Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of econometrics*, 63(1):289–306.
- Taylor, S. (1980). Conjectured models for trends in financial prices, tests and forecasts. *Journal of the Royal Statistical Society. Series A (General)*, pages 338–362.

Taylor, S. (1986). *Modelling Financial Time Series*. John Wiley & Sons.

Van Bellegem, S. (2012). *Locally stationary volatility models*, in *Bauwens, L. and Hafner, C.M. and Laurent, S.* Wiley Handbooks in Financial Engineering and Econometrics: Handbook of Volatility Models and Their Applications. Wiley.

von Sachs, R. and MacGibbon, B. (2000). Non-parametric curve estimation by wavelet thresholding with locally stationary errors. *Scand. J. Statist.*, 27(3):475–499.

Zaffaroni, P. (2009). Whittle estimation of egarch and other exponential volatility models. *Journal of econometrics*, 151(2):190–200.