ASYMPTOTICS FOR FRACTIONAL PROCESSES

JAMES DAVIDSON

Version of 23 May 2024

Contents

Preface

The object of this book is to develop an approach to the large-sample analysis of the so-called fractional partial-sum processes, featuring long memory increments. Long memory, equivalently called strong dependence, is commonly defined to mean that the autocovariance sequence is non-summable.

These increment processes have a linear moving average representation with a single parameter measuring the degree of long-run persistence. In the econometrics literature this parameter is denoted by d although in statistics the symbol H, to denote $d + \frac{1}{2}$, is well established. Long memory means that d is positive, while negative d defines a special type of short memory known as antipersistence, for which the autocovariance sequence sums to zero. Antipersistent processes are treated here in parallel with the long memory case.

By contrast to the time domain focus adopted in this book, an extensive literature on long memory takes the route of harmonic analysis. The spectral density depends on d in a characteristic manner, providing the basis for various semiparametric estimators. Time domain analyses do not generally entertain the possibility of a nonparametric representation in which the rate of decay of the autocovariances is not linked to a functional form, although as explained in Chapter 1, such a gain in generality would be more apparent than real. It is the elegant (albeit misnamed) concept of the 'fractional root', as the natural generalization of the unit root, that has captured the imagination of researchers.

Chapters 1, 2 and 3 treat the weak convergence of certain normalized partial sums to fractional Brownian motion, otherwise known as fBM. This is a Gaussian, almost surely continuous process that, unlike regular Brownian motion, exhibits correlated increments. The proof of Gaussianity uses well established methods after rearranging the partial sum so as to isolate the coordinates of the driving shock process. The approach is then similar to the conventional analysis of unit root processes converging to regular Brownian motion, except that the increments are heteroscedastic. To show uniform tightness of the distributions also follows a conventional path up to a point, but as well as showing that the squared process increments are uniformly integrable it is necessary to specify the rate of this convergence to zero. The requirement entails a minimum moment condition which binds in the antipersistent case.

Two versions of these and subsequent results are given, assuming first that the driving shocks are independent and identically distributed and then that they possess a nonparametric form of weak dependence (near-epoch dependence on a mixing process). No apology is offered for this dual approach. Dealing with short-range dependence adds a substantial layer of complication, making it harder for the reader to grasp the essential features of the strong dependence model. Since no restrictions are placed on the models studied other than linearity and the behaviour of the moving average coefficients at long range, even with i.i.d. shocks the theory embraces a very extensive class of time series processes. It is really only when nonlinear features such as conditional heteroscedasticity have to be confronted that these results would prove strictly inadequate. While not overlooking the reduced rate of convergence to the limit (effective sample size) in such cases, it is gratifying to find that in large samples the replacement of the shock variance by its long-run counterpart is in nearly every case the only consequence of the generalization.

Chapters 4, 5 and 6 treat different aspects of the most technically challenging problem, which is the limiting distribution of stochastic integrals where both the integrand and the integrator processes exhibit either long memory or antipersistence. These are not Itô integrals, so the approach has to be quite different from the by-now conventional analysis of unit root processes. The key step is to show that the limit distributions feature the sum of distinct Itô-distributed components. Chapter 7 reviews applications of the theory to regression with fractional processes, in particular cointegrating regressions involving nonstationary (partial sum) processes.

All these results are also developed initially under the assumption of i.i.d. shocks. The various extensions required to cover the dependent shocks case are gathered together in Chapter 8. This arrangement allows the reader to decide at any point whether the additional effort of dealing with the general case is worthwhile.

The final two chapters are included to provide context, with accounts of some closely related results. Chapter 9 sets out the essentials of the harmonizable representation of fractional processes. The main revelation is that asymptotic results obtained in this framework match those already found in the time domain analysis, showing how the same limit properties can be demonstrated in apparently very different ways. Chapter 10 is about local-to-unity autoregression, cases where the limit distribution is an Ornstein-Uhlenbeck process instead of fractional Brownian motion. In neither case does the treatment claim to be comprehensive, the chief intention being to inform readers curious to compare these alternative methodologies and alternative generalizations of the unit root class of nonstationary processes.

There are two appendices. Appendix A contains some items of probability theory not directly related to fractional models but important for deriving the limit distributions. Appendix B lists various trigonometric and related identities arising in the text for the benefit of those desiring an aide memoir, together with a number of integrals to be encountered whose solutions are not elementary. These latter are to be found in one or other of the legendary formula compendia of Gradsteyn and Ryzhik ([27]) and Abramowitz and Stegun ([1]), but the formulae are reproduced here for the convenience of readers.

These chapters were originally conceived as forming a section of *Stochastic Limit Theory* 2nd Edition ([14]), henceforth abbreviated as SLT. In the event a free-standing contribution proves to be the better option, but SLT plays an important role as a source for certain necessary results that can be cited rather than reproved, as well a useful reference for background material. Readers are referred there for certain theorems arising in Chapters 3 and 6, in particular. Citations of other relevant proofs are supplied in footnotes at various points.

A note about format and presentation. Numbered assumptions, theorems, and lemmas are terminated by $\square$ unless followed immediately by a proof, while proofs are terminated by $\blacksquare$. References to these items appear in the text in boldface while references to chapter sections have the prefix §. Certain frequently used notations are relatively non-standard and so should be defined. Thus, in a relation involving sequences x_n and y_n, the notation $x_n \ll y_n$ replaces the more commonly used $x_n \leq Cy_n$ for all n, for a positive constant C. The notation $x_n \simeq y_n$ indicates that there are constants $C_1 > 0$ and $C_2 < \infty$ such that $C_1 \leq x_n/y_n \leq C_2$ for all n, generalizing the standard tilde notation $\sim$ denoting that $x_n/y_n \to 1$. Braces are most commonly used to denote the sequence or array whose generic element is enclosed, as in $\{x_i\}$, but they are also used to define a random event. The indicator of an event A is written 1_A. $[x]$ is the floor function of x, the largest integer below x.

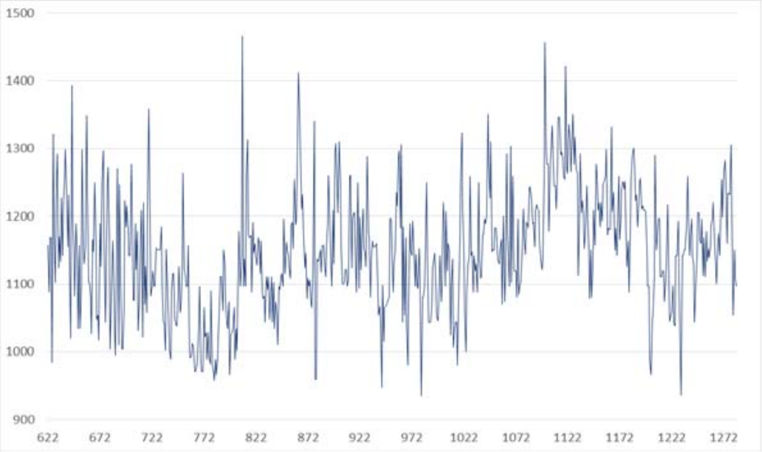
Some of the results presented here are revisions, extensions and corrections of material from published journal articles by myself and coauthors Robert de Jong and Nigar Hashimzade. The contributions of these colleagues to the enterprise, as well as that of Morten Nielsen who most helpfully commented on an early draft, are most gratefully acknowledged. I have also received some extremely valuable comments from anonymous referees.

James Davidson
University of Exeter

May 2024

Chapter 1

The Fractional Model



There is still no better illustration of the long memory phenomenon than the first example to be identified in the literature. The plot above shows the annual minima of depth measurements of the river Nile, taken at the Roda gauge near Cairo, for the 663 years from 622 AD to 1284 AD, the longest unbroken run available, as recorded in Omar Tousson's history of the Nile ([70]). Famously, these data were studied by the hydrologist Harold Edwin Hurst (1880-1978) who was engaged in the design of the Aswan high dam. The problem was to specify a dam high enough (but no higher) such that the lake to be created behind it would neither empty nor overflow under variations in rainfall at the river's source, given a constant rate of discharge downstream. This question was easily answered if it could be assumed that the measured flow generated an independent, or at worst weakly dependent, sequence over time. The distribution of the range of a centred and normalized Brownian motion over a fixed period is easily calculated, and would provide a good approximation if the period were long enough. What Hurst showed decisively by means of his well-known rescaled range (R/S) test (see

[38], [47], [67]) was that this distribution was the wrong one.

Two features stand out in the plot of what must be thought of as an annual record of precipitation in central Africa. The first is the complete absence of trend over this lengthy historical period. This is, from any reasonable perspective, a stationary time series. But at the same time the persistence of the variations is very striking. The series median is 1148 centimetres, but in the 86 years 720–806, this level was exceeded only twice. At the other extreme, in the 33 years 1099–1132 the level did not once fall below the median. The Egyptians living through these prolonged periods of drought and flood, respectively, may have thought with some reason that "the climate was changing". The full record says otherwise.

It is to build models that can describe and predict series of this type that the theory of fractionally integrated processes has been developed. Early contributions to this literature include the classic papers of Clive Granger ([29], [28]), Jonathan Hosking ([33], [34]) and Peter Robinson ([59]). A useful general account is the monograph by Jan Beran ([4]) and prominent among more recent contributions to the statistical analysis are the comprehensive monographs by Giraitis et al. ([25]) Pipiras and Taqqu ([52]) and Beran et al. ([5]), the latter featuring an extensive bibliography. Among other valuable references are the books by Palma ([50]), Hassler ([31]), Samorodnitsky ([63]) and Doukhan et al. ([22]).

1.1 The Model

Working with the paradigm of the linear moving average process driven by a stationary shock sequence, say $\{u_i\}_{i=-\infty}^{\infty}$ having mean of zero, the basic setup gives the form of the process at date i as

$$x_i = \sum_{j=0}^{\infty} b_j u_{i-j} \tag{1.1}$$

where $b_0 = 1$. The structure defining a fractional (long memory) process can be expressed in various ways, but for maximum generality let

$$b_j \sim dj^{d-1} L(j) \tag{1.2}$$

as $j \to \infty$, where $d > 0$ is a parameter and the sequence $\{L(j)\}_{j=0}^{\infty}$ is positive and at worst slowly varying at infinity. Observe that b_j is allowed to take arbitrary form when j is finite since (1.2) specifies only the limiting behaviour of the sequence at long range. In leading cases $L(j)$ is, or converges to, a positive constant. Since this is a one-sided (causal) model, relating the present to the past, it may in certain contexts be necessary to specify that $b_j = 0$ for $j < 0$ in a two-sided sum.

The distribution of the sequence $\{x_i\}_{i=-\infty}^{\infty}$ is not well defined if the value of d is too large and in the next section it is explained how $d < \frac{1}{2}$ generally defines the boundary for existence in the mean-squared sense, depending on the specification of $\{u_i\}_{i=-\infty}^{\infty}$. If $u_i = 0$ for $i \leq 0$ then (1.1) for any $d > 0$ defines a nonstationary process with finite starting point, notably including the well-known unit root process when $d = 1$.

Subject to these considerations, nothing needs to be added to (1.1)+(1.2) to define what is often called a fractionally integrated process, although strictly this terminology is reserved for the special case to be examined in §1.3. Maximum likelihood estimation of a parametric model conforming to this setup, applied to the Nile minima series, give values of d in the vicinity of 0.4 (see e.g. [18]). While the multiplier d in formula (1.2) is optional since it might be incorporated into L with no loss of generality, it will be convenient at various points in the development to have it appear explicitly. Thus, formula (1.2) applies to the case $d = 0$ since it yields simply $x_i = u_i$. Any dependence in the process then has to be modelled by a different mechanism.

A key feature of model (1.1)+(1.2) with $d > 0$ is that the sequence $\{b_j\}_{j=0}^{\infty}$ is nonsummable. It is this condition that technically defines the categorization of the process $\{x_i\}$ as long memory and, equivalently, to be strongly dependent. The assumption $\mathrm{E}(u_i) = 0$ is therefore essential, since with a nonzero mean included the sum in (1.1) would diverge. A fractional process with nonzero mean μ can exist, but it must take the form $x_i + \mu$ for $i \geq 1$ where x_i follows (1.1)+(1.2).

Another leading implication of strong dependence is the failure of the conventional central limit theorem. Thus,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i = \sum_{j=0}^{\infty} b_j \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_{i-j} \right) = \sum_{j=0}^{\infty} b_j Z_n(j) \tag{1.3}$$

(say) where under various conditions on the shock sequence $\{u_i\}$, such as serial independence, the variables $Z_n(j)$ would approach the same Gaussian limit for each j, as $n \to \infty$. In the case $|\sum_{j=0}^{\infty} b_j| < \infty$, this Gaussian property would extend to $n^{-1/2}\sum_{i=1}^{n} x_i$ itself, the only effect being a change of scale. However, under long memory, $|\sum_{j=0}^{\infty} b_j| = \infty$ and nothing useful can be said about the limiting distribution of (1.3). Specifically, the normalization by $n^{-1/2}$ is inappropriate and the variance of the sum diverges.

The permitted inclusion of slowly varying components in (1.2) involves something of a sleight of hand, because to do the asymptotics requires that the L functions for the series have known form. These functions must be included in the normalizing divisors to obtain the known limit distributions of statistics. If neglected, a slowly varying drift would render the limit distributions invalid, although hopefully by a small enough amount that finite-n approximations would be useful. Since such components are rarely if ever specified in empirical applications, there is a sense in which they are just window-dressing and the mathematics would surely be simplified by their omission. However, on the other side there is the question "What if?" If slowly varying components do exist, at least the mathematics tells the practitioner what to do about them. There is the larger issue, that the parameter d must also be known to make use of the limit results. This is not a problem arising in the conventional unit root analysis, but at least there do exist a range of procedures to estimate d from the data. Estimation falls outside the subject matter covered in this book, but among innumerable references see for example [24], [48], [60], [66], [32], and [18].

1.2 Shocks and Dependence

Regarding the distribution of the shock sequence $\{u_i\}$, two alternative sets of assumptions are invoked at different points in the development that follows.

1.1 Assumption The sequence $\{u_i, -\infty < i < \infty\}$ is identically and independently distributed and L_r-bounded for $r \geq 2$, with $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(u_i^2) = \sigma_u^2$ where $0 < \sigma_u^2 < \infty$.

1.2 Assumption The sequence $\{u_i, -\infty < i < \infty\}$ is

(a) strictly stationary and L_r-bounded for $r \geq 2$, with $\mathrm{E}(u_i) = 0$, $\mathrm{E}(u_i^2) = \sigma_u^2$ where $0 < \sigma_u^2 < \infty$ and for $k > 0$, $\mathrm{E}(u_i u_{i+k}) = \gamma_u(k) = O(k^{-1-\delta})$ for $\delta > 0$.

(b) L_2-near epoch dependent (NED)[1] of size $-\frac{1}{2}$ on either a α-mixing sequence of size $-r/(r-2)$ with $r > 2$, or a ϕ-mixing sequence of size $-r/(2r-2)$.

Assumption **1.2**(a), and in particular the requirement $\delta > 0$, specifies that the autocovariances are absolutely summable, the condition defining the boundary between weak and strong dependence. It implies that the standard deviations of partial sums of the process grow like $\sqrt{n}$ and hence that $\omega_u^2 < \infty$ where

$$\omega_u^2 = \lim_{n\to\infty} \frac{1}{n}\mathrm{E}\left(\sum_{j=1}^{n} u_i\right)^2 = \sigma_u^2 + 2\sum_{k=1}^{\infty} \gamma_u(k). \tag{1.4}$$

Assumption **1.2**(b) implies the autocovariance summability condition of **1.2**(a),[2] but this is nonetheless stated explicitly to define notation and focus attention on the important implications of weak dependence.

In calculations involving second moments, much simplicity in exposition is achieved by neglecting autocovariances, which Assumption **1.1** permits. Since no restriction beyond (1.2) is imposed on the moving average coefficients, allowing them to take arbitrary forms for finite lags, assuming independent shocks does no more than confine attention to linear forms of serial dependence, less restrictive than might be supposed. The strategy adopted here, when Assumption **1.1** is convenient for exposition, is to collect the generalizations needed to move to Assumption **1.2** in one place. What will be shown, in Section 2.6 and Chapter 8 in particular, is that relaxing the i.i.d. assumption in many cases involves no change in asymptotic results other than replacement of σ_u^2 by ω_u^2 in formulae.

Another class of models with the moving average representation (1.1) is the familiar ARMA class, where the expression corresponding to (1.2) is $b_j \backsimeq \alpha_{\max}^j$ where $\alpha_{\max}$ is the largest autoregressive root, assuming the parameterization that places stable roots inside the unit circle. If $|\alpha_{\max}| < 1$, ARMA coefficients of order

[1] For an explanation of the mixing and NED dependence concepts see SLT, Chapters 15 and 18.

[2] See SLT Chapter 17.5.

$\alpha_{\max}^j$ as $j \to \infty$ form an absolutely summable and hence also square-summable geometric sequence. The ARMA process is in this case weakly dependent and covariance stationary, having a finite variance not dependent on the time index. That Assumption **1.2** is a relatively mild condition on dependence is evident from the fact that geometric memory decay corresponds at long range to the case where δ is arbitrarily large.

The coefficients in (1.2) are square summable if $0 < d < \frac{1}{2}$, which is a sufficient condition for covariance stationarity. The cited estimates of d for the Nile minima series fall in the covariance stationary region although not too far from its boundary, which is what the eyeball analysis of the series would also suggest. Under Assumption **1.1** the variance of the process $\{x_i\}$ in the stationary case is easily calculated as

$$\sigma_x^2 = \mathrm{E}(x_i^2) = \sigma_u^2 \sum_{j=0}^{\infty} b_j^2 < \infty. \tag{1.5}$$

The autocovariances are also defined in this case. Since this symbol arises frequently, for economy of notation define γ_k to stand in for $\gamma_x(k) = \mathrm{E}(x_i x_{i+k})$, for $k > 0$. It is easily verified that these have the property

$$\gamma_k = \sigma_u^2 \sum_{j=0}^{\infty} b_j b_{j+k} \backsimeq \sum_{j=k+1}^{\infty} j^{2d-2}\left(1 - \frac{k}{j}\right)^{d-1} L(j)^2 = O(k^{2d-1\cdot}). \tag{1.6}$$

The term weakly dependent is generally defined to mean a process with summable autocovariances, the ARMA class being the best-known example, whereas under stationary strong dependence the autocovariances do not form a summable sequence, as is apparent from (1.6).

Even if stationary, long memory processes do not generally satisfy useful mixing and near-epoch dependence (NED) conditions.[3] Under Assumption **1.1**, the L_2-NED criterion for x_i can be evaluated as

$$\|x_i - \mathrm{E}_{i-m}^{i+m} x_i\|_2 = \sigma_u \left(\sum_{j=m+1}^{\infty} b_j^2 \right)^{1/2} = O(m^{d-1/2}).$$

This does vanish with $d < \frac{1}{2}$ although not fast enough to meet the usual conditions for the central limit theorem to operate. In the L_p-bounded case with $p < 2$, and also with u_i dependent under Assumption **1.2**, where in either case it is necessary to appeal to the Minkowski inequality, no NED property of any kind can be demonstrated for x_i. The usual nonparametric techniques for modelling dependence are not available here.

The nonstationary fractional process with $\frac{1}{2} \le d < 1$, in spite of exhibiting too much persistence in its variations to possess a variance, does eventually lose dependence on initial conditions since $b_j \to 0$ as $j \to \infty$. Its behaviour is therefore distinct from that of a random walk. On the other hand the case $d = 1$ matches the

[3] Consider the sufficient mixing condition under linearity of SLT Theorem 15.9. Under (1.2), there is no positive value of d to satisfy condition (b) of that theorem.

ARMA case $\alpha_{\max} = 1$, the so-called unit root process. A major reason for interest in the fractional process is that it provides an alternative model class within which the unit root can be embedded. However, d is not the root of any equation and the 'fractional root' terminology sometimes encountered in the literature is best avoided.

The single model feature intrinsically associated with long memory is the parameter d and according to the equality in (1.6), under Assumption **1.1** the autocovariance sequence $\{\gamma_k\}_{k=0}^{\infty}$ provides a complete specification of the memory properties of x_i in (1.1). It is reasonable to ask whether the long memory property might be specified directly by the autocovariances or, more conveniently, in terms of the spectral density (the Fourier transform of the autocovariances). Some details of this approach are given in Chapter 9. Working with the spectrum, dependence on d can be expressed semiparametrically. Might there be a gain in generality to be achieved by avoiding the linear formulation in (1.1) altogether?

The answer to this question is best provided by consideration of the Wold decomposition theorem[4] according to which a stationary and nondeterministic time series always has a linear moving average representation with *white noise* shocks, that is, covariance stationary and uncorrelated with zero mean. Except in the Gaussian case this is not the same thing as i.i.d., as specified in Assumption **1.1**. However, what it does imply is that the only form of shock dependence permitted must coexist with uncorrelatedness and hence must be nonlinear, prominently including the case of conditional heteroscedasticity. Volatility models of the ARCH and GARCH type (see for example SLT §18.6) are permitted under Assumption **1.2**. In other words, defining long memory as a property of the second moments of the joint distribution of a process is equivalent to specifying (1.1) with (1.2) under Assumption **1.1** and nearly so under Assumption **1.2**, in the sense that the cases ruled out under Assumption **1.1** are of a specific character, that is to say, they must feature nonlinear dependence. Arguably, the loss of generality involved in working in the linear framework is little more than notional.

It is also noteworthy that while long memory in volatility is an important phenomenon, the moving average framework is also adopted in fractional generalizations of GARCH models.[5] The centred squares of a process are modelled as fractional linear processes, even though with random signs the process itself could be white noise. If the centred squares are covariance stationary, the same basic considerations apply.

The focus on the stationary case of long memory adopted henceforth is not because the nonstationary case is uninteresting, but because it is connected to the stationary case by the simple operations of cumulation and differencing. If x_i is stationary with parameter $d < \frac{1}{2}$ then $S_i = \sum_{j=1}^{i} x_j$ is a nonstationary fractional process, having $1 + d$ as the fractional exponent. The integration operation can be repeated any finite number of times, to define d to have any desired positive value. The original series is easily retrieved from S_i by differencing, with $x_i = \Delta S_i =$

[4] SLT Theorem 13.14.

[5] For example, see [16].

$S_i - S_{i-1}$ after setting $S_0 = 0$. Any other starting value S_0 can be substituted for 0 when forming the integral, although the initial date must always be finite. It is legitimate to define a stationary long memory process $\{x_i, -\infty < i < \infty\}$, but take care to note that $\{S_i, -\infty < i < \infty\}$ does not have a well-defined distribution. It is perhaps needless to say that it is the behaviour of partial sum processes $\{S_i, i \geq 1\}$, after suitable normalization, that is the focus of interest in this book. Their properties derive from those of the increment processes just as the behaviour of Brownian motion is understood as deriving from the partial summation of weakly dependent increments.

1.3 Fractional Integration

The best-known mechanism for inducing the behaviour summarized in (1.2) is represented by the recursion

$$b_j = \frac{j+d-1}{j} b_{j-1},\ j > 0 \tag{1.7}$$

where $b_0 = 1$. (No slowly varying component in this instance). These b_j are the coefficients of the binomial expansion of the so-called fractional integral

$$(1-B)^{-d} = \sum_{j=0}^{\infty} b_j B^j. \tag{1.8}$$

where B denotes the backshift operator (or lag operator) defined by the relation $Bx_i = x_{i-1}$, where x_i is a time series coordinate as in (1.1).[6] The fractional integral generalizes the unit root operator $(1-B)^{-1}$. Its solution as an infinite-order lag polynomial follows from Newton's generalized binomial theorem, which says that for $|t| < 1$,

$$(1+t)^r = \sum_{j=0}^{\infty} \frac{(r)_j}{j!} t^j \tag{1.9}$$

where r is any real number and

$$(r)_j = r(r-1)\cdots(r-j+1). \tag{1.10}$$

The latter function is known as a falling factorial and $(\cdot)_k$ is called the Pochhammer symbol.

The relation $j! = \Gamma(j+1)$ where Γ denotes the gamma function, defined in (B.12) of Appendix B, follows by recursive application of (B.13) starting from $\Gamma(1) = 1$. While $(r)_j = r!/(r-j)!$ is only true when r is a positive integer and $r \geq j$, (B.13) shows that the equality

$$(r)_j = \frac{\Gamma(r+1)}{\Gamma(r-j+1)} \tag{1.11}$$

---

[6] While in econometrics L is the symbol commonly assigned for the lag operator, B is adopted here to avoid confusion with the slow variation process $L(j)$.

holds for any real r. Define

$$b_j = \frac{(d+j-1)_j}{j!} = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} = \frac{d\Gamma(j+d)}{\Gamma(d+1)\Gamma(j+1)} \tag{1.12}$$

and notice that the recursion in (1.7) generates this series for $j \geq 1$. In view of the fact that

$$(-r)_j = (-1)^j (r+j-1)_j$$

it can be verified from the first equality of (1.12) that

$$\frac{(-d)_j}{j!} = b_j(-1)^j.$$

Setting $r = -d$ and replacing t by $-B$, the expansion in (1.9) assumes the form of (1.8). It is clear why with $d > 0$ the process generated by (1.8) is referred to as 'fractionally integrated', the 'fully integrated' case with $d = 1$ corresponding to the unit root. It follows from (1.10) and (1.11) for fixed j and increasing r that for a fixed value of a,

$$\frac{\Gamma(r)r^a}{\Gamma(r+a)} \to 1 \text{ as } r \to \infty. \tag{1.13}$$

Therefore (1.12) implies

$$b_j \sim \frac{dj^{d-1}}{\Gamma(d+1)} \tag{1.14}$$

which may be compared with (1.2). The class of models having lag structure (1.8) with $d \in (0, \frac{3}{2})$ provide a continuum of dependence properties within which the unit root case is embedded, with $0 < d < \frac{1}{2}$ representing the covariance stationary members of the class.

One advantage of special case (1.12) is that the autocovariances can be evaluated exactly. The following derivation adapts Lemma 1 of [35].

1.3 Theorem

$$\gamma_k = \sigma_u^2 \frac{\Gamma(1-2d)\Gamma(d+k)}{\Gamma(1-d+k)} \frac{\sin \pi d}{\pi}. \tag{1.15}$$

Proof Substituting from (1.12) into the formula in (1.6) gives

$$\begin{aligned}
\gamma_k &= \sigma_u^2 \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \frac{\Gamma(j+k+d)}{\Gamma(d)\Gamma(j+k+1)} \\
&= \sigma_u^2 \frac{\Gamma(d+k)}{\Gamma(d)\Gamma(k+1)} F(d, k+d; k+1; 1) \\
&= \sigma_u^2 \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)} \frac{\Gamma(d+k)}{\Gamma(1-d+k)}
\end{aligned} \tag{1.16}$$

where the second equality of (1.16) notes the match with the hypergeometric series defined in (B.22) and the third is got by simplification after applying identity (B.23). The formula in (1.15) is lastly obtained by use of (B.15). ∎

Applying (1.13) shows that $\gamma_k = O(k^{2d-1})$, agreeing with the calculation in (1.6). The case of particular interest, easily shown using identity (B.15) after setting $k = 0$, is

$$\gamma_0 = \sigma_x^2 = \sigma_u^2 \frac{\Gamma(1-2d)}{\Gamma(1-d)^2}. \tag{1.17}$$

As expected, this formula diverges as d approaches the stationarity boundary of $\frac{1}{2}$. Another version of formula (1.15) sometimes cited is

$$\gamma_k = \sigma_u^2 \frac{\Gamma(1-2d)(-1)^k}{\Gamma(1-d+k)\Gamma(1-d-k)}.$$

Switching between the two is a simple application of identities (B.15) and (B.5).

1.4 Antipersistence

The case $d < 0$ in formula (1.2) is not by itself of special interest since it merely implies summable lag coefficients and hence, weak dependence in general. This is however the distinguishing feature of the fractionally differenced process, otherwise known as antipersistent. The special additional requirement defining the antipersistent case is that $\sum_{j=0}^{\infty} b_j = 0$. According to the first equality in (1.6), under Assumption **1.1** this has the implication (noting $\gamma_k = \gamma_{-k}$ under stationarity) that

$$\sum_{k=-\infty}^{\infty} \gamma_k = \sigma_u^2 \left(\sum_{j=0}^{\infty} b_j \right)^2 = 0. \tag{1.18}$$

With $\gamma_0 = \sigma_x^2 > 0$ from (1.5), it follows that antipersistent processes must exhibit negative autocorrelation.

The most familiar case of antipersistence is the simple difference of an i.i.d. process, $x_i = u_i - u_{i-1}$. Here, the autocovariance sequence is of the form $\gamma_0 = 2\sigma_u^2$, $\gamma_1 = \gamma_{-1} = -\sigma_u^2$ and zero otherwise, which agrees with (1.18). This is the case $d = -1$. More generally, a fractional process with d in the range $(-\frac{1}{2}, 0)$ arises as the differences of a covariance-nonstationary fractional process where the fractional exponent has the form $d + 1 \in (\frac{1}{2}, 1)$. Suppose that $y_i = \sum_{j=0}^{\infty} a_j u_{i-j}$ where $a_j \sim j^d L(j)$ with d in the indicated range and $L(j) > 0$, and so define $x_i = y_i - y_{i-1} = \sum_{j=0}^{\infty} b_j u_{i-j}$ where $b_0 = a_0 = 1$ and $b_j = a_j - a_{j-1}$ for $j > 0$. Given the assumption of slow variation with $L(j)/L(j-1) \to 1$, it can be verified that as $j \to \infty$,

$$j^d L(j) - (j-1)^d L(j-1) \sim d j^{d-1} L(j) \tag{1.19}$$

which matches the representation in (1.2). It is also easy to see that with $d < 0$,

$$\sum_{j=0}^{m} b_j = a_m \sim m^d L(m) \to 0 \tag{1.20}$$

as $m \to \infty$. In the case of the fractional integral defined by (1.12), the relation in (1.19) holds not just as a tendency but identically for every $j > 0$. Applying

(B.13) once again, note that

$$\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} = \frac{\Gamma(j+d)}{\Gamma(d+1)\Gamma(j)}\left(\frac{j+d}{j} - 1\right)$$
$$= \frac{\Gamma(j+d+1)}{\Gamma(d+1)\Gamma(j+1)} - \frac{\Gamma((j-1)+d+1)}{\Gamma(d+1)\Gamma((j-1)+1)}.$$

The fact that these coefficients sum to zero when $d < 0$ follows directly on setting $B = 1$ in (1.8). Also note that (1.15) gives $\gamma_k < 0$ for every $k > 0$ when $d \in (-\frac{1}{2}, 0)$. The fact that the γ_k sum to zero over the range $k = -\infty, \ldots, \infty$ is less obvious from the formula, but this must follow from (1.18) and (1.20).

The differencing operation can be iterated just like the integration operation so in principle, fractional processes are defined for any real value of d at all. Since

$$1 + (b_1 - 1) + (b_2 - b_1) + (b_3 - b_2) + \cdots = 0$$

for any sequence $\{b_j\}_{j=1}^{\infty}$ converging to zero, the antipersistence property continues to hold under iterated differencing.

By contrast with integration, differencing loses information about the current location of the process and a starting value has to be supplied for each integration operation. If $x_i = S_i - S_{i-1}$ for $i = 1, 2, \ldots$ are the differences of a nonstationary process $\{S_i,\, i = 1, 2, \ldots\}$, it is always correct to set $x_1 = S_1$ since the differencing is in effect undoing a cumulation operation that must, by construction, have had a finite initial point. This is true whether or not $\{x_i\}$ is itself stationary. By contrast, since no finite initial date can be specified, the operation of differencing a stationary sequence loses the initial value. For this reason, it is said to form an overdifferenced process. In particular, an antipersistent process with $d < -\frac{1}{2}$ is overdifferenced since cumulating it from any initial date yields a fractional process with exponent $1 + d < \frac{1}{2}$, hence eventually stationary.

A convenient vehicle for understanding the antipersistent case is the fractional integral $x_i = (1 - B)^{-d}u_i$ from (1.8), since inverting the polynomial is straightforward. The autoregressive representation of the process is $(1 - B)^d x_i = u_i$, or equivalently

$$x_i = \sum_{j=1}^{\infty} \phi_j x_{i-j} + u_i \tag{1.21}$$

where the lag coefficients are

$$\phi_j = -\frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} = O(j^{-d-1}).$$

If $d \le -1$, this latter sequence is either non-vanishing or divergent as $j \to \infty$, which implies that the process $\{x_i\}_{i=-\infty}^{\infty}$ does not have a well-defined distribution. Such a process is said to be non-invertible. On the other hand, processes for which $-1 < d \le -\frac{1}{2}$ are invertible, noting that $\phi_j \to 0$ as $j \to \infty$ and the representation in (1.21) is well defined. (See [49] and [8] for alternative analyses of this question). In spite of being invertible, these cases are excluded from consideration in the sequel since it is the bounds $-\frac{1}{2} < d < \frac{1}{2}$ that fix the region over which normalized partial sum processes have the limiting properties of interest.

1.5 Linearity

The pure fractional model (1.8) depending on a single parameter is excessively restrictive. If the first route to a parametric generalization is to let $\{u_i\}$ be a weakly dependent process, maintaining linearity should be the first consideration. In other words, let a moving average representation of the shocks be $u_i = \varphi(B)\varepsilon_i$ where $\{\varepsilon_i\}$ is i.i.d. with zero mean and finite variance and $\varphi(B) = \sum_{j=0}^{\infty} \varphi_j B^j$ denotes a polynomial in the lag operator with absolutely summable coefficients. The natural choice is the ARMA(p, q), for which $\varphi(B) = \theta(B)/\phi(B)$ and $\phi(B)$ and $\theta(B)$ are lag polynomials, of finite orders p and q respectively, having roots strictly outside the unit circle. The rate of decay of the lag coefficients is geometric in this case. The ARFIMA(p, d, q) (autoregressive fractionally integrated moving average) model with representation

$$\phi(B)(1-B)^d x_i = \theta(B)u_i$$

is a popular choice in empirical studies of long memory.

The following result shows that linear weak dependence of the forcing sequence has no effect on the long-range behaviour of the process beyond changes of scale. For simplicity's sake the slowly varying component is omitted in this instance, but its inclusion would change nothing material.

1.4 Theorem If $a(B) = b(B)\varphi(B)$ denotes the composite lag polynomial where $b_j \sim Kj^{d-1}$ for $|d| < \frac{1}{2}$ and a constant $K > 0$ and $|\varphi_j| = O(j^{-1-\delta})$ for $\delta > 0$, then

$$a_j \sim \varphi(1)Kj^{d-1} \tag{1.22}$$

as $j \to \infty$.

Proof Gathering terms of the product with matching powers of L shows that

$$a_j = \sum_{k=0}^{j} \varphi_k b_{j-k}$$

for each $j \geq 0$. Choose η from the interval $(1/(1+\delta), 1)$ and so write

$$a_j \sim K\sum_{k=0}^{j-1} \varphi_k (j-k)^{d-1} = Kj^{d-1}\Big(\frac{j-j^\eta}{j}\Big)^{d-1} \sum_{k=0}^{j-1} \varphi_k \Big(\frac{j-k}{j-j^\eta}\Big)^{d-1}. \tag{1.23}$$

Break up the sum on the right-hand side of (1.23) as

$$\Bigg(\sum_{k=0}^{[j^\eta]-1} + \sum_{k=[j^\eta]}^{j-1}\Bigg)\varphi_k \Big(\frac{j-k}{j-j^\eta}\Big)^{d-1} = A(j) + B(j). \tag{1.24}$$

Since $(j-k)/(j-j^\eta) \to 1$ as $j \to \infty$ for any fixed $k < [j^\eta]$, it is immediate that $A(j) \to \varphi(1)$. In view of (1.23) and the fact that $(j-j^\eta)/j \to 1$, the proof is completed by showing $B(j) \to 0$.

To do this, it is convenient to re-order the terms of (1.24) after taking the absolute value. Substitute $|\varphi_k| \backsimeq k^{-1-\delta}$ and also let $m = j - k$, so as to write

$$\begin{aligned}
|B(j)| &\ll \sum_{m=1}^{j-[j^\eta]} (j-m)^{-1-\delta}\Big(\frac{m}{j-j^\eta}\Big)^{d-1} \\
&= (j-j^\eta)^{1-d} j^{-\eta(1+\delta)} \sum_{m=1}^{j-[j^\eta]} \Big(\frac{j-m}{j^\eta}\Big)^{-1-\delta} m^{d-1} \\
&\leq (j-j^\eta)^{1-d} j^{-\eta(1+\delta)} \sum_{m=1}^{j-[j^\eta]} m^{d-1} \\
&\ll (j-j^\eta) j^{-\eta(1+\delta)}.
\end{aligned}$$

The second inequality holds because $j - m \geq j^\eta$ so that the factors in the terms of the sum with exponent $-1-\delta$ are all smaller than 1. The sum of the terms m^{d-1} is of order of magnitude $(j-j^\eta)^d$ by integral approximation[7] and the conclusion $|B(j)| \to 0$ follows noting that $\eta(1+\delta) > 1$ by choice of η. ∎

The implication is that linear short-run dependence has no asymptotic implications. The coefficients of the infinite order moving average (MA(∞)) process

$$x_i = (1-B)^{-d}\varphi(B)u_i \tag{1.25}$$

with $|d| < \frac{1}{2}$ satisfy the conditions specified in (1.2) where $\varphi(B)$ is any lag polynomial with summable coefficients and $L(j) \to \varphi(1)/\Gamma(d+1)$ as $j \to \infty$. Any form of shock dependence that can be removed by linear filtering is compatible under (1.2) with Assumption **1.1**.

[7] See SLT Theorem 2.17(i).

Chapter 2

Fractional Asymptotics

From the point of view of inference the main interest in fractional processes is with the associated partial sum processes, since it is the time aggregation that gives rise to known distributions in the limit. The familiar case of Brownian motion B has the feature that $B(at) \sim_{\mathrm{d}} a^{1/2}B(t)$ for $t \geq 0$ and $a \geq 0$, where $\sim_{\mathrm{d}}$ denotes equivalence in distribution. This is an example of the self-similarity property of a process, by which the distribution is preserved under changes of scale with suitable normalization. John Lamperti ([40]) first suggested extending the class of self-similar processes X to cases of the type $X(at) \sim_{\mathrm{d}} a^{\alpha}X(t)$, with the self-similarity index α falling either above or below $\frac{1}{2}$. The same idea was explored by Benoit Mandelbrot ([43], [44]), who defined fractional Brownian motion, commonly abbreviated as fBM, to be the Gaussian case of this generalization. For an excellent survey of its properties by Murad Taqqu, see [69]. As previously remarked, the symbol H (after Hurst) is commonly used for the self-similarity index in the statistical literature, corresponding to $d+\frac{1}{2}$, so that $0 < H < 1$ is the range of interest matching $|d| < \frac{1}{2}$.

2.1 Fractional Brownian Motion

Fractional Brownian motion is the almost surely continuous stochastic process $X : [0,1] \mapsto \mathbb{R}$ that is represented in the time domain by

$$X(t) = \int_0^t (t-\xi)^d \mathrm{d}U(\xi) + \int_{-\infty}^0 ((t-\xi)^d - (-\xi)^d)\mathrm{d}U(\xi), \; t \in [0,1] \tag{2.1}$$

where $|d| < \frac{1}{2}$ and U denotes a Brownian motion process with variance $\mathrm{E}(U(1)^2) = \sigma_u^2$, where without loss of generality it can be stipulated that $U(0) = 0$. The slightly counter-intuitive notion of $U(\xi)$ evolving from a starting point at $\xi = -\infty$ can be better appreciated by noting that a Brownian motion has independent increments and is accordingly reversible. The distribution on the negative half-line can be paired with that on the positive half-line initialized at zero. While of course U

diverges in the limit with probability 1, only the stationary increments dU enter the formula in (2.1), having square-integrable weights in the second term.

The case of regular Brownian motion is embedded in (2.1) as the case $d = 0$, the formula reducing to $U(t)$ for $t \in [0, 1]$ with the presample component cancelling. The convention introduced by Mandelbrot and van Ness ([44]) is to include the scale factor $1/\Gamma(d + 1)$ in the formula. As explained in Chapter 9, this scale factor aligns the variance of the process in (2.1) with that of the harmonizable representation of fBM in the frequency domain, although this is only true if the process driving the harmonizable representation has variance matching that of U. This may be an attractive assumption, as when both variances can be set to 1, but the factor otherwise achieves nothing material and tends to clutter algebraic expressions, hence it is omitted here.

The striking feature of (2.1) is the dependence of the process on the entire past history of the driving process U. This is of course a characteristic of (1.1) generally and hence of the ARMA class of processes, but weak dependence means that in large samples the presample contribution to the variation becomes negligible. Setting $d = 0$ in (2.1) gives regular Brownian motion $B(t) = \int_0^t dU(\xi)$, the second term in (2.1) vanishing. In fractional processes having $d > 0$, on the other hand, the presample contribution persists into the limit distribution.

Sometimes the expression in (2.1) is referred to as 'fractional Brownian motion of Type I' since an alternative formulation known as the Type II case (see [46]) is

$$X(t) = \int_0^t (t - \xi)^d \mathrm{d}U(\xi). \tag{2.2}$$

This is the asymptotic counterpart of the partial sum process driven by the truncated version of (1.1) in which $u_i = 0$ for $i \leq 0$. Since it features nonstationary increments, where the distributions are linked to the time elapsed since what is usually regarded as the first observation of a sample (i.e., unconnected with the phenomenon being modelled, in general) the Type II process is evidently a less satisfactory platform for empirical time series modelling than (2.1) (see [18]). Nonetheless, there is a substantial literature based around these truncated processes so it may be useful to point out that weak convergence to (2.2) can be analyzed by a somewhat more compact version of the arguments to be deployed here. This suggests a useful exercise for the reader.

2.2 The Variance

An increment of the fBM process of Type I for $0 \leq t < s \leq 1$ has the form

$$X(s) - X(t) = \int_t^s (s - \xi)^d \mathrm{d}U(\xi) + \int_{-\infty}^t \left((s - \xi)^d - (t - \xi)^d\right) \mathrm{d}U(\xi). \tag{2.3}$$

The signature property of fBM is the variance function of these increments.

2.1 Theorem If $|d| < \frac{1}{2}$, $\mathrm{E}(X(s) - X(t))^2 = \sigma_u^2|s-t|^{2d+1}\Upsilon_d < \infty$ for any $s, t \in [0,1]$ where

$$\Upsilon_d = \frac{1}{2d+1} + \int_0^\infty \left((1+\tau)^d - \tau^d\right)^2 \mathrm{d}\tau. \tag{2.4}$$

Proof Assume $s > t$ without loss of generality. The orthogonality of Brownian motion increments implies that

$$\mathrm{E}(X(s) - X(t))^2 = \sigma_u^2 \int_t^s (s-\xi)^{2d}\mathrm{d}\xi + \sigma_u^2 \int_{-\infty}^t \left((s-\xi)^d - (t-\xi)^d\right)^2 \mathrm{d}\xi. \tag{2.5}$$

Apply changes of variable $\tau = (s-\xi)/(s-t)$ in the first integral in (2.5) and $\tau = (t-\xi)/(s-t)$ in the second integral, noting $(s-\xi)/(s-t) = \tau + 1$. The integral in (2.4) converges for $|d| < \frac{1}{2}$. ∎

The increments of fBM are thus shown to be stationary, the variance depending on the width of the interval $|s-t|$, but not on t. Self-similarity implies that the distribution of the increment $X(t+\delta) - X(t)$ matches that of $a^{-d-1/2}(X(t+a\delta) - X(t))$ for any $a > 0$ (see [44]).

For Theorem **2.1** to be practically useful, a closed form for the expression Υ_d in (2.4) is desirable. The formula in question is as follows.

2.2 Theorem If $|d| < \frac{1}{2}$,

$$\Upsilon_d = \frac{\Gamma(1-2d)\Gamma(1+d)}{(2d+1)\Gamma(1-d)} \qquad \square \tag{2.6}$$

This formula is sometimes quoted with the additional factor $1/\Gamma(1+d)^2$, for the reasons explained in the remarks following (2.1). Note the salient fact that $\Upsilon_d = \infty$ in both of the boundary cases, $d = \frac{1}{2}$ and $d = -\frac{1}{2}$. As discussed in greater detail in §2.4, these divergences represent the breakdown of a.s. boundedness of X on the one hand and the failure of a.s. continuity of X on the other.

To prove (2.6) requires a digression into methods of complex analysis that do not otherwise feature in this theory and its study is therefore optional, but it may be of found of interest in its own right. The following reworks Lemma 5.1 of [17].

2.3 Lemma If $|d| < \frac{1}{2}$,

$$\int_0^\infty \left((\tau+1)^d - \tau^d\right)^2 \mathrm{d}\tau = \frac{1}{2d+1}\left(\frac{\Gamma(1-2d)\Gamma(1+d)}{\Gamma(1-d)} - 1\right) \tag{2.7}$$

Proof Consider the function of a complex variable

$$\phi(z) = \left((z+1)^d - z^d\right)^2, \; z \in \mathbb{C}.$$

Recalling $\mathrm{e}^{\mathrm{i}\pi} = -1$ define, for a real argument x,

$$z(x) = \begin{cases} x, & x \geq 0 \\ \mathrm{e}^{\mathrm{i}\pi}|x|, & x < 0. \end{cases} \tag{2.8}$$

Then,

$$f(x) = \phi \circ z = \left((z(x)+1)^d - z(x)^d\right)^2$$

is a well-defined function of a real variable $x \in \mathbb{R}$ whose integral over $[0, \infty)$ is the solution sought.

In the case $0 < d < \frac{1}{2}$, it can be verified both that $f(x) = O(|x|^{2d-2})$ as $x \to \pm\infty$ and that $\lim_{x\to 0} |f(x)| = 1$, so f is integrable for both positive and negative x. In the case $-\frac{1}{2} < d < 0$, f has a singularity at $x = 0$, but $\lim_{x\to 0} \left|f(x)/z(x)^{2d}\right| = 1$ where $2d > -1$ so it is integrable on the nonnegative half-line. It also has a singularity at $x = -1$ but $\lim_{x\to -1} |f(x)/(z(x)+1)^{2d}| = 1$ so it is likewise integrable on the negative half-line. Therefore, the integral

$$\mathcal{L}_d = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$$

exists for $|d| < \frac{1}{2}$.

Introduce the change of variable $y = -1 - x$, with inverse relation $x = -1 - y$. For each of the three cases $x > 0$, $-1 < x < 0$, and $x < -1$, it can be verified that $z(x) = \mathrm{e}^{\mathrm{i}\pi}(1 + z(y))$ and also that $1 + z(x) = \mathrm{e}^{\mathrm{i}\pi} z(y)$. Hence,

$$\begin{aligned}
\mathcal{L}_d &= \int_{-\infty}^{\infty} \left((z(x)+1)^d - z(x)^d\right)^2 \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left((\mathrm{e}^{\mathrm{i}\pi} z(y))^d - (\mathrm{e}^{\mathrm{i}\pi}(z(y)+1))^d\right)^2 \mathrm{d}y = \mathrm{e}^{2\mathrm{i}\pi d}\mathcal{L}_d. \qquad (2.9)
\end{aligned}$$

It follows from (2.9) that $\mathcal{L}_d = 0$ for $d \neq 0$. Direct inspection shows that $\mathcal{L}_0 = 0$ also holds.

Let $\mathcal{L}_{1d}$, $\mathcal{L}_{2d}$, and $\mathcal{L}_{3d}$ denote the integrals of $f(x)$ over the intervals $(-\infty, -1)$, $(-1, 0)$, and $(0, \infty)$ respectively, so that $\mathcal{L}_{3d}$ is the formula sought and $\mathcal{L}_d = \mathcal{L}_{1d} + \mathcal{L}_{2d} + \mathcal{L}_{3d} = 0$ for all d. Analogously to (2.9) the change of variable $y = -1-x$ gives

$$\begin{aligned}
\mathcal{L}_{1d} &= \int_{-\infty}^{-1} \left((z(x)+1)^d - z(x)^d\right)^2 \mathrm{d}x \\
&= \int_{0}^{\infty} \left((-z(y))^d - (-z(y)-1)^d\right)^2 \mathrm{d}y = \mathrm{e}^{2\mathrm{i}\pi d}\mathcal{L}_{3d}.
\end{aligned}$$

Next, multiply out $f(x)$ and use the facts that with x in the interval $(-1, 0)$ it is possible to write both $z(x)+1 = 1 - |x|$ and $z(x) = \mathrm{e}^{\mathrm{i}\pi}|x|$, also with $|x|$ becoming x under the change of variable $x = -x$. Thus,

$$\begin{aligned}
\mathcal{L}_{2d} &= \int_{-1}^{0} (z(x)+1)^{2d}\mathrm{d}x + \int_{-1}^{0} z(x)^{2d}\mathrm{d}x - 2\int_{-1}^{0} (z(x)+1)^d z(x)^d \mathrm{d}x. \\
&= \frac{1-0}{2d+1} + \frac{0 - \mathrm{e}^{\mathrm{i}\pi(2d+1)}}{2d+1} - 2\mathrm{e}^{\mathrm{i}\pi d}\int_0^1 x^d(1-x)^d\,\mathrm{d}x \\
&= \frac{1+\mathrm{e}^{2\mathrm{i}\pi d}}{2d+1} - 2\mathrm{e}^{\mathrm{i}\pi d}B(d+1, d+1) \qquad (2.10)
\end{aligned}$$

where the last equality notes the definition of the Beta function from (B.14) of Appendix B. Adding over the three intervals yields the relation

$$\mathrm{e}^{2\mathrm{i}\pi d}\mathcal{L}_{3d} + \frac{1+\mathrm{e}^{2\mathrm{i}\pi d}}{2d+1} - 2\mathrm{e}^{\mathrm{i}\pi d}B(d+1,d+1) + \mathcal{L}_{3d} = 0 \tag{2.11}$$

which after rearrangement using (B.2) solves as

$$\begin{aligned}\mathcal{L}_{3d} &= \frac{1}{1+\mathrm{e}^{2\mathrm{i}\pi d}}\Big(-\frac{1+\mathrm{e}^{2\mathrm{i}\pi d}}{2d+1} + 2\mathrm{e}^{i\pi d}B(d+1,d+1)\Big)\\ &= \frac{B(d+1,d+1)}{\cos(\pi d)} - \frac{1}{2d+1}.\end{aligned} \tag{2.12}$$

By identity (B.8) and double applications of (B.15) and (B.13),

$$\cos(\pi d) = \frac{\sin(2\pi d)}{2\sin(\pi d)} = \frac{\Gamma(1-d)\Gamma(d+1)}{\Gamma(1-2d)\Gamma(2d+1)} \tag{2.13}$$

and also, by the second equality of (B.14) and (B.13),

$$B(d+1,d+1) = \frac{\Gamma(d+1)^2}{(2d+1)\Gamma(2d+1)}. \tag{2.14}$$

After substitution from (2.13) and (2.14) and simplification, (2.12) is seen to match the formula in (2.7). ∎

Proof of 2.2 Immediate on substituting (2.7) into (2.4) and simplifying. ∎

An alternative derivation of formula (2.6) is given in Chapter 9, using the harmonizable representation of the process; see Theorem **9.2**. Another version of the formula that is sometimes quoted is

$$\Upsilon_d = \frac{\Gamma(d+1)^2}{\Gamma(2d+2)\cos\pi d}. \tag{2.15}$$

The proof that (2.15) matches (2.6) is a simple matter of applying (2.13) in reverse.

A useful identity, defined for any pair of time intervals $s_1 > t_1$ and $s_2 > t_2$, is

$$\begin{aligned}&(X(s_1) - X(t_1))(X(s_2) - X(t_2))\\ &\quad= \tfrac{1}{2}\Big((X(s_1) - X(t_2))^2 + (X(s_2) - X(t_1))^2\\ &\qquad\qquad - (X(t_2) - X(t_1))^2 - (X(s_2) - X(s_1))^2\Big).\end{aligned} \tag{2.16}$$

With Theorem **2.1** this allows the covariance of any pair of process increments to be calculated. Setting $t_1 = t_2 = 0$ in (2.16), and also $s_1 = s$ and $s_2 = t$, gives

$$\mathrm{E}(X(s)X(t)) = \tfrac{1}{2}\sigma_u^2\Upsilon_d\big(s^{2d+1} + t^{2d+1} - |s-t|^{2d+1}\big) \tag{2.17}$$

which with $d = 0$ reduces to the familiar case of regular Brownian motion, having $\mathrm{E}(X(s)X(t)) = \sigma_u^2\min(t,s)$. For another example, set $t_1 = s_2 = t$, $s_1 = t+\delta$, and

$t_2 = t - \delta$ in (2.16) to give the covariance of a pair of adjacent non-overlapping increments,

$$\mathrm{E}((X(t+\delta) - X(t))((X(t) - X(t-\delta)) = \sigma_u^2 \Upsilon_d (2^{2d} - 1)\delta^{2d+1}.$$

This is 0 if $d = 0$ and otherwise has the sign of d.

It is reasonable to ask how a stochastic process such as (2.1) can be understood to connect with the infinitely remote past. To elucidate this, imagine a version of (2.1) in which the Brownian motion driving the process has a finitely remote starting point, say $-N$ for some $N \in \mathbb{N}$. That is, define a process X^N by

$$X^N(t) = \int_0^t (t-\xi)^d \mathrm{d}U(\xi) + \int_{-N}^0 ((t-\xi)^d - (-\xi)^d)\mathrm{d}U(\xi),\ t \in [0,1]. \tag{2.18}$$

A minor amendment of Theorem **2.1** now gives the following.

2.4 Corollary $\mathrm{E}(X^N(t)^2) = \sigma_u^2 \Upsilon_d^N t^{2d+1}$ where

$$\Upsilon_d^N = \frac{1}{2d+1} + \int_0^N \big((1+\tau)^d - \tau^d\big)^2 \mathrm{d}\tau. \quad \square \tag{2.19}$$

It is easy to verify that for $t \in (0,1]$,

$$\begin{aligned}\mathrm{E}\big(X(t) - X^N(t)\big)^2 &= \sigma_u^2 t^{2d+1}(\Upsilon_d - \Upsilon_d^N) \\ &= \sigma_u^2 t^{2d+1} \int_N^\infty \big((1+\tau)^d - (\tau)^d\big)^2 \mathrm{d}\tau \\ &= O(N^{2d-1}) \end{aligned} \tag{2.20}$$

as $N \to \infty$. In other words, subject to the condition $d < \frac{1}{2}$, X exists as the mean square limit of a sequence of finite-lag processes X^N. This decomposition plays a major role in the treatment of the functional central limit theorem in Chapter 3.

2.3 The Linear Structure

The next objective is to show that when the shock sequence in (1.1) satisfies either Assumption **1.1** or Assumption **1.2**, the normalized partial sum of a fractional moving average that is stationary and not overdifferenced converges to a limit having the variance function specified in Theorem **2.1**. The weak convergence proof itself is given in Chapter 3. The remainder of this preliminary chapter explores some essential properties of the objects under study and develops techniques of analysis.

With x_i given by (1.1) let

$$S_n = \sum_{i=1}^n x_i. \tag{2.21}$$

The representation in (1.3) suggests that this sum would be Gaussian in the limit if supplied with an appropriate normalization, although also showing that this normalization could not be $n^{-1/2}$. However, in the framework of (2.21) the mechanics of calculating moments and so forth would be daunting. S_n is a sum of sums and the essential trick, underlying all that follows, is to aggregate the components in a different order. Thus,

$$\begin{aligned} S_n &= \sum_{i=1}^{n}\sum_{j=0}^{\infty} b_j u_{i-j} \\ &= (b_0 u_n + b_1 u_{n-1} + \cdots) + (b_0 u_{n-1} + b_1 u_{n-2} + \cdots) \\ &\qquad\qquad + \cdots + (b_0 u_1 + b_1 u_0 + \cdots) \\ &= b_0 u_n + (b_0 + b_1) u_{n-1} + \cdots + (b_0 + \cdots + b_{n-1}) u_1 \\ &\qquad + (b_1 + \cdots + b_n) u_0 + (b_2 + \cdots + b_{n+1}) u_{-1} + \cdots \\ &= \sum_{i=-\infty}^{n} a_{ni} u_i \end{aligned} \tag{2.22}$$

where

$$a_{ni} = \sum_{j=\max\{0,1-i\}}^{n-i} b_j. \tag{2.23}$$

In this representation of S_n the terms are either independent or, at worst, weakly dependent, but they form a nonstationary sequence in view of (2.23) and are infinite in number.

More generally, the notation to be used in the sequel is

$$S_{[ns]} - S_{[nt]} = \sum_{i=-\infty}^{[ns]} a_{ni}(s,t) u_i \tag{2.24}$$

for any $0 \le t < s \le 1$, where

$$a_{ni}(s,t) = \sum_{j=\max\{0,[nt]-i+1\}}^{[ns]-i} b_j. \tag{2.25}$$

In particular, note that $a_{ni}(t,t) = 0$ for any i, representing an empty sum, so that $S_{[ns]} - S_{[nt]} = 0$ if $s = t$. Also, $a_{n[ns]}(s,t) = b_0 = 1$ when n is large enough that $[ns] > [nt]$. Where convenient, a_{ni} as in (2.23) will continue to be used as shorthand for the case $a_{ni}(1,0)$.

Assuming for simplicity that the coefficients b_j vary like (1.2) for all j and not just in the limit, consider how the sequence of weights $a_{ni}(s,t)$ in (2.25) varies in the contrasting cases $d > 0$ and $d < 0$. When $d > 0$, $a_{ni}(s,t)$ would take its maximum at $i = [nt] + 1$, at which point it has $[ns] - [nt]$ terms and diverges at the rate n^d. As i increases, the number of terms in $a_{ni}(s,t)$ decreases down to the single term $b_0 = 1$ at the point $i = [ns]$. Moving in the other direction, with

$i \leq [nt]$, the terms of the sum in (2.25) are $b_{1-i}, \ldots, b_{[ns]-i}$. While there are $[ns]$ of these terms, for a given finite n their sum is tending to zero, with $a_{ni} = O(|i|^{d-1})$ as $i \to -\infty$.

In the antipersistent case with $d < 0$, $a_{n[ns]}(s,t) = b_0 = 1$ but the b_j are negative for $j > 0$, with $\sum_{j=0}^{\infty} b_j = 0$. Letting i decrease from $[ns]$ down to 1 the sum accumulates negative terms and so declines towards zero. There is a discontinuity in the function at the point $i = [nt]$ where it jumps to approximately -1 when n is large, being the sum of the first $[ns]$ coefficients starting from $j = 1$. Thereafter it forms an increasing (absolutely decreasing) negative sequence, tending to 0 with $a_{ni} = O(|i|^{d-1})$ as $i \to -\infty$.

With this framework established, let

$$\kappa(n) = n^{d+1/2}L(n) \tag{2.26}$$

and hence define the normalized partial sum process $X_n : [0,1] \mapsto \mathbb{R}$ where

$$X_n(t) = \frac{S_{[nt]}}{\kappa(n)} = \frac{1}{\kappa(n)} \sum_{i=-\infty}^{[nt]} a_{ni}(t,0)u_i. \tag{2.27}$$

Following (2.24) the variation over an interval $(t,s]$ has the form

$$X_n(s) - X_n(t) = \frac{1}{\kappa(n)} \sum_{i=-\infty}^{[ns]} a_{ni}(s,t)u_i. \tag{2.28}$$

The process defined by (2.27) is a step function, being constant but for jumps at the points where $t = [nt]/n$ and taking its value at the terminal points of the jumps. While not a continuous function of time it is right-continuous, with every point having a limit point to its left. Such processes are commonly referred to by the colourful French acronym 'càdlàg'. If the process has the unit interval as domain, as is commonly the case, the space of càdlàg processes is denoted $D_{[0,1]}$, of which the space of continuous processes, $C_{[0,1]}$, forms a subset.

A technical detail that can be overlooked in the present context, but matters for convergence proofs, is that $D_{[0,1]}$ is customarily endowed with the Skorokhod J1 topology. Distances between elements x and y of $C_{[0,1]}$ are defined by the so-called uniform metric, $d_U(x,y) = \sup_t |x(t) - y(t)|$, but this has undesirable consequences in the presence of discontinuities. Without going into too much detail,[1] the Skorokhod distance between càdlàg processes x and y is the smallest number $d_S(x,y)$, by suitable choice of homeomorphism $\lambda : [0,1] \mapsto [0,1]$, such that both $\sup_t |x(t) - y(\lambda(t))| \leq d_S(x,y)$ and $\sup_t |\lambda(t) - t| \leq d_S(x,y)$. Thus, functions are close in the Skorokhod topology if their discontinuities are close in time as well as in magnitude. In practice, distances are assigned by a separable complete metric embodying the topology, such as that constructed by Billingsley ([6], [7]).

[1] A full account can be found in SLT Chapter 30.

2.4 Limiting Forms

Allowing the general formulation in (1.2) for the lag coefficients, the decomposition in (2.25) exhibits the following tendencies as n increases. In this context, recalling the interpretation of series such as $\{b_j\}_{j=0}^{\infty}$, note that 0^{d-1} is always assigned the value 1.

2.5 Theorem Let $a_{ni}(s,t)$ be defined by (2.25) for $0 \le t < s \le 1$ where the b_j satisfy (1.2) for $|d| < \frac{1}{2}$ and $\sum_{j=0}^{\infty} b_j = 0$ if $d < 0$. For $t < x < s$,

$$a_{n[nx]}(s,t) \sim (n(s-x))^d L(n(s-x)) \tag{2.29}$$

and for $-\infty < x \le t$,

$$a_{n[nx]}(s,t) \sim \big((n(s-x))^d - (n(t-x))^d\big) L(n(s-x)) \tag{2.30}$$

as $n \to \infty$.

Proof First, suppose $0 < d < \frac{1}{2}$. If m is a positive integer,

$$\int_0^m y^{d-1}L(y)\mathrm{d}y = \sum_{j=0}^{m-1} \int_j^{j+1} y^{d-1}L(y)\mathrm{d}y.$$

Let $L(y)$ denote a slowly varying function of a real-valued argument, that matches the definition of b_j in (1.2) at the integer points and that for large enough j satisfies the inequalities

$$(j+1)^{d-1}L(j+1) \le \int_j^{j+1} y^{d-1}L(y)\mathrm{d}y \le j^{d-1}L(j). \tag{2.31}$$

These terms are nonsummable over j so there exists m large enough that

$$\sum_{j=1}^{m} j^{d-1}L(j) \le \int_0^m y^{d-1}L(y)\mathrm{d}y \le \sum_{j=0}^{m-1} j^{d-1}L(j). \tag{2.32}$$

For the case $t < x < s$, set $m = [ns] - [nx]$ in (2.32). It follows according to (2.25) and (1.2) that

$$a_{n[nx]}(s,t) = \sum_{j=0}^{[ns]-[nx]} b_j \sim d \int_0^{[ns]-[nx]} y^{d-1}L(y)\mathrm{d}y. \tag{2.33}$$

By Theorem **A.8**(i) of Appendix A, the integral in (2.33) has solution corresponding to (2.29) as $n \to \infty$. Similarly if $x \le t$ then

$$a_{n[nx]}(s,t) = \sum_{j=[nt]+1-[nx]}^{[ns]-[nx]} b_j \sim d \int_{[nt]-[nx]}^{[ns]-[nx]} y^{d-1}L(y)\mathrm{d}y \tag{2.34}$$

with solution matching (2.30).

Now consider $d < 0$ and $\sum_{j=0}^{\infty} b_j = 0$. For $t < x < s$, given (2.31) for large enough j,

$$a_{n[nx]}(s,t) = \sum_{j=0}^{[ns]-[nx]} b_j = -\sum_{j=[ns]+1-[nx]}^{\infty} b_j \sim -d\int_{[ns]-[nx]}^{\infty} y^{d-1}L(y)\mathrm{d}y$$

and the solution is in this case given by Theorem **A.8**(ii). If $x \le t$, similarly,

$$\begin{aligned} a_{n[nx]}(s,t) = \sum_{j=[nt]+1-[nx]}^{[ns]-[nx]} b_j &= \sum_{j=[nt]+1-[nx]}^{\infty} b_j - \sum_{j=[ns]+1-[nx]}^{\infty} b_j \\ &\sim d\int_{[nt]-[nx]}^{[ns]-[nx]} y^{d-1}L(y)\mathrm{d}y. \quad \blacksquare \end{aligned}$$

The following limit result applies the tendencies established by Theorem **2.5** to show the implication of having formula (1.2) constrain the moving average coefficients at long range.

2.6 Theorem If $|d| < \frac{1}{2}$,

$$\sum_{i=-\infty}^{[ns]} a_{ni}(s,t)^2 \sim \Upsilon_d (n(s-t))^{2d+1} L(n(s-t))^2. \tag{2.35}$$

Proof Define sums M_{1n} and M_{2n} by

$$\sum_{i=-\infty}^{[ns]} a_{ni}(s,t)^2 = \sum_{i=[nt]+1}^{[ns]} a_{ni}(s,t)^2 + \sum_{i=-\infty}^{[nt]} a_{ni}(s,t)^2 = M_{1n} + M_{2n}. \tag{2.36}$$

For fixed values of s and t, not indicated explicitly, let A_{ni} denote the approximator in (2.29) for $i = [nx]$ and let

$$g_{ni} = \frac{|a_{ni}^2(s,t) - A_{ni}^2|}{A_{ni}^2}. \tag{2.37}$$

Summing over indices $i = [nt]+1, \ldots, [ns]$ specified for M_{1n},

$$\left| \frac{\sum_i a_{ni}^2(s,t)}{\sum_i A_{ni}^2} - 1 \right| \le \frac{\sum_i g_{ni} A_{ni}^2}{\sum_i A_{ni}^2} \le \max_{[nt]<i\le[ns]} g_{ni} \to 0 \tag{2.38}$$

as $n \to \infty$, by the modulus inequality and Theorem **2.5**. Approximating sum by integral analogously to the argument leading to (2.33) and applying Theorem **A.8**(i), with $d > -\frac{1}{2}$ (2.38) implies

$$M_{1n} \sim \sum_{i=[nt]+1}^{[ns]} A_{ni}^2 \sim \int_0^{n(s-t)} y^{2d} L(y)^2 \mathrm{d}y \sim \frac{(n(s-t))^{2d+1}}{2d+1} L(n(s-t))^2. \tag{2.39}$$

For the case $-\infty < i \leq [nt]$, substitute the approximator from (2.30) into (2.37) and construct the counterpart of (2.39) for M_{2n} to get

$$\begin{aligned} M_{2n} &\sim \int_{-\infty}^{nt} \big((ns-y)^d - (nt-y)^d\big)^2 L(ns-y)^2 \mathrm{d}y \\ &\sim (n(s-t))^{2d+1} L(n(s-t))^2 \int_0^{\infty} ((1+\tau)^d - \tau^d)^2 \mathrm{d}\tau. \end{aligned} \tag{2.40}$$

The second asymptotic equivalence makes the change of variable $\tau = (nt-y)/(ns-nt)$ and uses the fact that $L(n(s-t)x)/L(n(s-t)) \to 1$ as $n \to \infty$ for both $x = \tau$ and $x = 1+\tau$. ∎

Had b_j been defined to have the form $dj^{d-1}L(j)/\,\Gamma(d+1)$ in place of (1.2), such as might be motivated by the particular case in (1.12), the formula in (2.35) would need to include the factor $1/\Gamma(d+1)^2$. However, notwithstanding the remark following Theorem **2.2**, formula (2.6) for Υ_d is maintained in what follows.

The limiting variance of an increment of the process X_n is going to be a central feature of the asymptotic analysis and the following are immediate consequences of (2.24) and Theorem **2.6**.

2.7 Corollary Under Assumption **1.1** and with $|d| < \frac{1}{2}$,

$$\mathrm{E}\Bigg(\sum_{i=[nt]+1}^{[ns]} x_i\Bigg)^2 \sim \sigma_u^2 \Upsilon_d (n(s-t))^{2d+1} L(n(s-t))^2. \;\square \tag{2.41}$$

2.8 Corollary Under Assumption **1.1** and $|d| < \frac{1}{2}$, for $t \in [0,1]$ and $\delta \in (0, 1-t]$,

$$\lim_{n\to\infty} \mathrm{E}\big(X_n(t+\delta) - X_n(t)\big)^2 = \sigma_u^2 \Upsilon_d \delta^{2d+1}. \tag{2.42}$$

Proof Definition (2.27) implies

$$\mathrm{E}\big(X_n(t+\delta) - X_n(t)\big)^2 = \frac{\sigma_u^2}{\kappa(n)^2} \sum_{i=-\infty}^{[n(t+\delta)]} a_{ni}(t+\delta, t)^2. \tag{2.43}$$

Substitute into (2.43) from (2.35) and simplify. ∎

These calculations go some way to explaining the behaviour of the process at the boundary points $-\frac{1}{2}$ and $+\frac{1}{2}$. In the nonstationary case $d = \frac{1}{2}$ the integral in (2.40) diverges, showing that the remote data points are so influential that the partial sum process eventually blows up and cannot approach the limit distribution (2.1). The rate at which the tail component in (2.20) disappears with N is an indicator of what happens as the boundary is approached. In the case $d = -\frac{1}{2}$ the integral in (2.39) diverges. The reasoning in this case is a little more subtle, but the limit in (2.42) gives the clue. The constant Υ_d might be removed by choice of normalization and it is not so much that this quantity diverges with $d = -\frac{1}{2}$, as the fact that δ^{2d+1} does not approach 0 as $\delta \to 0$. This shows that it is almost sure continuity that fails.

The boundary points are awkward since except at these points a well-defined limit distribution of the partial sums can always be obtained, if necessary by differencing or cumulating the underlying process provided in the latter case that an initial value (i.e. constant of integration) is known. However, the difference between the case $d = \frac{1}{2}$ and a stationary alternative is at worst a logarithmic divergence. In the modelling context, a test of the point hypothesis $d = \frac{1}{2}$ has no obvious rationale comparable to (for example) the cases $d = 0$ or $d = 1$.

2.5 The Multivariate Model

Extending the fractional model to describe the interactions of two or more related processes is very largely a matter of setting up appropriate notation. Let $\boldsymbol{\Delta}(B)$ denote a diagonal $m \times m$ matrix polynomial whose diagonal elements are of the form $b_k(B) = \sum_{j=0}^{\infty} b_{kj} B^j$ for $k = 1, \ldots, m$, where B denotes the backshift operator and $b_{kj} \sim d_k j^{d_k - 1} L_k(j)$, with $|d_k| < \frac{1}{2}$ for each k. The parameters $d_1, \ldots, d_m$ are assumed to be in order of magnitude with $d_1 \le \cdots \le d_m$ and the functions $L_1(j), \ldots, L_m(j)$ are either slowly varying or constant at infinity. A m-vector of fractional processes is then represented by

$$\boldsymbol{x}_i = \boldsymbol{\Delta}(B)\boldsymbol{u}_i \quad (m \times 1). \tag{2.44}$$

Define $m \times m$ matrices

$$\boldsymbol{D}_n = \mathrm{diag}(n^{d_1 + 1/2} L_1(n), \ldots, n^{d_m + 1/2} L_m(n)) \tag{2.45}$$

and

$$\boldsymbol{A}_{ni}(s,t) = \mathrm{diag}(a_{1ni}(s,t), \ldots, a_{mni}(s,t)) \tag{2.46}$$

where $a_{kni}(s,t)$ is the case of (2.25) with $d = d_k$ and so define the vector $\boldsymbol{X}_n(t)$ $(m \times 1)$ for $0 < t \le 1$ in terms of the process increment by

$$\boldsymbol{X}_n(s) - \boldsymbol{X}_n(t) = \sum_{i=[nt]+1}^{[ns]} \boldsymbol{D}_n^{-1} \boldsymbol{x}_i = \sum_{i=-\infty}^{[ns]} \boldsymbol{D}_n^{-1} \boldsymbol{A}_{ni}(s,t) \boldsymbol{u}_i. \tag{2.47}$$

Assume that the elements of the m-vector shock process $\{\boldsymbol{u}_i, -\infty < i < \infty\}$ individually satisfy the conditions of Assumption **1.1**, and have the contemporaneous covariance matrix

$$\boldsymbol{\Sigma}_u = \{\sigma_{kl}\} = \mathrm{E}(\boldsymbol{u}_i \boldsymbol{u}_i') \quad (m \times m). \tag{2.48}$$

Also define the matrix $\boldsymbol{\Upsilon} = \{\Upsilon_{kl}\}$ $(m \times m)$ where

$$\Upsilon_{kl} = \frac{1}{d_k + d_l + 1} + \int_0^{\infty} ((1+\tau)^{d_k} - \tau^{d_k})((1+\tau)^{d_l} - \tau^{d_l}) \mathrm{d}\tau. \tag{2.49}$$

The Hadamard product $\boldsymbol{\Sigma}_u \odot \boldsymbol{\Upsilon}$ is the $m \times m$ matrix having elements $\sigma_{kl}\Upsilon_{kl}$ for $k, l = 1, \ldots, m$. Finally, for $0 < \delta \leq 1$ let

$$\boldsymbol{K}(\delta) = \operatorname{diag}(\delta^{d_1+1/2}, \ldots, \delta^{d_m+1/2}).$$

This setup permits the following generalization of Corollary **2.8**.

2.9 Theorem If Assumption **1.1** holds for each element of $\boldsymbol{u}_i$,

$$\lim_{n\to\infty} \mathrm{E}(\boldsymbol{X}_n(t+\delta) - \boldsymbol{X}_n(t))(\boldsymbol{X}_n(t+\delta) - \boldsymbol{X}_n(t))' = \boldsymbol{K}(\delta)(\boldsymbol{\Sigma}_u \odot \boldsymbol{\Upsilon})\boldsymbol{K}(\delta) \quad (2.50)$$

for $0 < \delta \leq 1$ and $0 \leq t < 1 - \delta$.

Proof Modify the proof of Theorem **2.6** as follows. Replace the term $a_{ni}(s,t)^2$ in (2.36) by $a_{kni}(s,t)a_{lni}(s,t)$ and replace A_{ni}^2 by the product $A_{kni}A_{lni}$ where these factors represent the approximator functions (2.29) and (2.30) evaluated at parameters d_k and d_l. The result that

$$\Upsilon_{kl} = \lim_{n\to\infty} \frac{1}{(n\delta)^{d_k+d_l+1}L_k(n)L_l(n)} \sum_{i=-\infty}^{[n(t+\delta)]} a_{kni}(t+\delta, t)a_{lni}(t+\delta, t) \qquad (2.51)$$

for each pair k, l and any choice of t and δ is then a straightforward extension. The modifications of the asymptotic equivalences in (2.39) and (2.40) are then direct, with $d_k + d_l$ replacing $2d$ and the squared terms replaced by products in the cases $k \neq l$.

Given (2.51), the argument of Corollary **2.8** can be applied to each element of the expected outer product of (2.50) with one modification, that according to (2.47) the $(k,l)^{\text{th}}$ divisor has the form $n^{d_k+d_l+1}L_k(n)L_l(n)$ in place of $\kappa(n)^2$ as appears in equation (2.43). The conclusion is therefore that, for each pair (k,l),

$$\mathrm{E}(X_{kn}(t+\delta) - X_{kn}(t))(X_{ln}(t+\delta) - X_{ln}(t)) \to \sigma_{kl}\Upsilon_{kl}\delta^{d_k+d_l+1}. \qquad (2.52)$$

These limits constitute the elements of the limit matrix in (2.50). ∎

A closed form for (2.49) is obtainable by extending the technique of Lemma **2.3**, but this will be most conveniently given in Chapter 4, as Theorem **4.6**. The context of this latter result is a bivariate analysis, but of course a covariance is necessarily a pairwise construction. Adapted to the present notation, the formula in (4.34) takes the form

$$\Upsilon_{kl} = \frac{\Gamma(d_k+1)\Gamma(d_l+1)\cos(\pi(d_k-d_l)/2)}{\Gamma(d_k+d_l+2)\cos(\pi(d_k+d_l)/2)}. \qquad (2.53)$$

As required, $\Upsilon_{kl} = \Upsilon_{lk}$ and the formula collapses to Υ_d, as in (2.15), on setting $d_k = d_l = d$. Another derivation, based on the harmonizable representation of the processes, is Theorem **9.3**.

It is tempting to ask whether formula (2.16) has a generalization to allow the calculation of covariances for increments over differing time intervals. Unequivocally the answer to this question is no, covariances can shed no light on the time-ordering of events. This fact can be illustrated by modifying (2.16) for a sequence pair (X_k, X_l), replacing the squares of increments with contemporaneous products. For simplicity's sake set $t_1 = t_2 = 0$ assuming $X_k(0) = X_l(0) = 0$, also putting $s_1 = s$ and $s_2 = t$. This gives the equality

$$\begin{aligned} X_k(s)X_l(s) + X_k(t)X_l(t) - &(X_k(s) - X_k(t))(X_l(s) - X_l(t)) \\ &= X_k(s)X_l(t) + X_k(t)X_l(s). \end{aligned} \tag{2.54}$$

The expectation of the left-hand side of (2.54) can be calculated by (2.52), but it is not possible to determine either $\mathrm{E}(X_k(s)X_l(t))$ or $\mathrm{E}(X_k(t)X_l(s))$ individually, only their sum. More generally, it can be pointed out that formula (2.52) is also shown by Theorem **9.3**. By construction, the harmonizable representation of the processes can contain no information about temporal orderings.

2.6 Shock Dependence

While there are many situations in which the result of replacing Assumption **1.1** by Assumption **1.2** implies no more changes to asymptotic results than the replacement of σ_u^2 by ω_u^2 in formulae, this can be quite tedious to demonstrate. As an example consider Corollary **2.8**, which is an easy consequence of Theorem **2.6**. Its generalization, given here for the case $t = 0$ and $\delta = 1$, depends on a blocking argument.

2.10 Theorem Under Assumption **1.2**,

$$\lim_{n\to\infty} \mathrm{E}\left(\frac{1}{\kappa(n)}\sum_{i=1}^{n} x_i\right)^2 = \omega_u^2 \Upsilon_d.$$

Proof For brevity, write as before a_{ni} for $a_{ni}(1, 0)$ defined in (2.25). Choose an increasing integer sequence $\{B_n\}$ such that $B_n \to \infty$ but $B_n/n \to 0$ and let $r_n = [n/B_n]$. Define, for $j = -\infty, \ldots, r_n$,

$$S_{nj} = \frac{1}{B_n^{1/2}} \sum_{i=(j-1)B_n+1}^{jB_n} u_i \tag{2.55}$$

and also

$$S_{nj}^* = \sum_{i=(j-1)B_n+1}^{jB_n} \frac{a_{ni} - a_{n,jB_n}}{B_n^{3/2} g_{nj}} u_i \tag{2.56}$$

where $g_{nj} > 0$ is an array to be chosen. Then, define A_{1n}, A_{2n}, and A_{3n} by

$$\begin{aligned} \sum_{i=1}^{n} x_i &= \sum_{i=-\infty}^{n} a_{ni}u_i \\ &= B_n^{1/2} \sum_{j=-\infty}^{r_n} a_{n,jB_n} S_{nj} + B_n^{3/2} \sum_{j=-\infty}^{r_n} g_{nj} S_{nj}^* + \sum_{i=r_nB_n+1}^{n} a_{ni}u_i \end{aligned}$$

$$= A_{1n} + A_{2n} + A_{3n}. \tag{2.57}$$

The object of the proof is to show that

$$\lim_{n\to\infty} \frac{\mathrm{E}(A_{1n}^2)}{\kappa(n)^2} = \omega_u^2 \Upsilon_d \tag{2.58}$$

where (2.26) defines $\kappa(n)$, whereas $\mathrm{E}(A_{2n}^2)$ and $\mathrm{E}(A_{3n}^2)$ and hence also the expected cross-products are of smaller order.

Since the autocovariances are summable by Assumption **1.2**(a),

$$\sup_{-\infty<j\le r_n} |\mathrm{E}(S_{nj}^2) - \omega_u^2| \to 0 \text{ as } n \to \infty. \tag{2.59}$$

Next, consider the covariance of blocks of observation S_{nj} and $S_{n,j-m}$ with $m \ge 1$. There are B_n^2 pairings of coordinates from each block, whose date separations $B_n m + p$ vary from $B_n(m-1)+1$ with $p = 1 - B_n$ to $B_n(m+1)-1$ with $p = B_n - 1$. Since $B_n - |p|$ of the pairs have separation $B_n m + p$ for each p,

$$|\mathrm{E}(S_{nj}S_{n,j-m})| \le \frac{1}{B_n} \sum_{p=1-B_n}^{B_n-1} (B_n - |p|)|\gamma_u(B_n m + p)| = O(m^{-1-\delta}B_n^{-\delta}). \tag{2.60}$$

From this it follows that

$$\sup_{-\infty<j\le r_n} \sum_{m=1}^{\infty} |\mathrm{E}(S_{nj}S_{n,j-m})| = O(B_n^{-\delta}). \tag{2.61}$$

Next, define

$$w_{nj} = \frac{B_n a_{n,jB_n}}{\sqrt{\sum_{i=-\infty}^{n} a_{ni}^2}}, \quad -\infty < j \le r_n.$$

From Theorem **2.5** and applying the argument of Theorem **2.6**,

$$\sum_{j=-\infty}^{r_n} a_{r_n j}^2 \sim \Upsilon_d r_n^{2d+1} L(r_n)^2. \tag{2.62}$$

It follows, since $(n - B_n j)^d \sim B_n^d (r_n - j))^d$ for $0 \le j \le r_n$ and $(-B_n j)^d = B_n^d(-j))^d$ for $j < 0$, that

$$B_n \sum_{j=-\infty}^{r_n} a_{n,jB_n}^2 \sim \Upsilon_d n^{2d+1} L(r_n)^2.$$

Since $L(r_n)/L(n) \to 1$ it further follows that $\sum_{j=-\infty}^{r_n} w_{nj}^2 \to 1$, as $n \to \infty$. Write

$$\left(\sum_{i=-\infty}^{n} a_{ni}^2\right)^{-1} \mathrm{E}(A_{1n}^2) = T_{1n} + 2T_{2n} \tag{2.63}$$

where

$$T_{1n} = \sum_{j=-\infty}^{r_n} w_{nj}^2 \mathrm{E}(S_{nj}^2), \quad T_{2n} = \sum_{j=-\infty}^{r_n-1} \sum_{m=1}^{r_n-j} w_{nj} w_{n,j+m} \mathrm{E}(S_{nj} S_{n,j+m}).$$

By (2.59),

$$|T_{1n} - \omega_u^2| \leq \sup_{-\infty<j\leq r_n} |\mathrm{E}(S_{nj}^2 - \omega_u^2)| \sum_{j=-\infty}^{r_n} w_{nj}^2 + \omega_u^2 \left| \sum_{j=-\infty}^{r_n} w_{nj}^2 - 1 \right| \to 0 \quad (2.64)$$

as $n \to \infty$. Also by (2.61), noting that $\sum_{j=-\infty}^{r_n-1} |w_{nj} w_{n,j+m}| \leq 1$ for $1 \leq m \leq r_n - j$,

$$|T_{2n}| \leq \sum_{j=-\infty}^{r_n-1} \sum_{m=1}^{r_n-j} |w_{nj} w_{n,j+m} \mathrm{E}(S_{nj} S_{n,j+m})| = o(1). \quad (2.65)$$

The implication of (2.63), (2.64), and (2.65), together with (2.26) and Theorem **2.6**, is that (2.58) is confirmed.

Next, consider A_{2n} in (2.57). A bound must be found the terms $|a_{ni} - a_{n,B_nj}|$, and in view of formula (2.25) with $t = 0$ and $s = 1$ these contain the sums of the $B_nj - i$ successive coefficients $b_{n-B_nj+1}, \ldots, b_{n-i}$. If the sequence $\{|b_j|\}_{j=1}^\infty$ is monotone decreasing then $|a_{ni} - a_{n,B_nj}| \leq (B_nj - i)|b_{n-B_nj+1}|$. In this case, let the choice in (2.56) be

$$g_{nj} = |b_{n-B_nj+1}| \quad (2.66)$$

so that $|a_{ni} - a_{nB_nj}| \leq B_n g_{nj}$ for $B_n(j-1) < i \leq B_nj$. More generally, set $g_{nj} = \max_{m\geq n-B_nj+1} |b_m|$. Given (1.2), the resulting shift is by at most a finite number of steps and leaves the asymptotic argument unaffected. With this setup, the weights appearing in S_{nj}^* in (2.56) are bounded by $B_n^{-1/2}$. By arguments paralleling those for S_{ni},

$$\sup_{-\infty<j\leq r_n} \mathrm{E}(S_{nj}^{*2}) = O(1) \quad (2.67)$$

and

$$\sup_{-\infty<j\leq r_n} \sum_{m=1}^{\infty} |\mathrm{E}(S_{nj}^* S_{n,j-m}^*)| = o(1). \quad (2.68)$$

Next define

$$w_{nj}^* = \left(\frac{B_n^3 g_{nj}^2}{\sum_{i=-\infty}^{n} a_{ni}^2} \right)^{1/2}, \ -\infty < \ j \leq r_n.$$

Since $|b_{n-B_nj+1}| \sim (n - B_nj)^{d-1} L(n - B_nj)$ by (1.2) and $n \sim r_n B_n$, the choice in (2.66) implies

$$\sum_{j=-\infty}^{r_n} B_n^3 g_{nj}^2 = O\left(B_n^{2d+1} \sum_{j=-\infty}^{r_n} (r_n - j)^{2d-2} L(n - B_nj)^2 \right) = O(B_n^{2d+1}).$$

Hence by Theorem **2.6**,

$$\sum_{j=-\infty}^{r_n} w_{nj}^{*2} = O\Big(\frac{B_n^{2d+1}}{n^{2d+1}L(n)^2}\Big) = O(r_n^{-2d-1}). \tag{2.69}$$

From (2.57), (2.67), (2.68), and (2.69), the conclusion is that

$$\begin{aligned}\Big(\sum_{i=-\infty}^{n} a_{ni}^2\Big)^{-1} &\mathrm{E}(A_{2n}^2)\\ &= \sum_{j=-\infty}^{r_n} w_{nj}^{*2}\mathrm{E}(S_{nj}^{*2}) + 2\sum_{j=-\infty}^{r_n-1}\sum_{m=1}^{r_n-j} w_{nj}^{*}w_{n,j+m}^{*}\mathrm{E}(S_{nj}^{*}S_{n,j+m}^{*})\\ &= o(1) \end{aligned} \tag{2.70}$$

which in view of Theorem **2.6** confirms $\mathrm{E}(A_{2n}^2) = o(\mathrm{E}(A_{1n}^2))$.

Finally, it is easy to verify in view of Assumption **1.2**(a) that $\mathrm{E}(A_{3n}^2) = O(B_n^{2d+1}) = o(\mathrm{E}(A_{1n}^2))$. Combining this fact with (2.70) and (2.58) completes the proof. ∎

For clarity of exposition this result has been given for the case of $a_{ni}(1,0)$, but the following extension is immediate.

2.11 Corollary Under Assumption **1.2**,

$$\lim_{n\to\infty} \mathrm{E}\Big(\frac{1}{\kappa(n)}\sum_{i=[nt]+1}^{[ns]} x_i^2\Big) = \omega_u^2\Upsilon_d(s-t)^{2d+1}.$$

Proof In equation (2.57), and subsequently, a_{ni} is replaced by $a_{ni}(s,t)$ and n by $[ns]$, with the definitions of B_n and r_n modified to match. The argument of Theorem **2.6** allows equation (2.62) to be replaced by

$$\sum_{j=-\infty}^{[r_n s]} a_{r_n j}^2(s,t) \sim \Upsilon_d r_n^{2d+1}(s-t)^{2d+1}L(r_n(s-t))^2.$$

With these substitutions, the proof of Theorem **2.10** is now replicated almost unchanged. ∎

The next corollary of Theorem **2.10** extends the result to the multivariate case, in the manner of Theorem **2.9**. To replace the contemporaneous covariance matrix $\boldsymbol{\Sigma}_u$ defined in (2.48), the long run covariance matrix may be defined as

$$\boldsymbol{\Omega}_u = \{\omega_{kl}\} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}(\boldsymbol{u}_i\boldsymbol{u}_j') \quad (m\times m) \tag{2.71}$$

which is finite under Assumption **1.2**.

2.12 Corollary If Assumption **1.2** holds for each element of $\boldsymbol{u}_i$,

$$\lim_{n\to\infty} \mathrm{E}(\boldsymbol{X}_n(t+\delta)-\boldsymbol{X}_n(t))(\boldsymbol{X}_n(t+\delta)-\boldsymbol{X}_n(t))' = \boldsymbol{K}(\delta)(\boldsymbol{\Omega}_u \odot \boldsymbol{\Upsilon})\boldsymbol{K}(\delta). \quad (2.72)$$

Proof This is by extension of the arguments of Theorems **2.9** and **2.10**. In place of A_{1n}, A_{2n} and A_{3n} in (2.57), define in the obvious manner the pairs A_{k1n}, A_{l1n}, A_{k2n}, A_{l2n}, and A_{k3n}, A_{l3n}. In (2.59), replace $\mathrm{E}(S_{nj}^2)$ by $\mathrm{E}(S_{knj}S_{lnj})$ and in (2.60), $\mathrm{E}(S_{nj}S_{n,j-m})$ by $\mathrm{E}(S_{knj}S_{ln,j-m})$, with the starred sums redefined in the same manner. In (2.63), $A_{k1n}A_{l1n}$ replaces A_{1n}^2 and $a_{kni}a_{lni}$ replaces a_{ni}^2, with the corresponding substitutions in (2.70). With these substitutions, the logic of the proof of Theorem **2.10** holds for each of the diagonal and off-diagonal elements of the matrix, with $\omega_{kl}\Upsilon_{kl}\delta^{d_k+d_l+1}$ replacing $\omega_u^2\Upsilon_d\delta^{2d+1}$. ∎

Chapter 3

The FCLT for Fractional Processes

Proofs of functional weak convergence of fractional partial sum processes, with fractional Brownian motion X in (2.1) as the limit, were were given originally by Yu. Davydov ([20]) and Murad Taqqu ([68]). Other contributions to this literature include ([26]) and more recently [71] and the relevant chapters of [25] and [5]. FCLTs for the Type II form of the limit process are given in [65], [37], and [45]. The material of the present chapter is based on joint work with Robert de Jong, in [21] and specifically Theorem 3.1 of [15], although the proof presented here is a much expanded and corrected revision of the original, benefitting in particular from invaluable commentary in [39].

What distinguishes the present approach is that it deals with the Type I limit process and also that it provides a version (shown in §3.5) for general nonparametric dependence of the shocks as specified in Assumption **1.2**(b). The other results cited above all depend in one way or another on a linear short-memory structure based on independent shocks. Assumption **1.1** is also invoked here initially, for simplicity and so that a complete proof can be given without too much reliance on cited results. Theorem **3.11** in §3.5, which is the generalization to Assumption **1.2**, depends on material from SLT at key steps of the argument.

3.1 The Main Result

Let X_n be as defined in (2.27) with x_i given by (1.1) and (1.2) with $|d| < \frac{1}{2}$ and u_i satisfying Assumption **1.1**. The following extra assumption for the shock process in relation to d is also made, imposing a further moment restriction when $d < 0$ (the antipersistent case).

3.1 Assumption $\{u_i\}_{i=-\infty}^{\infty}$ is L_r-bounded for $r > \max\{2, 1/(\frac{1}{2} + d)\}$. $\square$

Given the results of Chapter 2, under Assumptions **1.1** and **3.1** a functional central limit theorem can be proved establishing the weak convergence of X_n to the fBM X with parameter d. That the limit process has the covariance function of X is already known, given the agreement between Corollary **2.8** and Theorem **2.1**.

3.2 Theorem If $|d| < \frac{1}{2}$ and Assumptions **1.1** and **3.1** hold, $X_n \to_{\mathrm{d}} X$ where X is the fBM in (2.1). □

The proof of Theorem **3.2** is shown by means of a series of lemmas. As a preliminary, it is illuminating to write an increment of the normalized limit process as the sum of three Brownian functionals. For a constant $N \in \mathbb{N}$, further decomposing (2.3) yields

$$\begin{aligned} X(s) - X(t) = &\int_t^s (s-\xi)^d \mathrm{d}U(\xi) + \int_{-N}^t \big((s-\xi)^d - (t-\xi)^d\big)\mathrm{d}U(\xi) \\ &+ \int_{-\infty}^{-N} \big((s-\xi)^d - (t-\xi)^d\big)\mathrm{d}U(\xi). \end{aligned} \tag{3.1}$$

The first two terms of this decomposition constitute the increment $X^N(s) - X^N(t)$ according to the definition in (2.18). According to Corollary **2.4,** under the assumption on d, by taking N large enough the final term can of (3.1) can be treated as negligible in L_2 norm, as shown in (2.20).

Defining $\kappa(n)$ by (2.26), from (2.21) and (2.24) the finite-n counterparts of these limit expressions are

$$\begin{aligned} X_n(s) - X_n(t) &= \left(\sum_{i=[nt]+1}^{[ns]} + \sum_{i=1-nN}^{[nt]} + \sum_{i=-\infty}^{-nN}\right)\frac{a_{ni}(s,t)}{\kappa(n)}u_i \\ &= R_{1n}(s,t) + R_{2n}(s,t) + R_{3n}(s,t). \end{aligned} \tag{3.2}$$

The objective is now to show that the increment $X^N(s) - X^N(t)$ represents the weak limit of the first two terms of (3.2). Let the sample counterpart be denoted by

$$X_n^N(t) = R_{1n}(t,0) + R_{2n}(t,0) \tag{3.3}$$

so that $R_{3n}(t,0) = X_n(t) - X_n^N(t)$. The status of this latter term is established similarly to Corollary **2.4**.

3.3 Lemma Under Assumption **1.1** and with $|d| < \frac{1}{2}$, $\lim_{n\to\infty} \mathrm{E}(R_{3n}(s,t)^2) = O(N^{2d-1})$ as $N \to \infty$.

Proof Following the approach of Theorem **2.6**, similarly to (2.40) and (2.20),

$$\begin{aligned} \mathrm{E}(R_{3n}(s,t))^2 &= \frac{\sigma_u^2}{\kappa(n)^2}\sum_{i=-\infty}^{-nN} a_{ni}^2(s,t) \\ &\sim \sigma_u^2(s-t)^{d+1}\int_N^\infty ((1+\tau)^d - \tau^d)^2 \mathrm{d}\tau. \end{aligned} \tag{3.4}$$

Noting that the integral in (3.4) is $\Upsilon_d - \Upsilon_d^N$, the conclusion follows by (2.20). ∎

This result distinguishes two modes of convergence as n and N respectively increase, these convergences being treated as strictly sequential. While it could be proved that $\mathrm{E}(R_{3n}(1,0))^2 \to 0$ as $N \to \infty$ for any finite value of n, it is only necessary to show this for the limit case. In the usual way with the functional convergence of partial sums, increasing n represents the compression of a growing set of observations into a finite interval. For any given N, the discrete collection of points approaches a continuum as the gaps separating them shrink like $1/n$, leading to the asymptotic equivalence indicated in (3.4). By contrast, the divergence of N represents the inclusion of ever more remote lags into the sum. The domain of the function $R_{2n}(s,t)$ grows with N although more slowly than the number $nN + [nt]$ of corresponding data points.

Theorem **2.6** leads to the conclusions that

$$\mathrm{E}(R_{1n}(s,t))^2 = \frac{\sigma_u^2}{\kappa(n)^2} \sum_{i=[nt]+1}^{[ns]} a_{ni}^2(s,t) \to \frac{\sigma_u^2(s-t)^{2d+1}}{2d+1} \tag{3.5}$$

and

$$\begin{aligned}\mathrm{E}(R_{2n}(s,t))^2 &= \frac{\sigma_u^2}{\kappa(n)^2} \sum_{i=1-nN}^{[nt]} a_{ni}^2(s,t) \\ &\to \sigma_u^2(s-t)^{2d+1} \int_0^N ((1+\tau)^d - \tau^d)^2 \mathrm{d}\tau \end{aligned} \tag{3.6}$$

as $n \to \infty$. Putting these limits together, it is possible to write

$$\mathrm{E}(X^N(t)^2) = \sigma_u^2 \Upsilon_d^N t^{2d+1} \tag{3.7}$$

where Υ_d^N is defined in (2.19) and $\Upsilon_d - \Upsilon_d^N = O(N^{2d-1})$ according to (2.4) and (2.20). The procedure at this point is to derive the limiting distributions of the increments $X_n^N(s) - X_n^N(t) = R_{1n}(s,t) + R_{2n}(s,t)$ where the approximation to (3.2) is controlled by the choice of N.

3.2 Finite Dimensional Distributions

For the results of this section, Assumption **3.1** is not required.

3.4 Lemma Under Assumption **1.1**,

$$R_{1n}(s,t) + R_{2n}(s,t) \xrightarrow{\mathrm{d}} \mathrm{N}\big(0, \sigma_u^2 \Upsilon_d^N (s-t)^{2d+1}\big)$$

for any s and t with $0 \le t < s \le 1$.

Proof The terms $R_{1n}(s,t)$ and $R_{2n}(s,t)$ are shown to have Gaussian limits individually and since these pairs have no time periods in common, they are independent of each other by assumption so that their limiting sum is also Gaussian.

The limiting variance follows directly from (3.5) and (3.6). The terms of the sums $R_{1n}(s,t)$ and of $R_{2n}(s,t)$ are also independent of each other under Assumption **1.1**, but are heterogeneously distributed, with moving average weights $a_{ni}(s,t)/\kappa(n)$. The proof is given for the case $t = 0$ and $s = 1$, but according to Theorem **2.5** the generalization is direct.

Under Assumption **1.1** the Lindeberg condition is sufficient for the CLT to hold[1] and for $R_{1n}(1,0)$, this has the form that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{\kappa(n)^2} \sum_{i=1}^{n} a_{ni}^2 \mathrm{E}\big(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}}\big) = 0. \tag{3.8}$$

To check this condition, consider

$$\begin{aligned}\frac{1}{\kappa(n)^2} \sum_{i=1}^{n} a_{ni}^2 \mathrm{E}\big(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}}\big)&\\ \le \max_{1\le i\le n} \mathrm{E}\big(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}}\big) \frac{1}{\kappa(n)^2} \sum_{i=1}^{n} a_{ni}^2.&\end{aligned} \tag{3.9}$$

It follows by Theorem **2.6** and Corollary **2.8** that $\kappa(n)^{-2}\sum_{i=1}^{n} a_{ni}^2 = O(1)$. Therefore, to bound the majorant of (3.9) apply Theorem **A.4** in Appendix A with $\eta = \varepsilon\kappa(n)/|a_{ni}|$. There are two cases to be considered. In the case $d > 0$, Theorem **2.5** gives $\max_{1\le i\le n}|a_{ni}|/\kappa(n) = O(n^{-1/2})$, not overlooking that in ratios of this sort the slowly varying components, if any, cancel in the limit. According to Theorem **A.4**, the bound on (3.9) therefore has the form

$$\max_{1\le i\le n} \mathrm{E}\big(u_i^2 1_{\{|u_i|>\varepsilon\kappa(n)/|a_{ni}|\}}\big) = o\Big(\max_{1\le i\le n} (|a_{ni}|/\kappa(n))^{r-2}\Big) = o(n^{1-r/2}) \tag{3.10}$$

which vanishes when $r \ge 2$. If $d < 0$, Theorem **2.5** says that $\max_{1\le i\le n}|a_{ni}| = 1$, and so $\max_{1\le i\le n}|a_{ni}|/\kappa(n) = O(n^{-d-1/2})$. In this case Theorem **A.4** gives

$$\max_{1\le i\le n} \mathrm{E}(u_i^2 1_{\{|u_i|>\varepsilon\kappa(n)/|a_{ni}|\}}) = o(n^{(d+1/2)(2-r)}). \tag{3.11}$$

Therefore, provided $d > -\frac{1}{2}$ the Lindeberg condition is satisfied in either case when $r \ge 2$, although when $d < 0$ the convergence is correspondingly slower.

In the case of $R_{2n}(1,0)$, the Lindeberg condition assumes the form

$$\lim_{n\to\infty} \frac{1}{\kappa(n)^2} \sum_{i=1-nN}^{0} a_{ni}^2 \mathrm{E}(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}}) = 0 \tag{3.12}$$

and Theorem **2.5** and Corollary **2.8** give

$$\frac{|a_{ni}|}{\kappa(n)} \sim \frac{|(n-i)^d - (-i)^d|}{n^{d+1/2}} \tag{3.13}$$

[1] See SLT Theorem 24.6.

which is decreasing in $-i$ for both $d > 0$ and $d < 0$. A convenient way to handle the Lindeberg condition test is to break up the terms of the sum into N blocks of length n, itemized by $k = 0, \ldots, N-1$. Consider initially the case $k = 0$. The first n terms of $R_{2n}(1,0)$ contribute the sum

$$\frac{1}{\kappa(n)^2}\sum_{i=1-n}^{0} a_{ni}^2 \mathrm{E}(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}})$$
$$\leq \max_{1-n\leq i\leq 0} \mathrm{E}(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}} \sum_{i=1-n}^{0} \frac{a_{ni}^2}{\kappa(n)^2}$$

where (3.13) implies

$$\frac{1}{\kappa(n)^2}\sum_{i=1-n}^{0} a_{ni}^2 \sim \int_0^1 ((\tau+1)^d - \tau^d)^2 \mathrm{d}\tau.$$

Similarly to (3.10) and (3.11), (3.13) further implies

$$\max_{1-n\leq i\leq 0} \mathrm{E}(u_i^2 1_{\{|u_i|>\varepsilon\kappa(n)/|a_{ni}|\}}) = \begin{cases} o(n^{(1-r/2)}), & d > 0 \\ o(n^{(d+1/2)(2-r)}), & d < 0. \end{cases} \tag{3.14}$$

For the cases $k \geq 1$, the maximum value of $|a_{ni}|$ over the range $1 - n(k+1) \leq i \leq -nk$ is found when n is large enough at $i = -nk$, for d of either sign. According to (3.13), $|a_{n,-nk}| \sim n^d|(k+1)^d - k^d|$ and after simplification it is found that

$$\begin{aligned}
&\sum_{k=1}^{N-1}\sum_{i=1-n(k+1)}^{-nk} \frac{a_{ni}^2}{\kappa(n)^2}\mathrm{E}(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}}) \\
&\qquad \leq n\sum_{k=1}^{N-1} \max_{-n(k+1)<i\leq -nk} \frac{a_{ni}^2}{\kappa(n)^2}\mathrm{E}(u_i^2 1_{\{|a_{ni}u_i|/\kappa(n)>\varepsilon\}}) \\
&\qquad \ll n\sum_{k=1}^{N-1} \frac{n^{2d}|(k+1)^d - k^d|^2}{n^{2d+1}}\left(\frac{n^{d+1/2}}{n^d|(k+1)^d - k^d|}\right)^{2-r} \\
&\qquad = n^{1-r/2}\sum_{k=1}^{N-1} |(k+1)^d - k^d|^r.
\end{aligned} \tag{3.15}$$

The sum in the majorant of (3.15) is bounded for arbitrary N by $\int_1^\infty((\tau+1)^d - \tau^d)^2\mathrm{d}\tau$, which with $|d| < \frac{1}{2}$ is finite by Lemma **2.3**. Putting together the convergences of (3.15) and (3.14) shows that (3.12) holds so that $R_{2n}(1,0)$ also satisfies the Lindeberg condition.

Replacing $a_{ni}(1,0)$ by $a_{ni}(s,t)$, and the sum limit n by $[ns]$ in (3.8) and (3.12) extends these results with straightforward modification of the formulae to $R_{1n}(s,t)$ and $R_{2n}(s,t)$, completing the proof. ∎

3.3 Uniform Boundedness and Uniform Integrability

The hypothesized limit process (2.1) is almost surely continuous. It remains to show that the sequence of empirical distributions defined by $\{X_n^N, n \in \mathbb{N}\}$, where X_n^N is defined in (3.3), has a limit that is almost surely continuous likewise. The key concept here, that of uniform tightness of the sequence of distributions, is reviewed in §3.4. First some fundamental limit properties have to be established, specifically, that process increments of a fixed width δ are uniformly bounded and uniformly square-integrable. The tightness analysis then considers what happens as δ shrinks to zero.

Consider an increment $X_n^N(t+\delta) - X_n^N(t)$ for $\delta > 0$ and $0 \le t \le 1-\delta$. For $s \in (t, t+\delta]$ define

$$T_n(s,t) = X_n^N(s) - X_n^N(t) = R_{1n}(s,t) + R_{2n}(s,t) \tag{3.16}$$

and decompose the variance $\mathrm{E}(T_n(t+\delta,t)^2) = \nu_n^2(t,\delta)$ correspondingly, where

$$\begin{aligned}\nu_n^2(t,\delta) &= \frac{\sigma_u^2}{\kappa(n)^2}\Bigg(\sum_{i=[nt]+1}^{[n(t+\delta)]} + \sum_{i=1-nN}^{[nt]}\Bigg) a_{ni}^2(t+\delta,t) \\ &= \nu_{1n}^2(t,\delta) + \nu_{2n}^2(t,\delta). \end{aligned} \tag{3.17}$$

Under Assumption **1.1**, $\nu_{1n}^2(t,\delta)$ and $\nu_{2n}^2(t,\delta)$ are the variances of $R_{1n}(t+\delta,t)$ and $R_{2n}(t+\delta,t)$ respectively and in view of (2.35) and (2.26) it is evident that $\nu_n^2(t,\delta) = O(1)$.

It is helpful to set up a compact notation for the objects under study. For $\delta > 0$ and $0 \le t \le 1-\delta$, define

$$\widetilde{T}_n(t,\delta) = \sup_{\{s:|s-t|<\delta\}} \frac{|T_n(s,t)|}{\nu_n(t,\delta)}. \tag{3.18}$$

The notations $\widetilde{R}_{1n}(t,\delta)$, $\widetilde{R}_{2n}(t,\delta)$, and so forth are defined likewise for the processes in question. In each case, the supremum over s of the sum's absolute value is normalized by the standard deviation of the increment.

The squared supremum of the absolute value is the same thing as the supremum of the square. For given t and δ, the uniform square-integrability of (3.18) is the condition that for some $n_0 < \infty$,

$$\sup_{n>n_0} \mathrm{E}\big(\widetilde{T}_n^2(t,\delta) 1_{\{\widetilde{T}_n(t,\delta)>B\}}\big) \to 0 \text{ as } B \to \infty. \tag{3.19}$$

In fact a stronger condition can be shown, specifying the rate of convergence of the expectations to zero, and this extension will be needed in the sequel. While $n_0 = 1$ may generally be a reasonable assumption, it is possible for the required conditions to be imposed on the limiting case but left unspecified in finite samples.

Uniform integrability cannot hold without uniform boundedness in probability and the following result is therefore informative, for while Assumption **3.1** is shown in Theorem **3.8** to be sufficient for uniform tightness, the fact of its necessity in the antipersistent case is an important feature of the fractional asymptotics. The following necessity proof adapts a counter-example in [72], cited by [39] in commenting on [15].

3.5 Theorem Under Assumption **1.1**, for $\delta > 0$ and $0 \le t \le 1-\delta$ the collection $\{\widetilde{T}_n^2(t,\delta),\ n \in \mathbb{N}\}$ is uniformly bounded in probability if and only if Assumption **3.1** holds.

Proof If u_i is L_r-bounded then according to Lemma **A.3**, $P(|u_i| > \eta) = O(\eta^{-r}\log(\eta)^{-1-\mu})$ as $\eta \to \infty$ for $\mu > 0$. If the u_i are independently and identically distributed then for any $\varepsilon > 0$,

$$
\begin{aligned}
P\biggl(\sup_{t\le s\le t+\delta}\frac{|u_{[ns]}|}{\kappa(n)} \le \varepsilon\biggr) &= P\biggl(\bigcap_{i=[nt]+1}^{[n(t+\delta)]}\{|u_i| \le \varepsilon\kappa(n)\}\biggr)\\
&= \bigl(1 - P(|u_i| > \varepsilon\kappa(n))\bigr)^{[n(t+\delta)]-[nt]}\\
&= O\biggl(\Bigl(1 - \frac{1}{\varepsilon^r\kappa(n)^r\log(\varepsilon\kappa(n))^{1+\mu}}\Bigr)^{n\delta}\biggr)\\
&= O\biggl(\exp\Bigl\{-\frac{n\kappa(n)^{-r}}{\varepsilon^r\log(\varepsilon\kappa(n))^{1+\mu}}\Bigr\}^{\delta}\biggr). \qquad (3.20)
\end{aligned}
$$

If $n\kappa(n)^{-r} \to \infty$ then this probability converges to zero as $n \to \infty$ for all $\varepsilon > 0$ and the supremum over $[nt] \le i \le [n(t+\delta)]$ of $|u_i|/\kappa(n)$ accordingly diverges with probability 1.

According to (2.26) the condition $n\kappa(n)^{-r} = O(1)$ holds by Assumption **3.1**. The proof is completed by showing that this condition is both necessary and sufficient for $\sup_{t\le s\le t+\delta}|T_n(s,t)|$ to be bounded in probability. According to (2.25), the partial sum from $[nt]+1$ to $[ns]$ has the form

$$
\begin{aligned}
T_n(s,t) &= \sum_{i=1-nN}^{[ns]}\frac{a_{ni}(s,t)}{\kappa(n)}u_i\\
&= \frac{b_0}{\kappa(n)}u_{[ns]} + \frac{b_0+b_1}{\kappa(n)}u_{[ns]-1} + \frac{b_0+b_1+b_2}{\kappa(n)}u_{[ns]-2} + \cdots \qquad (3.21)
\end{aligned}
$$

where $b_0 = 1$ and if $d < 0$ the lag coefficients in (3.21) are converging to 0 as in (1.20). To show necessity, rearrange (3.21) by pulling out the leading term, taking absolute values, applying the triangle inequality, and then taking the sup of both sides, to give

$$
\sup_{t\le s\le t+\delta}\frac{|u_{[ns]}|}{\kappa(n)} \le \sup_{t\le s\le t+\delta}|T_n(s,t)| + \sup_{t\le s\le t+\delta}\biggl|\sum_{i=1-nN}^{[ns]-1}\frac{a_{ni}(s,t)}{\kappa(n)}u_i\biggr|. \qquad (3.22)
$$

The shocks are independent and the second majorant term does not contain $u_{[ns]}$. Since $\nu_n(t,\delta) = O(1)$, the divergence of $\widetilde{T}_n^2(t,\delta)$ with probability 1 must follow from the divergence of the minorant of (3.22).

To show sufficiency, consider relation (3.20) with $t = -N$ and δ set to $N + 1$. Under Assumption **3.1**, this has the implication that as $\varepsilon \to \infty$,

$$-\log P\Big(\sup_{-N\le s\le 1} \frac{|u_{[ns]}|}{\kappa(n)} \le \varepsilon\Big) = O(\varepsilon^{-r}\log(\varepsilon\kappa(n))^{-1-\mu}). \tag{3.23}$$

The probability itself therefore tends to 1. By definition the bound in question applies to every $u_i/\kappa(n)$ in the sum (3.21) and it follows that $|T_n(s,t)| = O_p(1)$ under Assumptions **1.1** and **3.1**. This is true for any s and in particular for $\sup_{t\le s\le t+\delta}|T_n(s,t)|$. The same order of magnitude extends to (3.18) and, as noted, the supremum of the square is the square of the supped absolute value. ∎

To see the implications of this result for (3.19), rearrange the Chebyshev inequality as

$$\mathrm{E}(\widetilde{T}_n^2(t,\delta)) > \varepsilon^2\big(1 - P(\widetilde{T}_n(t,\delta) \le \varepsilon)\big). \tag{3.24}$$

Unless Assumption **3.1** holds, (3.22) implies that if $d < 0$, with large enough n the probability in the minorant of (3.24) must shrink so far that $\mathrm{E}\big(\widetilde{T}_n^2(t,\delta)\big) \ge \varepsilon^2$. Since as $n \to \infty$ this remains true even as ε is taken arbitrarily large, (3.19) is ruled out.

The natural route to demonstrating that (3.19) holds is by verifying a uniform bound for squared partial sums, such as Theorem **A.5** in Appendix A. However, the argument is not straightforward due to the fact that the moving average weights appearing in $T_n(s,t)$ are not merely an array of constants but depend on s, over which the supremum is to be taken. Therefore the collection $\{\widetilde{T}_n^2(t,\delta),\ n \in \mathbb{N}\}$ does not meet the specified conditions of Theorem **A.5**. What can be shown is that this sequence is dominated by sequences that do meet the conditions. The key feature of Theorem **A.5** in this context is that when the summands are independently distributed the Doob inequality, which is a result for martingales, holds irrespective of their ordering. The following lemma develops a roundabout application of the theorem to obtain the required result.

3.6 Lemma Under Assumptions **1.1** and **3.1**, for all $0 < \delta < 1$ and $t \in [0, 1-\delta]$ the collection $\{\widetilde{T}_n^2(t,\delta), n \ge n_0\}$ for $n_0 < \infty$ is uniformly integrable with

$$\mathrm{E}\big(\widetilde{T}_n^2(t,\delta)1_{\{\widetilde{T}_n(t,\delta)>B\}}\big) = o(B^{2-r}). \tag{3.25}$$

Proof Consider the sum (3.16). Taking absolute values and applying the triangle inequality, also noting $\nu_n(t,\delta) \ge \nu_{1n}(t,\delta)$ and $\nu_n(t,\delta) \ge \nu_{2n}(t,\delta)$, gives the relation

$$\frac{|T_n(s,t)|}{\nu_n(t,\delta)} \le \frac{|R_{1n}(s,t)|}{\nu_{1n}(t,\delta)} + \frac{|R_{2n}(s,t)|}{\nu_{2n}(t,\delta)}. \tag{3.26}$$

Taking the sup over s of each term of (3.26) and then squaring both sides leads to the inequality

$$\widetilde{T}_n^2(t,\delta) \le \Big(\widetilde{R}_{1n}(t,\delta) + \widetilde{R}_{2n}(t,\delta)\Big)^2 \tag{3.27}$$

which holds with probability 1. According to Theorems **A.7** and **A.6**, to prove the lemma it is sufficient to show that the collections $\{\widetilde{R}_{1n}^2(t,\delta),\ n \ge n_0\}$ and

$\{\widetilde{R}_{2n}^2(t,\delta), n \geq n_0\}$ are uniformly integrable, with order of magnitude $o(B^{2-r})$ as $B \to \infty$ similarly to (3.25).

Consider the moving average coefficients of $R_{1n}(s,t)$. According to (2.25) these depend on s, having the form $a_{ni}(s,t) = b_0 + \cdots + b_k$, where $k = [ns] - i$ with i running from $[nt]+1$ to $[ns]$ and hence k running from 0 to $[ns]-[nt]-1$. The complete set of these coefficients, by descending order of i, is

$$\left\{\frac{b_0}{\kappa(n)}, \frac{b_0+b_1}{\kappa(n)}, \frac{b_0+b_1+b_2}{\kappa(n)}, \dots, \frac{b_0+\cdots+b_{[n(t+\delta)]-[nt]-1}}{\kappa(n)}\right\}. \tag{3.28}$$

For any $s \in (t, t+\delta]$ the coefficients appearing in $R_{1n}(s,t)$ belong to the set (3.28), being the first $[ns]-[nt]$ members of the indicated sequence, but in reverse order (increasing i), so that the successive partial sums as s increases are not a simple cumulation of terms. This is the problem with the application of Theorem **A.5**.

Therefore, consider for $s \in (t, t+\delta]$ the modified array

$$R_{1n\delta}^*(s,t) = \frac{1}{\kappa(n)} \sum_{i=[nt]+1}^{[ns]} a_{ni}(t+\delta,t)u_i \tag{3.29}$$

in which the moving average coefficients depend on fixed t and δ, but not on s. While $R_{1n\delta}^*(t+\delta,t) = R_{1n}(t+\delta,t)$, for other values of s the coefficients in (3.29) are also drawn from (3.28) in decreasing order as i increases, but now starting from the end. Define

$$c_{ni} = \frac{|a_{ni}(t+\delta,t)|}{\kappa(n)\nu_{1n}(t,\delta)}$$

so that $\sum_{i=[nt]+1}^{[n(t+\delta)]} c_{ni}^2 = 1$ by (3.17) and

$$\frac{R_{1n\delta}^*(s,t)}{\nu_{1n}(t,\delta)} = \sum_{i=[nt]+1}^{[ns]} c_{ni}u_i. \tag{3.30}$$

This is the cumulation of the first $[ns]-[nt]$ terms of a sequence with fixed coefficients. It follows by Theorem **A.5** that, on the assumptions, the collection $\{\widetilde{R}_{1n\delta}^*(t,\delta)^2, n \in \mathbb{N}\}$ is uniformly integrable with the rate of convergence in (3.25).

This property does not depend on the ordering of the moving average weights in the sum. By assumption the shock process is stationary and Theorem **A.5** continues to apply under arbitrary permutations of the c_{ni} in (3.30). Let $R_{1np}^*(s,t)$ denote the partial sum corresponding to (3.30) under such a permutation, noting that $v_{1n}^2(t,\delta)$ is the sum of the squared elements of (3.28) and is unchanged. In particular, a permutation p^* can be chosen such that, with probability 1,

$$\sup_{\{s:|s-t|<\delta\}} R_{1n}(s,t)^2 \leq \sup_{\{s:|s-t|<\delta\}} R_{1np^*}^*(s,t)^2. \tag{3.31}$$

To see that such a p^* exists, note how the sup in the minorant of (3.31) is constructed. For each s the sum of the terms $\{a_{ni}(s,t)u_i,\ i = [nt]+1,\dots,[ns]\}$ is formed and then s is chosen to maximize the square of this sum. By contrast, on the majorant side of (3.31) the partial sums are formed by picking coefficients in the chosen order p^* from the set (3.28). This set includes all those in the minorant and equality in (3.31) is always achieved by having the two sets of coefficients matching. Free choice of p^* can never do worse.

Let $\{p_n^*\}$ denote a sequence of permutations that satisfy inequality (3.31), for each n. By construction, each element of the collection $\{\widetilde{R}_{1np_n^*}^*(t,\delta)^2, n \in \mathbb{N}\}$ is drawn from a sequence with the property shown in (3.25) and hence shares that property. In view of (3.31) and Theorem **A.6**, the same property applies to the collection $\{\widetilde{R}_{1n}(t,\delta)^2, n \in \mathbb{N}\}$.

Next, consider $R_{2n}(s,t)$ for $t \leq s \leq t+\delta$. These sums all contain $[nt]+nN$ terms, but they differ because the moving average coefficients are sums of $[ns]-[nt]$ terms according to (2.25). Thus, for each $i = 1-nN,\ldots,[nt]$, $a_{ni}(s,t) = b_{[nt]+1-i} + \cdots + b_{[ns]-i}$. The second term of (3.17), $\nu_{2n}^2(t,\delta)$, is the sum of the coefficients $a_{ni}^2(t+\delta,t)/\kappa(n)^2$. The sum of the terms $a_{ni}^2(s,t)/\kappa(n)^2$ may similarly be written $\nu_{2n}^2(t,s-t)$. If the weights c_{ni} are defined for given s by the equality

$$\frac{R_{2n}(s,t)}{\nu_{2n}(t,\delta)} = \sum_{i=1-nN}^{[nt]} c_{ni}u_i \tag{3.32}$$

then for this same s,

$$\sum_{i=1-nN}^{[nt]} c_{ni}^2 = \frac{\nu_{2n}^2(t,s-t)}{\nu_{2n}^2(t,\delta)}. \tag{3.33}$$

If the ratio in (3.33) is $O(1)$ as $n \to \infty$, the sums (3.32) satisfy the conditions of Theorem **A.5**. Since the maxima of the squared partial sums are thereby uniformly integrable, this is certainly also true of the complete sums, according to Theorem **A.6**. This is true for any choice of s in the closed interval $[t,t+\delta]$ and hence, in particular, for the collection $\{\widetilde{R}_{2n}^2(t,\delta), n \in \mathbb{N}\}$.

It therefore remains to show that the ratio (3.33) is indeed $O(1)$. The sums $\nu_{2n}^2(t,s-t)$ and $v_{2n}^2(t,\delta)$ contain the same number of terms, respectively the terms $a_{ni}^2(s,t)$ and $a_{ni}^2(t+\delta,t)$ defined by (2.25) and with large enough n, $a_{ni}^2(s,t) \leq a_{ni}^2(t+\delta,t)$ for every i. This is because the coefficients b_j have the sign of d at long range according to (1.2) and hence sums with more terms (larger s for given i) are absolutely larger. This is true even with $d<0$, since $i \leq [nt]$ and hence b_0 is always excluded. Thus, there exists n_0 large enough that $\nu_{2n}^2(t,s-t) \leq \nu_{2n}^2(t,\delta)$ for $n \geq n_0$. ∎

n_0 is left unspecified in this result only because (1.2) leaves the finite-order lag coefficients unspecified. In the case of (1.12), to take the leading example, the coefficients are a monotone sequence and n_0 can be set to 1. Note however that without Assumption **3.1**, with $d<0$ the expectations in (A.11) are undefined in the limit as $n \to \infty$, as demonstrated by (3.24), so that the arguments based on Theorem **A.5** fail. The assumption is accordingly necessary.

3.4 Uniform Tightness

A tight distribution is one having the property that there exists a compact subset of the sample space having probability arbitrarily close to 1, so that the random outcomes do not ‘escape to infinity’. Uniform tightness is the property of a sequence of tight distributions, that tightness is preserved in the limit. A familiar counter-example is the sequence of uniform distributions on the intervals $[-n, n]$ of $\mathbb{R}$, which are well defined for every finite n but not in the limit.

The empirical process X_n defined by (2.27) is an element of the space $D_{[0,1]}$ of càdlàg functions on the unit interval, in other words, a step function. As the sample size n increases, the width of the steps shrinks like $1/n$ while the heights of the jumps are diminishing in line with the standard deviation of the increments, at the rate $n^{-1/2-d}$. Notwithstanding the Gaussianity, to match fBM in (2.1) the limit distribution of the process must be confined to the space $C_{[0,1]}$ with probability 1. This property of a sequence of random empirical processes is called stochastic equicontinuity, A formal definition and additional discussion of this concept can be found in, for example, §22.3 of SLT.

A step function can be transformed into a continuous function by the simple expedient of joining the vertices with straight lines. Hence, it is also possible to think in terms of a sequence of distributions with domain $C_{[0,1]}$ and the issue is then whether the limit process is also in $C_{[0,1]}$, with probability 1. If some lines do not shrink in length to zero in the limit, but instead become vertical jumps as the step width goes to zero, processes under the limit distribution are said to escape from $C_{[0,1]}$. Uniform tightness fails if this occurrence has positive probability. By analogy, it is also commonly said that a sequence with domain $D_{[0,1]}$ is uniformly tight if the limit is in $C_{[0,1]}$ almost surely, which is the property required in the present case.

To allow additional applications, specifically those arising in Chapter 6, the domain of the empirical processes is taken to be an interval $[L, U]$ that is finite and includes the origin, but otherwise is unspecified. However, for the present purpose of proving Theorem **3.2** the bounds are set to $L = 0$ and $U = 1$. For a random process $Y_n \in D_{[L,U]}$, the so-called modulus of continuity is defined as

$$w_n(\delta) = \sup_{t\in[L,U]} \sup_{\{s:0<|s-t|<\delta\}} |Y_n(s) - Y_n(t)|. \tag{3.34}$$

If $w_n(\delta) \to 0$ as $\delta \to 0$, the function Y_n must be continuous in its argument. If stochastic equicontinuity holds, the probability of this event can be made as close to one as desired by taking n large enough and uniform tightness of the sequence can be defined in these terms. Sufficient conditions are proved as Theorem 30.19 of SLT, which is itself adapted from Theorem 15.5 of [6]. Adapted to the present applications and omitting some topological detail, this theorem can be cast in the following form.

3.7 Theorem Let $\{\mu_n\}$ denote a sequence of probability measures on $D_{[L,U]}$ where $U - L < \infty$. If for all $n \geq n_0$ with $n_0 < \infty$,

(a) for $\eta > 0$ there exists $M < \infty$ such that $\mu_n(x : \sup_t |x(t)| > M) \leq \eta$

(b) for $\varepsilon > 0$ and $\eta > 0$ there exists $\delta > 0$ such that $\mu_n(w_n(\delta) \geq \varepsilon) \leq \eta$

then $\{\mu_n\}$ is uniformly tight and any cluster point has $\mu_n(C_{[L,U]}) = 1$. $\square$

SLT Theorem 30.19 treats the case $L = 0$ and $U = 1$ and also replaces condition **3.7**(a) with the condition of boundedness in probability at a point of the interval, which is sufficient given condition **3.7**(b) although condition **3.7**(a) is convenient for the present case since it has already been shown to hold by Theorem **3.5**.

The next result gives conditions for stochastic equicontinuity by way of validating condition **3.7**(b). The familiar P is here interpreted to denote μ_n when the argument is a random element bearing subscript n.

3.8 Theorem For a random process $Y_n \in D_{[L,U]}$ with $U - L < \infty$, define

$$\widetilde{T}_n(t,\delta) = \sup_{\{s:|s-t|<\delta\}} \frac{|Y_n(s) - Y_n(t)|}{\nu_n(t,\delta)} \tag{3.35}$$

where $\nu_n(t,\delta)^2 = \mathrm{E}(Y_n(t+\delta) - Y_n(t))^2$. If $|d| < \frac{1}{2}$ and for each $\delta > 0$ and $t \in [L, U-\delta]$,

(a) $\nu_n^2(t,\delta) = O(\delta^{\min\{1,2d+1\}})$ as $n \to \infty$

(b) $\mathrm{E}\big(\widetilde{T}_n^2(t,\delta)1_{\{\widetilde{T}_n(t,\delta)>B\}}\big) = o(B^{2-r})$ for $n \geq n_0$ where $n_0 < \infty$ and $r \geq \max\{2, 1/(d+\frac{1}{2})\}$

then, for $\varepsilon > 0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty} P(w_n(\delta) \geq \varepsilon) = 0 \tag{3.36}$$

where $w_n(\delta)$ is defined in (3.34).

Proof Define the set of integers $J(\delta) = \{[L/\delta]+1, [L/\delta]+2, \ldots, [L/\delta]+[(U-L)/\delta]\}$ so that the collection $\{j\delta \in [L,U] : j \in J(\delta)\}$ defines a set of $[(U-L)/\delta]$ points of the interval within a distance δ of their nearest neighbours. For $j \in J(\delta)$, let

$$w_{nj}(\delta) = \sup_{\{s:|s-j\delta|<2\delta\}} |Y_n(s) - Y_n(j\delta)|. \tag{3.37}$$

For any pair s, t such that $L \leq t < s < t+\delta \leq U$, there exists $j \in J(\delta)$ such that $|t-j\delta| < 2\delta$ and $|s-j\delta| < 2\delta$. Noting that $\{|x+y| > \varepsilon\} \subseteq \{|x| > \frac{1}{2}\varepsilon\} \cup \{|y| > \frac{1}{2}\varepsilon\}$ for variables x and y, subadditivity implies that

$$\begin{aligned} P(w_n(\delta) > \varepsilon) &= P\Big(\sup_{t\in[L,U]}\sup_{\{s:|s-t|<\delta\}} |Y_n(s) - Y_n(j\delta) + Y_n(j\delta) - Y_n(t)| > \varepsilon\Big) \\ &\leq 2P\Big(\max_{j\in J(\delta)} w_{nj}(\delta) > \tfrac{1}{2}\varepsilon\Big). \end{aligned} \tag{3.38}$$

Noting that $w > M$ for $M > 0$ if and only if $w1_{\{w>M\}} > M$, subadditivity and then the Markov inequality also give

$$
\begin{aligned}
P\Big(\max_{j\in J(\delta)} w_{nj}(\delta) > \tfrac{1}{2}\varepsilon\Big) &\le \sum_{j\in J(\delta)} P\big(w_{nj}^2(\delta) > \tfrac{1}{4}\varepsilon^2\big) \\
&\le \frac{4}{\varepsilon^2} \sum_{j\in J(\delta)} \mathrm{E}\big(w_{nj}^2(\delta) 1_{\{w_{nj}^2(\delta)>\varepsilon^2/4\}}\big).
\end{aligned} \tag{3.39}
$$

For brevity define $\bar{\nu}_{nj} = \nu_n(j\delta, 2\delta)$, dependence on δ being understood, and similarly define

$$\bar{w}_{nj} = \frac{w_{nj}(\delta)}{\bar{\nu}_{nj}}$$

so that $\bar{w}_{nj} = \widetilde{T}_n(j\delta, 2\delta)$ according to (3.35) and (3.37). In view of condition (a), $\bar{\nu}_{nj}^2 = O((2\delta)^{\min\{1,2d+1\}})$ as $\delta \to 0$ and hence, since the set $j(\delta)$ has $[(U-L)/\delta]$ elements, $\sum_{j\in J(\delta)} \bar{\nu}_{nj}^2 \ll 2(U-L)\delta^{\min\{0,2d\}}$ when n is large enough. Hence there exists $n_0 < \infty$ such that for $n \ge n_0$,

$$
\begin{aligned}
\sum_{j\in J(\delta)} \mathrm{E}\big(w_{nj}^2 1_{\{w_{nj}^2(\delta)>\varepsilon^2/4\}}\big) &= \sum_{j\in J(\delta)} \bar{\nu}_{nj}^2 \mathrm{E}\big(\bar{w}_{nj}^2 1_{\{\bar{w}_{nj}^2>\varepsilon^2/4\bar{\nu}_{nj}^2\}}\big) \\
&\ll 2(U-L)\delta^{\min\{0,2d\}} \max_{j\in J(\delta)} \mathrm{E}\big(\bar{w}_{nj}^2 1_{\{\bar{w}_{nj}^2>\varepsilon^2/4\max_{k\in J(\delta)}\bar{\nu}_{nk}^2\}}\big).
\end{aligned} \tag{3.40}
$$

Passing back up the chain of inequalities (3.38), (3.39), and (3.40), the limsup as $n \to \infty$ of the probability in the minorant of (3.38) goes to zero if the same is true of the majorant of (3.40). This happens as $\delta \to 0$, confirming (3.36), if two conditions are met. The first is that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{t\in[0,1-\delta]} \nu_n^2(t,\delta) = 0 \tag{3.41}$$

which follows directly from condition (a) since $d > -\frac{1}{2}$. In this case the argument of the indicator function in the majorant of (3.40) diverges, as $\delta \to 0$. The second condition is that, for any $n \ge n_0$, the majorant of (3.40) vanishes as a result. However, according to condition (a) in conjunction with Lemma **3.6** and condition (b),

$$\delta^{\min\{0,2d\}} \max_{j\in J(\delta)} \mathrm{E}\big(\bar{w}_{nj}^2 1_{\{|\bar{w}_{nj}|>\varepsilon/2\max_{k\in J(\delta)}\bar{\nu}_{nk}\}}\big) = o\big(\delta^{\min\{0,2d\}-(1/2+d)(2-r)}\big) \tag{3.42}$$

as $\delta \to 0$. Since $\min\{0, 2d\} - (\frac{1}{2} + d)(2 - r) > 0$ under condition (b), the proof is complete. ∎

Proof of Theorem 3.2 It is sufficient to show that $X_n^N \to_{\mathrm{d}} X^N$ for $N \in \mathbb{N}$, where X_n^N is defined in (3.3). Since $X^N \to_{L_2} X$ as $N \to \infty$ by Lemma **3.3**, the limiting distribution of X_n^N may be made as close to the distribution of X as desired, by choice of N.

Lemma **3.4** establishes the Gaussianity of X^N. Applying Theorems **3.7** and **3.8** for the case $Y_n = X_n^N$ with $L = 0$ and $U = 1$ and letting $\widetilde{T}_n(s,t)$ in (3.35) be defined by (3.18) and (3.16), the fact that for $n \in \mathbb{N}$ and $\eta > 0$ there exists $M < \infty$ such that $P(\widetilde{T}_n(0,1) > M) \leq \eta$ follows from the sufficiency part of Theorem **3.5** and confirms condition (a) of Theorem **3.7**. Theorem **3.8**, whose assumptions are verified by (3.7) and Lemma **3.6**, in turn verifies condition (b) of Theorem **3.7**, so proving uniform tightness and the almost sure continuity of the limit distribution.

Corollary **2.8** establishes the covariance structure, to complete the proof. ∎

3.5 Dependent Shocks

Proving the FCLT under the assumption of independent shocks has the benefit of highlighting the special features of the problem while keeping complexities to a practicable minimum, but the ambition must be to extend the result to allow weak dependence of the shocks. One reason that it might be desirable to treat short-run autocorrelation separately from the specification of the moving average coefficients of equations (1.1) and (1.2) is so that the latter can be given a known functional form. The obvious example is the fractional integration formula (1.12), without which it is not possible to give the fractionally integrated process the neat representation

$$x_i = (1 - L)^{-d} u_i$$

from (1.8). There are circumstances where this formulation might be regarded as losing little generality over (1.1)+(1.2) provided u_i is not required to be an independent process.

A feature of the results of this section is a dependence on material from SLT. These involve, among other things, the following new assumption.

3.9 Assumption For scale constants $\{c_{nm}, m = 1, \ldots, n\}$, there exists $\alpha \in (0, 1]$ such that with $B_n = [n^{1-\alpha}]$ and $r_n = [n/B_n]$ for $n \in \mathbb{N}$, the following conditions hold: if

$$M_{nj} = \max_{(j-1)B_n+1 \leq m \leq jB_n} |c_{nm}| \tag{3.43}$$

for $j = 1, \ldots, r_n$ and $M_{n,r_n+1} = \max_{r_nB_n+1 \leq i \leq n} |c_{ni}|$ then

$$\max_{1 \leq j \leq r_n+1} M_{nj} = o(B_n^{-1/2}) \tag{3.44}$$

and

$$\sum_{j=1}^{r_n} M_{nj}^2 = O(B_n^{-1}). \quad \square \tag{3.45}$$

The following central limit theorem under dependence, with conditions given in a form adapted to the present applications, is proved as Theorem 25.12 of SLT.

3.10 Theorem $\sum_{m=1}^{n} c_{nm} u_m \to_{\mathrm{d}} \mathrm{N}(0, V)$ as $n \to \infty$ if

(a) $\mathrm{E}\big(\sum_{m=1}^{n} c_{nm}u_m\big)^2 \to V < \infty$

(b) $\{u_m\}_{m=1}^{n}$ satisfies Assumption **1.2**

(c) $\{c_{nm}\}_{m=1}^{n}$ satisfies Assumption **3.9**. □

A generic notation is used in Assumption **3.9** and Theorem **3.10** since the identity of the constants and the sample depend on the application. A leading case is $m = i$ for $i = 1, \dots n$, but to reprise the arguments of §3.2 the cases of $m = i + n(k+1)$ for $i = 1 - n(k+1), \dots, -kn$ also arise, for $k = 0, \dots, N-1$. In some applications the number of terms in the sum is not n but an integer multiple thereof, for example $[nt]$ for $t > 0$, but this detail is omitted from the formal statement in the interest of clarity.

Comparing the conditions of Theorem **3.10** with those of SLT Theorem 25.12 itself, condition (a) of the latter theorem specifies a limiting L_2-norm of 1, but this is an arbitrary choice of normalization and any finite constant value can be substituted. Conditions (b) and (c) impose the conditions of SLT Theorem 25.6, which is a CLT for mixingales.

The key fact is that the property specified in Assumption **1.2**(b), of L_2-near-epoch dependence (NED) on a mixing process, is sufficient for u_m to have the L_2-mixingale property of size $-\frac{1}{2}$.[2] With suitable scale adjustments, this property also extends to $c_{nm}u_m$. Formally, letting $\{\mathcal{F}_{nm}\}$ denote a filtration defined on the relevant probability space, an L_2-mixingale of size $-\varphi_0$ is a pair $\{x_{nm}, \mathcal{F}_{nm}\}$ satisfying the conditions[3]

$$\| \mathrm{E}(x_{nm}|\mathcal{F}_{n,m-v}) \|_2 \le |c_{nm}|\zeta_v$$
$$\| x_{nm} - \mathrm{E}(x_{nm}|\mathcal{F}_{n,m+v}) \|_2 \le |c_{nm}|\zeta_{v+1}$$

where $\zeta_v = O(v^{-\varphi})$ for $\varphi > \varphi_0$ and c_{nm} is an array of scale constants.[4] The filtration is typically of the form $\mathcal{F}_{nm} = \sigma(x_{nj}, j \le m)$. The series under present consideration have the general form (with $i = m$) of $x_{ni} = c_{ni}u_i$ where $c_{ni} = a_{ni}(s,t)/\kappa(n)$.

A special feature of SLT Theorem 25.6 is that it accommodates heterogeneity in the distribution of the process increments by fine-tuning the specification of the blocking parameters. This is the role of Assumption **3.9**. Like many limit results for dependent processes, Theorem 25.6 of SLT works by breaking the sample series into r_n blocks of B_n successive terms, similarly to Theorem **2.10**. The property being exploited is that partial sums of mixingales behave approximately like martingales, with the residual terms becoming negligible as n increases under suitable restrictions on the range of dependence. This allows a central limit theorem for martingales[5] to be invoked, although with heterogeneity in the underlying series

[2] Proved as SLT Theorem 18.7.

[3] The x_{nm} in this generic definition is a different object from the x_i of (1.1).

[4] The mixingale concept is reviewed in detail in SLT Chapter 17.

[5] For example, SLT Theorem 25.3.

the blocking construction must satisfy Assumption **3.9**. This condition is critical in the antipersistent case. For the case $d > 0$, any choice $B_n = o(n)$ meets the conditions but with $d < 0$, the restriction on the divergence rate of B_n implied by condition (3.44) is needed to ensure that the variance of every normalized block converges to zero. Comparison of (3.10) and (3.11) in the independent case shows a related contrast.[6]

The CLT can then be shown, subject to the further condition that the normalized block sums form a uniformly square-integrable array. The condition on the blocks is essentially equivalent to what was previously shown for the increments of the form (3.18), with Assumption **3.1** needed to validate the step. Here too, antipersistent processes are subject to an additional restriction relative to long memory processes.

With these preliminaries established, the main result to be shown is the following, where X_n is defined by (2.27) and X is the fBM in (2.1).

3.11 Theorem If Assumptions **1.2**, **3.1** and **3.9** hold then $X_n \to_{\mathrm{d}} X$. □

As before, the proof is constructed via a series of lemmas which mirror the development of the proof of **3.2**, working with the decomposition of (3.2). Of these, the first is the generalization of Lemma **3.3**.

3.12 Lemma Under Assumption **1.2**, $\lim_{n\to\infty} \mathrm{E}(R_{3n}(s,t)^2) = O(N^{2d-1})$ as $N \to \infty$.

Proof The result

$$\begin{aligned}\mathrm{E}(R_{3n}(s,t))^2 &\sim \frac{\omega_u^2}{\kappa(n)^2}\sum_{i=-\infty}^{-nN} a_{ni}(s,t)^2 \\ &\sim \omega_u^2(s-t)^{2d+1}\int_N^\infty ((1+\tau)^d-\tau^d)^2\mathrm{d}\tau \qquad (3.46)\end{aligned}$$

follows by the arguments of Theorem **2.10** and Corollary **2.11**, which can be applied to the complete sum $R_{1n} + R_{2n} + R_{3n}$, then noting that the division into three is simply accomplished by dividing Υ_d into its three components, of which the third is the integral in the last member of (3.46). Note that the equality in (3.4) now becomes the first limit result in (3.46) but the second step of the argument follows (3.4) directly. ∎

The next step is to prove a replacement for Lemma **3.4**, but to do this one further technical property has to be shown and this proof can conveniently be given first. This is, that increments of the empirical process $R_{1n}(s,t) + R_{2n}(s,t)$ are asymptotically uncorrelated.

[6] Lemma 25.10 of SLT offers some additional insight into the role of this restriction.

3.13 Lemma Let $z_{ni} = a_{ni}(s,t)u_i/\kappa(n)$ and for $r \in [-N,1]$ define $Z_n(r) = \sum_{i=-\infty}^{[nr]} z_{ni}$ so that for $q > r$, $Z_n(q) - Z_n(r) = \sum_{i=[nr]+1}^{[nq]} z_{ni}$ is an increment of the process $R_{1n}(s,t) + R_{2n}(s,t)$ in (3.2). For $-N \leq r_1 < q_1 \leq r_2 < q_2 \leq 1$,

$$\lim_{n\to\infty} \mathrm{E}\big((Z_n(q_1) - Z_n(r_1))(Z_n(q_2) - Z_n(r_2))\big) = 0.$$

Proof For any $\eta \in (0, q_2 - r_2)$,

$$\begin{aligned}&\big|\mathrm{E}(Z_n(q_1) - Z_n(r_1))(Z_n(q_2) - Z_n(r_2))\big| \\ &\quad \leq \Bigg|\sum_{i=[nr_1]+1}^{[nq_1]} \sum_{k=[nr_2]+1}^{[n(r_2+\eta)]} \mathrm{E}(z_{ni}z_{nk})\Bigg| + \Bigg|\sum_{i=[nr_1]+1}^{[nq_1]} \sum_{k=[n(r_2+\eta)]+1}^{[nq_2]} \mathrm{E}(z_{ni}z_{nk})\Bigg|. \qquad (3.47)\end{aligned}$$

Applying the modulus inequality to the first majorant term of (3.47) followed by the Cauchy-Schwarz inequality and then using that $z_{nk} = O_p(n^{-1/2})$ and the assumption of weak dependence,

$$\begin{aligned}\Bigg|\sum_{i=[nr_1]+1}^{[nq_1]} \sum_{k=[nr_2]+1}^{[n(r_2+\eta)]} \mathrm{E}(z_{ni}z_{nk})\Bigg| &\leq \mathrm{E}\Bigg|\sum_{i=[nr_1]+1}^{[nq_1]} z_{ni} \sum_{k=[nr_2]+1}^{[n(r_2+\eta)]} z_{nk}\Bigg| \\ &\leq \Bigg\|\sum_{i=[nr_1]+1}^{[nq_1]} z_{ni}\Bigg\|_2 \Bigg\|\sum_{k=[nr_2]+1}^{[n(r_2+\eta)]} z_{nk}\Bigg\|_2 = O(\eta^{1/2}).\end{aligned}$$

Next, consider the second majorant term of (3.47) where $|i-k| \geq n(r_2+\eta-q_1) \geq n\eta$. Assumption **1.2**(a) implies

$$\mathrm{E}(z_{ni}z_{nk}) = \gamma_u(|i-k|)\frac{a_{ni}(s,t)a_{nk}(s,t)}{\kappa(n)^2} = O\big((n(r_2+\eta-q_1))^{-1-\delta}n^{-1}\big) \qquad (3.48)$$

and hence the second majorant term of (3.47) is of $O((n\eta)^{-\delta})$. Now, choose $\eta = O(n^{-\pi})$ for $0 < \pi < 1$ so that $(n\eta)^{-\delta} = O(n^{-\delta(1-\pi)})$ and

$$\big|\mathrm{E}(Z_n(q_1) - Z_n(r_1))(Z_n(q_2) - Z_n(r_2))\big| = O(\max\{n^{-\pi/2}, n^{-\delta(1-\pi)}\}) = o(1). \quad \blacksquare$$

3.14 Lemma Under Assumptions **1.2** and **3.9**,

$$R_{1n}(s,t) + R_{2n}(s,t) \xrightarrow{\mathrm{d}} \mathrm{N}\big(0, \omega_u^2\Upsilon_d(s-t)^{2d+1}\big)$$

for any s and t with $0 \leq t < s \leq 1$.

Proof The case $s = 1$ and $t = 0$ is taken as the exemplar. Appealing to Theorem **3.10**, the main task is to show that Assumption **3.9** holds for the scale constants. In $R_{1n}(1,0)$, noting that the slowly varying components cancel in the limit,

$$|c_{ni}| = \frac{|a_{ni}(1,0)|}{\kappa(n)} \sim \frac{(n-i)^d}{n^{d+1/2}}. \qquad (3.49)$$

When $d > 0$, c_{ni} is decreasing in i and according to (3.43),

$$M_{nj} \sim \frac{(n-(j-1)B_n)^d}{n^{d+1/2}} \sim \frac{(r_n - j + 1)^d}{B_n^{1/2} r_n^{d+1/2}}.$$

Hence,

$$\max_{1\le j\le r_n+1} M_{nj} = M_{n1} = O(n^{-1/2}) = o(B_n^{-1/2}) \tag{3.50}$$

and

$$\sum_{j=1}^{r_n} M_{nj}^2 \backsimeq \frac{1}{B_n r_n^{2d+1}} \sum_{j=1}^{r_n} (r_n - j + 1)^{2d} = O(B_n^{-1}). \tag{3.51}$$

These results hold for $B_n = [n^{1-\alpha}]$ for any $\alpha \in (0,1]$.

On the other hand, if $d < 0$ then c_{ni} is increasing in i and takes its maximum at $i = n$, at which point $a_{ni} = 1$. In this case let $B_n = [n^{1-\alpha}]$ for $-2d < \alpha < 1$, which is a feasible choice with $d > -\frac{1}{2}$. Then,

$$\max_{1\le j\le r_n+1} M_{nj} = c_{nn} \sim \frac{1}{n^{d+1/2}} = o(B_n^{-1/2}). \tag{3.52}$$

In (3.51), j replaces $j-1$ in the definition of M_{nj} but (noting $0^d = 1$ in this context) the sum is still of $O(B^{-1})$. The conditions of Theorem **3.10** are accordingly satisfied for $R_{1n}(1,0)$ with $|d| < \frac{1}{2}$, the limiting variance V in Theorem **3.10**(a) being defined by (3.5) with ω_u^2 replacing σ_u^2 by application of Theorem **2.10**.

Next, consider the sum $R_{2n}(1,0)$ in which the terms are labelled $1-nN \le i \le 0$. Similarly to (3.15) break this sum into N blocks each of length n, denoted

$$R_{2n}^k(1,0) = \sum_{i=1-n(k+1)}^{-nk} \frac{a_{ni}(1,0)}{\kappa(n)} u_i, \quad k = 0, \dots, N-1.$$

The analysis comparable to that for R_{1n} can be applied to each of these blocks. Assumption **3.9** must be verified in effect for the case c_{nm} for $m = 1, \dots, n$, where $m = 1 - nk - i$ and according to (2.30), in this instance,

$$|c_{ni}| \sim \frac{|(n-i)^d - (-i)^d|}{n^{d+1/2}}. \tag{3.53}$$

To distinguish the different cases write M_{nj}^k for $k = 0, \dots, N-1$, where

$$\begin{aligned} M_{nj}^k &= \max_{1-nk-jB_n \le i \le -nk-(j-1)B_n} |c_{ni}| \\ &\sim \frac{|(n+nk+(j-1)B_n)^d - (nk+(j-1)B_n)^d|}{n^{d+1/2}}. \end{aligned} \tag{3.54}$$

For both $d > 0$ and $d < 0$, c_{ni} is decreasing as $-i$ increases and M_{nj}^k is maximized over $j = 1, \dots, r_n$ at $j = 1$. For each $k > 0$,

$$\max_{1\le j\le r_n+1} M_{nj}^k \sim \frac{n^d |(k+1)^d - k^d|}{n^{d+1/2}} = O(n^{-1/2}) = o(B_n^{-1/2}). \tag{3.55}$$

Making the change of variable $\tau = k + (j-1)/r_n$, it is also found from (3.54) that

$$\sum_{j=1}^{r_n}(M_{nj}^k)^2 \sim \frac{B_n^{2d}}{n^{2d+1}}\sum_{j=1}^{r_n}\big((r_n(k+1)+j-1)^d - (r_nk+j-1)^d\big)^2$$
$$\sim \frac{1}{B_n}\int_k^{k+1}\big((1+\tau)^d - \tau^d\big)^2 \mathrm{d}\tau = O(B_n^{-1}). \tag{3.56}$$

It follows that the conditions of Theorem **3.10** are satisfied by $R_{2n}^k(1,0)$ with

$$V = \lim_{n\to\infty} \mathrm{E}(R_{2n}^k(1,0))^2 = \omega_u^2 \int_k^{k+1}\big((1+\tau)^d - \tau^d\big)^2 \mathrm{d}\tau. \tag{3.57}$$

When $d > 0$, the indicated orders of magnitude in (3.55) and (3.56) hold for all $k \geq 0$. However, when $d < 0$, $k^d > (k+1)^d$ and in the case $k = 0$, with 0^d equated to 1 in the usual way, (3.55) has to be replaced by

$$\max_{1\leq j\leq r_n+1} M_{nj}^0 = M_{n1}^0 \sim \frac{1}{n^{d+1/2}} = O(n^{-d-1/2}) = o(B_n^{-1/2})$$

where the last equality holds provided that $B_n = [n^{1-\alpha}]$ for $-2d < \alpha < 1$. The calculation in (3.56) continues to apply in this case for $k = 0$, so that (3.45) is still validated.

To extend these arguments to the cases of $R_{1n}(s,t)$ and $R_{2n}(s,t)$, for any $s,t \in (0,1]$ is a matter of replacing n by $[ns]$ and setting $c_{ni} = a_{ni}(s,t)/\kappa(n)$.

The individual blocks R_{1n} and $R_{2n}^0, \ldots, R_{2n}^{N-1}$ are therefore Gaussian in the limit, and are also uncorrelated in the limit in view of Lemma **3.13**. Therefore the product of the marginal densities is identical in form to the joint density under independence and it follows that in the limit the blocks are jointly Gaussian and hence in particular their sum is Gaussian. The variances of the form (3.57) for each k are summable over $k = 0, \ldots, N-1$, hence the sum approaches the limit having the form of (3.6) with ω_u^2 replacing σ_u^2. ∎

The property of independent increments is important for another reason, which is to show that the dependence of the limit process X is determined exclusively by d, in the manner represented by equation (2.1), so validating Corollary **2.8** and the characteristic fractional dependence illustrated by (2.17). Such a demonstration was not needed in Lemma **3.4** because there the increments were independent by assumption.

The next step is to show uniform boundedness in probability as the preliminary to proving uniform tightness. Theorem **3.5** is for independent shocks and does not apply under Assumption **1.2**. However, the following assumption is shown to be a sufficient restriction on the dependence.

3.15 Assumption For $t \in [-N, 1]$, $\delta \in (0, 1-t]$, $n \in \mathbb{N}$, and $\eta > 0$,

$$P\Big(\sup_{t\leq s\leq t+\delta} |u_{[ns]}| \leq \eta\Big) \geq \big(1 - P(|u_i| > \eta)\big)^{[n(t+\delta)]-[nt]}. \quad \square \tag{3.58}$$

If the second equality of (3.20) in the proof of Theorem **3.5** is replaced by inequality (3.58), the sufficiency part of Theorem **3.5** continues to hold under dependence. This is shown formally as follows.

3.16 Corollary Under Assumptions **1.2**, **3.1**, and **3.15**, the collection $\{\widetilde{T}_n^2(t,\delta), n \in \mathbb{N}\}$ for $\delta > 0$ and $0 \le t \le 1-\delta$ is uniformly bounded in probability.

Proof The proof of Theorem **3.5** can be followed at most points, but under Assumption **3.15**, the second equality in equation (3.20) is replaced by an inequality as in (3.58). Equation (3.23) is therefore replaced by

$$-\log P\left(\sup_{-N\le s\le 1}\frac{|u_{[ns]}|}{\kappa(n)} \le \varepsilon\right) \le O(\varepsilon^{-r}\log(\varepsilon\kappa(n))^{-1-\mu}) \tag{3.59}$$

and under Assumption **3.1** the probability on the left-hand side of (3.59) must converge to 1 as $\varepsilon \to \infty$. ∎

If the inequality in (3.59) was not imposed the divergence of the supremum with positive probability could occur, but Assumption **3.15** rules out this eventuality.

Having regard to the plausibility of condition (3.58), the first thing to observe is that the existence or otherwise of autocorrelation in levels is irrelevant. The relation concerns only the joint distribution of the absolute values, or squares, of the process. Consider partitioning the sequence elements u_i in (3.58) into subsets S_A and S_B and so defining events $A = \bigcup_{S_A}\{|u_i| > \eta\}$ and $B = \bigcup_{S_B}\{|u_i| > \eta\}$. The intersection of the complements of these two sets is the event whose probability is on the left-hand side of (3.58). The relation

$$P(A^c \cap B^c) \ge P(A^c)P(B^c) = (1-P(A))(1-P(B)) \tag{3.60}$$

rearranges using the de Morgan law as $P(B|A) \ge P(B)$ and equivalently as $P(A|B) \ge P(A)$. If an element of one set exceeds the bound, it is at least as probable according to (3.60) that an element of the other set will do so. Suppose for example that the elements of S_A are odd-numbered in the time sequence and those of S_B are even-numbered. Then, a failure of (3.60) would represent the opposite of volatility clustering. Negative correlation of adjacent squares in the time sequence is not an impossible scenario, but few empirical studies of conditional heteroscedasticity suggest such outcomes.

Further, the partitioning exercise can be performed on each of sets S_A and S_B and an inequality obtained for the minorant of (3.60). Repeating the cycle until all the sets are singletons defines a chain of inequalities for products of probabilities, terminating in the right-hand side of (3.58). If (3.60) held for every such partition then Assumption **3.15** certainly follows, but this strong condition is only one of many ways the assumption might be validated. In the example of adjacent squares, note how the succeeding partitions would involve sample elements further separated in time and hence with a diminishing role under weak dependence. For Assumption **3.15** to be contradicted, inequality (3.60) would need to fail at a preponderance of the partitioning steps. These considerations suggest that the alternative to (3.58) can with some confidence be relegated to a special case.

A further issue is that Lemma **3.6** cannot be applied directly to show the condition imposed by Lemma **3.8**, since it uses Theorem **A.5** which does not work for mixingales. A uniform integrability proof for mixingales is given as Theorem 17.14 of SLT, and as with Theorem **A.5**, this result generalizes straightforwardly to lower and upper sum bounds of the form $[nt]+1$ and $[ns]$, for any $t<s$. The drawback with this theorem is that while it shows uniform integrability of the maximal sum it does not specify a rate of convergence in the manner of equation (A.10), which does need to be known for the application to Theorem **3.8**.

While the proof given for Theorem 17.14 of SLT is both lengthy and tricky, the extension is minor and can be given in terms of steps in that proof. Similarly to Theorem **3.10**, the interested reader is left to pursue the details with the relevant pages of SLT to hand (pages 357–360 of the 2nd Edition). Except for the addition of condition (3.61), the following statement matches SLT Corollary 17.15, which is the extension to arrays of the main result for mixingale sequences. With $m=i$ in the application, x_{ni}/c_{ni} here corresponds to u_i which is stationary and L_r-bounded for $r\geq 2$ under Assumption **1.2**.

3.17 Theorem For $r\geq 2$ let $\{x_{nm},\mathcal{F}_{nm}\}$ be a L_r-bounded, L_2-mixingale array of size $-\frac{1}{2}$. Let $S_{nk}=\sum_{m=1}^k x_{nm}$ and $v_n^2=\sum_{m=1}^n c_{nm}^2$ where c_{nm} are the mixingale scale constants. If the array $\{x_{nm}^2/c_{nm}^2\}$ is uniformly integrable the collection $\{\max_{1\leq k\leq n} S_{nk}^2/v_n^2,\ n\in\mathbb{N}\}$ is uniformly integrable, with

$$\mathrm{E}\Big(\big(\max_{1\leq k\leq n} S_{nk}^2/\nu_n^2\big)1_{\{\max_{1\leq k\leq n}|S_{nk}|/\nu_n>\eta\}}\Big)=o(\eta^{2-r}). \qquad (3.61)$$

Proof In view of SLT's Theorem 17.4, it suffices to show that when that proof is valid, (3.61) also holds. The proof proceeds by setting a number ε to be the common upper bound of three terms, there denoted $\mathrm{E}(\hat{u}_n^2)$, $\mathrm{E}(\hat{z}_n^2)$, and $\mathcal{E}_{M/6}(\hat{y}_n^2)$. The sum of these terms defines a bound on $\mathcal{E}_M(\hat{x}_n^2)=\mathrm{E}(\hat{x}_n^2 1_{\{|\hat{x}_n|>M\}})$ where $\hat{x}_n^2=\max_{1\leq k\leq n} S_{nk}^2/\nu_n^2$. It is shown that ε is a decreasing function of the bound M with $\varepsilon\to 0$ as $M\to\infty$.

The key relations are (17.65), (17.67), and (17.74). Equation (17.65) shows $\mathrm{E}(\hat{u}_n^2)=O(m^{-\delta})$, where m is the order of lead/lag defining the mixingale truncation in definitions (17.52)–(17.54); see also Definition 17.1 of SLT. The parameter δ can be equated with the serial dependence parameter from Assumption **1.2**(a). The notation $\backsimeq$ is used here to denote matching orders of magnitude as $\varepsilon\to 0$. Following (17.65), let the bound be represented as $\varepsilon\backsimeq m^{-\delta}$ and equivalently, let $m\backsimeq\varepsilon^{-1/\delta}$ denote the approximate value of m required to attain the bound.

Next, under the assumption that x_{nm} is an L_r-bounded random variable for $r\geq 2$, use the result $\mathrm{E}\big((x_{nm}/c_{nm})^2 1_{\{|X_{nm}/c_{nm}|>B\}}\big)=o(B^{2-r})$ from Theorem **A.4** of Appendix A. With this assumption, inequality (17.67) (where X_t/c_t can be read as x_{nm}/c_{nm}) can be extended by a step beyond the source proof and written as

$$\mathrm{E}(\hat{z}_n^2)\ll mB^{2-r}\backsimeq\varepsilon. \qquad (3.62)$$

Since the relations represent orders of magnitude the notation allows scale constants to be taken as implicit.

Substituting for m in (3.62) and rearranging gives $B \backsimeq \varepsilon^{(1+1/\delta)/(2-r)}$ as approximating the value of B needed to attain the bound. Inequality (17.74) can be rendered in the same manner as

$$\mathcal{E}_{M/6}(\hat{y}_n^2) \ll \frac{m^4 B^4}{M} \backsimeq \varepsilon. \tag{3.63}$$

Substituting for m and B in (3.63), rearranging, simplifying, and inverting gives for $0 < \varepsilon < 1$ and $r \geq 2$,

$$M^{2-r} \backsimeq \varepsilon^{2+r+4(r-1)/\delta} \leq \varepsilon. \tag{3.64}$$

Equating M with η to match the notation of Theorem **A.5**, Theorem **A.4** applied to the random variable $\max_{1\leq k\leq n} |S_{nk}|^2/\nu_n^2$ leads to (3.61) replacing (17.75). ∎

Like Theorem **A.5**, Theorem **3.17** places no restriction on the ordering of summands and can duplicate its role in the argument for the mixingale case. With $\widetilde{T}_n(t,\delta)$ defined as in (3.18) except that ω_u^2 replaces σ_u^2, the required result is as follows.

3.18 Lemma Under Assumptions **1.2**, **3.1**, and **3.15**, the collection $\{\widetilde{T}_n^2(t,\delta), n > n_0\}$ is uniformly integrable and satisfies (3.25), for $n_0 < \infty$ and all $0 < \delta < 1$ and $t \in [0, 1-\delta]$.

Proof Identical to the proof of Lemma **3.6**, subject only to replacement of the citations of Theorem **A.5** by citations of Theorem **3.17**. ∎

Proof of Theorem 3.11 Lemma **3.12** focuses attention on the terms R_{1n} and R_{2n} of (3.2) as before. Then, Lemmas **3.14**, and **3.13** establish the Gaussianity and increment independence of the limit distribution. Lemma **3.18** establishes condition (b) of Theorem **3.8,** which proves the almost sure continuity of the limit distribution in the same way as before. Theorem **2.10** and Corollary **2.11**, in combination with Theorem **2.6** and Corollary **2.8**, fix the covariance structure to complete the proof. ∎

The conclusion of these results is that the only consequence for the limit distribution of swapping Assumption **1.2** for Assumption **1.1** is the inflation of the variance parameter. In the important case of conditional heteroscedasticity, the shock series may be uncorrelated although dependent and in that case the limits in question may actually match the i.i.d. case. However, as can be deduced from the style of the proofs, the rate of convergence to the limit can be substantially slower. Much the most important consequence of shock dependence is that larger samples may be needed to attain comparable approximations to the limit.

3.6 The Multivariate FCLT

The matrix notation required to handle the multivariate case is detailed in §2.5. Let D^m denote the space of m-dimensional càdlàg functions and C^m the space of

m-dimensional continuous functions. The domain of the functions is not indicated explicitly to leave this option open, although $[0,1]$ is the case here.

For the space of càdlàg m-tuples endowed with the Skorokhod topology (see page 20 for details), the notation D^m is taken to convey that the times of discontinuities are coordinated. In other words, Skorokhod distances must be defined in terms of a common change-of-time function λ. Vectors of càdlàg functions whose jumps are uncoordinated are not treated in the present theory, but the elements $X_{1n},\dots,X_{mn}$ in the present case are step functions whose jump times are dictated by observation dates, hence coordinated.

The following fundamental result, proved as Theorem 30.13 of SLT, extends the well-known Cramér-Wold theorem to random functions and is at the heart of the multivariate FCLT, so deserves a formal statement in this context.

3.19 Theorem Let $\boldsymbol{X}_n \in D^m$ be an m-vector of random elements. $\boldsymbol{X}_n \to_{\mathrm{d}} \boldsymbol{X}$ where $P(\boldsymbol{X} \in C^m) = 1$ if and only if $\boldsymbol{\lambda}'\boldsymbol{X}_n \to_{\mathrm{d}} \boldsymbol{\lambda}'\boldsymbol{X}$, where $P(\boldsymbol{\lambda}'\boldsymbol{X} \in C) = 1$, for every fixed $\boldsymbol{\lambda}$ $(m \times 1)$ with $\boldsymbol{\lambda}'\boldsymbol{\lambda} = 1$. $\square$

With this setup the multivariate extension of Theorem **3.11** can be specified. Taking advantage of the developments in §3.5, it is possible to allow the conditions of Assumption **1.2** for full generality.

3.20 Theorem If the process $\boldsymbol{X}_n$ $(m \times 1)$ is defined by (2.47) where $-\frac{1}{2} < d_1 \leq \cdots \leq d_m < \frac{1}{2}$ and the elements of $\{\boldsymbol{u}_i\}$ each satisfy Assumptions **1.2**, **3.1**, and **3.9**, then $\boldsymbol{X}_n \to_{\mathrm{d}} \boldsymbol{X}$ where $\boldsymbol{X}$ is a vector whose k^{th} element is a fBM with parameter d_k, with covariance structure given by the limit in (2.72).

Proof By **3.19** the result follows if $\boldsymbol{\lambda}'\boldsymbol{X}_n$ converges in distribution to an a.s. continuous Gaussian limit $\boldsymbol{\lambda}'\boldsymbol{X}$ for all m-vectors $\boldsymbol{\lambda}$ of unit length. These limits are not fBMs in general, but as a consequence of Corollary **2.12** the cases $\boldsymbol{\lambda} = \boldsymbol{e}_k$ for $k = 1,\dots,m$ (the columns of the identity matrix of order m) yield fBMs with parameters d_k.

Establishing the finite dimensional distributions can proceed in tandem with the arguments of Theorem **3.11**. In parallel with (3.2) let

$$\begin{aligned}\boldsymbol{\lambda}'\boldsymbol{X}_n(s) - \boldsymbol{\lambda}'\boldsymbol{X}_n(t) &= \left(\sum_{i=[nt]+1}^{[ns]} + \sum_{i=1-nN}^{[nt]} + \sum_{i=-\infty}^{-nN}\right)\boldsymbol{\lambda}'\boldsymbol{D}_n^{-1}\boldsymbol{A}_{ni}(s,t)\boldsymbol{u}_i \\ &= R_{1n}^{\lambda}(s,t) + R_{2n}^{\lambda}(s,t) + R_{3n}^{\lambda}(s,t). \qquad (3.65)\end{aligned}$$

To verify the status of the remainder term $R_{3n}^{\lambda}(s,t)$, note that according to (2.49), $\gamma_{kl} = \gamma_{kl}^N + (\gamma_{kl} - \gamma_{kl}^N)$ where

$$\begin{aligned}\gamma_{kl} - \gamma_{kl}^N &= \int_N^{\infty} ((1+\tau)^{d_k} - \tau^{d_k})((1+\tau)^{d_l} - \tau^{d_l})\mathrm{d}\tau \\ &\leq \max_{k,l}\left\{\int_N^{\infty} ((1+\tau)^{d_k} - \tau^{d_k})^2\mathrm{d}\tau, \int_N^{\infty} ((1+\tau)^{d_l} - \tau^{d_l})^2\mathrm{d}\tau\right\}\end{aligned}$$

$$= O(N^{2\max\{d_k,d_l\}-1}).$$

It follows by Theorem **2.9** that

$$\begin{aligned}\lim_{n\to\infty} \mathrm{E}(R_{3n}^{\lambda}(s,t)^2) &= \sum_k \sum_l \lambda_k \lambda_l \omega_{kl}(\gamma_{kl} - \gamma_{kl}^N)(s-t)^{d_k+d_l+1} \\ &= O(N^{2d_m-1}).\end{aligned}$$

By assumption this is of small order in N and as in the univariate case, N can be taken large enough for this term to be negligible.

Consider the remaining sums in turn. As before, it is convenient to consider the case $s = 1$ and $t = 0$, the extension to other cases being easily supplied. As in Lemma **3.14**, the finite dimensional CLT is proved by verifying the conditions of Theorem **3.10**, the essential step being to specify the scale constants to be attached to each observation. For fixed $\boldsymbol{\lambda}$, write $W_{ni}^{\lambda} = \boldsymbol{\lambda}'\boldsymbol{D}_n^{-1}\boldsymbol{A}_{ni}(1,0)\boldsymbol{u}_i$ and so define the array $\{c_{ni}^{\lambda}\}$, being the positive square root of

$$(c_{ni}^{\lambda})^2 = \mathrm{E}(W_{ni}^{\lambda\,2}) = \boldsymbol{\lambda}'\boldsymbol{A}_{ni}(1,0)\boldsymbol{D}_n^{-1}\boldsymbol{\Omega}_u\boldsymbol{D}_n^{-1}\boldsymbol{A}_{ni}(1,0)\boldsymbol{\lambda} \tag{3.66}$$

where the long-run variance matrix $\boldsymbol{\Omega}_u$ is defined by (2.71). Then, for $i = 1,\dots,n$ corresponding to the terms of $R_{1n}^{\lambda}(1,0)$, applying Theorem **2.5** and noting that the slowly varying factors cancel,

$$(c_{ni}^{\lambda})^2 \sim \sum_{k=1}^{m}\sum_{l=1}^{m} \lambda_k\lambda_l\omega_{kl}\frac{(n-i)^{d_k+d_l}}{n^{d_k+d_l+1}} = \frac{(n-i)^{2d_1}}{n^{2d_1+1}}J_{ni}^{\lambda} \tag{3.67}$$

with $d_1 = \min\{d_1,\dots,d_m\}$ and

$$J_{ni}^{\lambda} = \sum_{k=1}^{m}\sum_{l=1}^{m} \lambda_k\lambda_l\omega_{kl}\frac{(n-i)^{d_k+d_l-2d_1}}{n^{d_k+d_l+1-2d_1}}. \tag{3.68}$$

Compare formula (3.67) with (3.49). The exponents in (3.68) are nonnegative by construction and the array J_{ni}^{λ} is both positive and finite for all $n > 1$. The conditions of Assumption **3.9** are therefore met, following the line of reasoning in the proof of Theorem **3.14**, subject to $d_1 > -\frac{1}{2}$. The finite-dimensional distributions of the processes R_{1n}^{λ} are proved in the same manner as before.

For $i = 1 - nN,\dots,0$, corresponding to terms of $R_{2n}^{\lambda}(1,0)$,

$$\begin{aligned}(c_{ni}^{\lambda})^2 &\sim \sum_{k=1}^{m}\sum_{l=1}^{m} \lambda_k\lambda_l\omega_{kl}\frac{((n-i)^{d_k} - (-i)^{d_k})((n-i)^{d_l} - (-i)^{d_l})}{n^{d_k+d_l+1}} \\ &= \frac{(n-i)^{2d_1}}{n^{2d_1+1}}J_{ni}^{\lambda 1} + \frac{(-i)^{2d_1}}{n^{2d_1+1}}J_{ni}^{\lambda 2} - 2\frac{(n-i)^{d_1}(-i)^{d_1}}{n^{2d_1+1}}J_{ni}^{\lambda 3}\end{aligned} \tag{3.69}$$

where

$$J_{ni}^{\lambda 1} = \sum_{k=1}^{m}\sum_{l=1}^{m} \lambda_k\lambda_l\omega_{kl}\left(\frac{n-i}{n}\right)^{d_k+d_l-2d_1}$$

$$J_{ni}^{\lambda 2}=\sum_{k=1}^{m}\sum_{l=1}^{m}\lambda_k\lambda_l\omega_{kl}\Big(\frac{-i}{n}\Big)^{d_k+d_l-2d_1}$$
$$J_{ni}^{\lambda 3}=\tfrac{1}{2}\sum_{k=1}^{m}\sum_{l=1}^{m}\lambda_k\lambda_l\omega_{kl}\bigg(\Big(\frac{n-i}{n}\Big)^{d_k-d_1}\Big(\frac{-i}{n}\Big)^{d_l-d_1}+\Big(\frac{-i}{n}\Big)^{d_k-d_1}\Big(\frac{n-i}{n}\Big)^{d_l-d_1}\bigg).$$

The terms $J_{ni}^{\lambda 1}$, $J_{ni}^{\lambda 2}$ and $J_{ni}^{\lambda 3}$ are strictly positive and finite for all $1-nN\le i\le 0$ and all $n\ge 1$. There therefore exist constants L^λ and U^λ, such that

$$0<L^\lambda\le\min(J_{ni}^{\lambda 1},J_{ni}^{\lambda 2},J_{ni}^{\lambda 3})\le\max(J_{ni}^{\lambda 1},J_{ni}^{\lambda 2},J_{ni}^{\lambda 3})\le U^\lambda<\infty$$

and such that when n is large enough, $L_{ni}^\lambda\le(c_{ni}^\lambda)^2\le U_{ni}^\lambda$ where

$$L_{ni}^\lambda=L^\lambda\frac{((n-i)^{d_1}-(-i)^{d_1})^2}{n^{2d_1+1}},$$
$$U_{ni}^\lambda=U^\lambda\frac{((n-i)^{d_1}-(-i)^{d_1})^2}{n^{2d_1+1}}.$$

Consider the stationary weakly dependent process $W_{ni}^\lambda/c_{ni}^\lambda$, for which scale constants of unity are appropriate. The argument of Lemma **3.14** applies to the increment processes $\sqrt{L_{ni}^\lambda}W_{ni}^\lambda/c_{ni}^\lambda$ and $\sqrt{U_{ni}^\lambda}W_{ni}^\lambda/c_{ni}^\lambda$, setting $\sqrt{L_{ni}^\lambda}$ and $\sqrt{U_{ni}^\lambda}$ respectively for the arrays of scale constants. In particular, the fact that the conditions of Assumption **3.9** are satisfied in these cases for $d_1>-\frac{1}{2}$ was established in the proof of Lemma **3.14**, leading to equations (3.55) and (3.56). (Note, in these arguments the magnitude of scale factors is irrelevant.) Since $\sqrt{L_{ni}^\lambda}/c_{ni}^\lambda<1<\sqrt{U_{ni}^\lambda}c_{ni}^\lambda$, W_{ni}^λ itself is tending to a convex combination of these processes as n increases. If both $\sum_{i=1-N_n}^{0}\sqrt{L_{ni}^\lambda}W_{ni}^\lambda/c_{ni}^\lambda$ and $\sum_{i=1-N_n}^{0}\sqrt{U_{ni}^\lambda}W_{ni}^\lambda/c_{ni}^\lambda$ are Gaussian in the limit then the same is true of R_{2n}^λ. These arguments establish the finite-dimensional distributions of the scalar processes.

To prove uniform tightness of the distributions, it is convenient to define

$$T_{nk}(s,t)=\sum_{i=1-nN}^{[ns]}\frac{a_{nki}(s,t)}{n^{d_k+1/2}L_k(n)}u_{ki}\tag{3.70}$$

so that $\sum_{k=1}^m\lambda_kT_{nk}(s,t)=R_{1n}^\lambda(s,t)+R_{2n}^\lambda(s,t)$ from (3.65). Also let

$$\nu_n^\lambda(t,\delta)^2=\sum_{k=1}^{m}\sum_{l=1}^{m}\lambda_k\lambda_l\omega_{kl}\sum_{i=1-nN}^{[n(t+\delta)]}\frac{a_{nki}(t+\delta,t)a_{nli}(t+\delta,t)}{n^{d_k+d_l+1}L_k(n)L_l(n)}.\tag{3.71}$$

Application of Theorem **2.9** gives, since $\min\{d_1,\dots,d_m\}>-\frac{1}{2}$ by assumption,

$$\lim_{n\to\infty}\nu_n^\lambda(t,\delta)^2=\sum_{k=1}^{m}\sum_{l=1}^{m}\lambda_k\lambda_l\omega_{kl}\Upsilon_{kl}\delta^{d_l+d_k+1}\to 0\text{ as }\delta\to 0.\tag{3.72}$$

It has to be shown that for $0<\delta<1$ and $t\in[0,1-\delta]$ the collections

$$\left\{\sup_{\{s:|s-t|<\delta\}}\frac{\left(\sum_{k=1}^m\lambda_kT_{nk}(s,t)\right)^2}{\nu_n^\lambda(t,\delta)^2},\ n\ge n_0\right\}\tag{3.73}$$

satisfy the conditions of Lemma **3.18**. If so, uniform tightness follows directly by Theorem **3.8**, given the convergence with δ indicated in (3.72).

To investigate (3.73), the approach is first to consider the terms of the sum one at a time. For $k = 1, \dots, m$, the particular case of $(c_{ni}^{\lambda})^2$ in (3.66) when $\boldsymbol{\lambda} = \boldsymbol{e}_k$ (the k^{th} column of the identity matrix) has the form

$$c_{nki}^2(s,t) = \frac{\omega_{kk} a_{nki}(s,t)^2}{n^{2d_k+1} L_k(n)^2}. \tag{3.74}$$

For each k, form the sums of these elements for the increment of width δ and note that

$$\nu_{nk}^2(t,\delta) = \sum_{i=-1-nN}^{[n(t+\delta)]} c_{nki}^2(t+\delta,t) \to \omega_{kk} \Upsilon_{kk} \delta^{2d_k+1} \tag{3.75}$$

as $n \to \infty$. Then, under Assumptions **1.2**, **3.1**, and **3.15** the collections

$$\left\{ \sup_{\{s:|s-t|<\delta\}} \frac{T_{nk}(s,t)}{\nu_{nk}(t,\delta)}, n \geq n_0 \right\}$$

satisfy the conditions of Lemma **3.18**, for each k. The conditions of Theorem **3.17** hold under Assumption **1.2**, noting that the squares of the terms of the sum (3.70) divided by the $c_{nki}^2(s,t)$ from (3.74) have the form u_{ki}^2/ω_{kk}.

Next, m successive applications of Theorem **A.7**, in conjunction with the argument leading to inequality (3.27) and then application of Theorem **A.6**, extend the required conditions to the collection

$$\left\{ \sup_{\{s:|s-t|<\delta\}} \left(\sum_{k=1}^{m} \frac{T_{nk}(s,t)}{\nu_{nk}(t,\delta)} \right)^2, \ n \geq n_0 \right\}. \tag{3.76}$$

The next step is to apply this reasoning to another case, in which, for the given set of weights $\lambda_1, \dots, \lambda_m$, u_{ki} in (3.70) is replaced for each $k = 1, \dots, m$ by

$$u_{ki}^* = \frac{\lambda_k \nu_{nk}(t,\delta)}{\nu_n^{\lambda}(t,\delta)} u_{ki}. \tag{3.77}$$

Let $T_{nk}^*(s,t)$ denote this case of (3.70) with u_{ki}^* from (3.77) replacing u_{ki}. Then, (3.75) together with (3.72) shows that for any $\delta > 0$ the ratios of scale factors in (3.77) converge to finite constant limits. Hence, under Assumption **1.2** the collection

$$\left\{ \sup_{\{s:|s-t|<\delta\}} \left(\sum_{k=1}^{m} \frac{T_{nk}^*(s,t)}{\nu_{nk}(t,\delta)} \right)^2, \ n \geq n_0 \right\} \tag{3.78}$$

satisfies the required conditions in the same way as the collection (3.76). This completes the proof, since the collection in (3.78) is identical with the collection in (3.73). ∎

Chapter 4

The Fractional Covariance

The material of the foregoing chapters lays the groundwork for posing various fundamental questions arising in econometrics, in particular, the distributions of regression coefficients when the variables entering the relationships in question may be fractional series, either regressors, or disturbance terms, or both. Interest therefore focuses on the limit distributions of sample means, variances and covariances under suitable normalizations.

Regressions may of course involve stationary series as defined by relations (1.1) and (1.2), but the cases of greatest interest inevitably involve partial sum processes of the form (2.21). This generalizes the usual cointegration methodology for non-stationary series with autoregressive roots of unity, which corresponds in the fractional context to the case $d = 0$. The suitably normalized covariance between a unit root process and a weakly dependent process is well-known to converge to a stochastic integral of the Itô type, having Brownian motion as the integrator process.

Generalizing these results to the case where either or both of the partial sum and the stationary integrator are fractional is now the primary object. The development occupies both the current chapter and Chapters 5 and 6 following. The three stages of the analysis are to calculate limiting expectations for the random variable in question, to develop heuristic representations of the limiting integrals and then, not least, to prove weak convergence to these limits. The developments are based on some original ideas sketched in [19], which is joint work with Nigar Hashimzade. Chapter 7 then briefly shows what the results obtained imply for regression analysis.

4.1 Assumptions and Preliminaries

Let x_i and y_i be linear processes having the MA(∞) forms

$$x_i = \sum_{j=0}^{\infty} b_j u_{i-j}, \qquad y_i = \sum_{l=0}^{\infty} c_l w_{i-l} \tag{4.1}$$

where $b_j \sim d_x j^{d_x-1} L_x(j)$ and $c_l \sim d_y l^{d_y-1} L_y(l)$. The analysis will initially focus on the following particular setup.

4.1 Assumption The pair (x_i, y_i) are defined by (4.1) where (a) $|d_x| < \frac{1}{2}$, $|d_y| < \frac{1}{2}$; (b) the sequences (u_i, w_i) are i.i.d. with means of zero, L_r-bounded with $r \geq 2$ and with contemporaneous covariance matrix

$$\mathrm{E}\begin{bmatrix} u_i \\ w_i \end{bmatrix}\begin{bmatrix} u_i & w_i \end{bmatrix} = \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_u^2 & \sigma_{uw} \\ \sigma_{uw} & \sigma_w^2 \end{bmatrix} \tag{4.2}$$

and $\mu_{uw}^4 = \mathrm{E}(u_i^2 w_i^2) < \infty$. $\square$

The case $u_i = w_i$ is allowed. These random variables define a filtration $\boldsymbol{F} = \{\mathcal{F}_i, i \in \mathbb{Z}\}$ on the probability space such that the processes $\{x_i, \mathcal{F}_i\}$ and $\{y_i, \mathcal{F}_i\}$ are adapted.

The assumption of independent shocks in (4.1) is restrictive but substantially simplifies the mathematics. The wholly linear framework has the benefit of simplicity while focusing interest on the long memory characteristics of the data. Since the coefficients b_j and c_l may have arbitrary forms for finite j and l, the range of models encompassed even by the present assumptions is wide. The extensions necessary to allow for weak dependence (specifically, autocorrelation) in the shock processes are treated separately, in Chapter 8.

The next two results relate to the empirical covariance of a pair of stationary fractional series. Define $\sigma_{xy} = \mathrm{E}(x_i y_i)$. Since under Assumption **4.1** the sequence

$$b_j c_j \sim d_x d_y j^{d_x+d_y-2} L_x(j) L_y(j)$$

is summable and the shocks are serially uncorrelated, the relation

$$\sigma_{xy} = \sigma_{uw} \sum_{j=0}^{\infty} b_j c_j < \infty \tag{4.3}$$

follows in a similar way to (1.5).

4.2 Theorem Under Assumption **4.1**, $n^{-1}\sum_{i=1}^n x_i y_i \to_{L_2} \sigma_{xy}$.

Proof

$$\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^n x_i y_i - \sigma_{xy}\right)^2 = \frac{1}{n^2}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}\sum_{p=0}^{\infty} b_j b_m c_l c_p$$
$$\times \sum_{i=1}^n \sum_{k=1}^n \big(\mathrm{E}(u_{i-j} w_{i-l} u_{k-m} w_{k-p}) - \mathrm{E}(u_{i-j} w_{i-l})\mathrm{E}(u_{k-m} w_{k-p})\big) \tag{4.4}$$

where Assumption **4.1** implies that

$$\mathrm{E}(u_{i-j} w_{i-l} u_{k-m} w_{k-p}) - \mathrm{E}(u_{i-j} w_{i-l})\mathrm{E}(u_{k-m} w_{k-p})$$

$$= \begin{cases} \mu_{uw}^4 - \sigma_{uw}^2 & i-j = i-l = k-m = k-p \\ \sigma_u^2\sigma_w^2 & i-j = k-m \neq i-l = k-p \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

It has to be shown that the terms of (4.4) that are subject to the two types of restriction in (4.5) are collectively of small order. Writing out the terms in which the first restriction holds, since $j = l$ and $m = p$ the fourfold sum reduces to a twofold sum. Since u and w are always contemporaneous in these terms, they can be combined as $x_i = u_i w_i - \mathrm{E}(u_i w_i)$ and by Assumption **4.1** this series is serially independent with mean zero and variance $\mu_{uw}^4 - \sigma_{uw}^2$. Also, let $h_j = b_j c_j$ which defines a summable sequence. If these substitutions are made in (4.4) the result is

$$\frac{1}{n^2}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty} h_j h_m \sum_{i=1}^{n}\sum_{k=1}^{n} \mathrm{E}(x_{i-j}x_{k-m}) = \frac{1}{n^2}\mathrm{E}\Bigg(\sum_{j=0}^{\infty}\sum_{i=1}^{n} h_j x_{i-j}\Bigg)^2. \tag{4.6}$$

Re-ordering the terms so that the coefficients attached to each x_i are gathered together (compare (2.22)), in view of the serial independence and summability of the coefficients the right-hand side of (4.6) is

$$\begin{aligned}
&\frac{1}{n^2}\mathrm{E}\Bigg(\sum_{j=0}^{\infty}\Bigg(\sum_{k=\max\{0,j-n+1\}}^{j} h_k\Bigg)x_{n-j}\Bigg)^2 \\
&\quad= \frac{1}{n^2}\mathrm{E}\Bigg(\sum_{j=0}^{n-1}\Bigg(\sum_{k=0}^{j} h_k\Bigg)x_{n-j} + \sum_{j=n}^{\infty}\Bigg(\sum_{k=j-n+1}^{j} h_k\Bigg)x_{n-j}\Bigg)^2 \\
&\quad= \frac{\mu_{uw}^4 - \sigma_{uw}^2}{n^2}\Bigg(\sum_{j=0}^{n-1}\Bigg(\sum_{k=0}^{j} h_k\Bigg)^2 + \Bigg(\sum_{j=n}^{2n-1} + \sum_{j=2n}^{\infty}\Bigg)\Bigg(\sum_{k=j-n+1}^{j} h_k\Bigg)^2\Bigg) \\
&\quad\ll \frac{1}{n^2}\Bigg(\sum_{j=0}^{n-1} j^{2(d_x+d_y-1)} + n + \sum_{j=2n}^{\infty} j^{2(d_x+d_y-2)}\Bigg) \\
&\quad= O(n^{-1}).
\end{aligned}$$

The order-of-magnitude inequality here neglects any slowly varying components which cannot affect the result, since $d_x + d_y < 1$.

The second restriction in (4.5) implies that

$$\begin{aligned}
\mathrm{E}(u_{i-j}w_{i-l}u_{k-m}w_{k-p}) - \mathrm{E}(u_{i-j}w_{i-l})\mathrm{E}(u_{k-m}w_{k-p}) \\
= \mathrm{E}(u_{i-j}u_{k-m})\mathrm{E}(w_{i-l}w_{k-p}).
\end{aligned}$$

Excluding the cases where one or other of these covariances is zero implies the further restrictions $m = j + k - i$ and $l = p + i - k$, which require respectively that $j \geq i - k$ and that $p \geq k - i$. The summands then do not depend on i and k independently, but only on $i - k$. Setting $q = i - k$, the sum has the form

$$\frac{1}{n^2}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}\sum_{p=0}^{\infty}\sum_{i=1}^{n}\sum_{k=1}^{n} b_j b_m c_l c_p \mathrm{E}(u_{i-j}u_{k-m})\mathrm{E}(w_{i-l}w_{k-p})$$

$$
\begin{aligned}
&= \frac{\sigma_u^2\sigma_w^2}{n^2}\sum_{q=1-n}^{n-1}\sum_{j=\max\{0,q\}}^{\infty}\sum_{p=\max\{0,-q\}}^{\infty} b_j b_{j-q} c_p c_{p+q} \\
&\ll \frac{1}{n^2}\sum_{q=1-n}^{n-1}\sum_{j=\max\{0,q\}}^{\infty} j^{2d_x-2}\Big(1-\frac{q}{j}\Big)^{d_x-1}\sum_{p=\max\{0,-q\}}^{\infty} p^{2d_y-2}\Big(1+\frac{q}{p}\Big)^{d_y-1} \\
&\ll \frac{1}{n^2}\sum_{q=0}^{n-1}(q^{2d_x-1}+q^{2d_y-1}) \\
&= O(n^{-1}). \quad \blacksquare
\end{aligned}
$$

A key application of this result is to the case $y_i = x_i$, with the limit in (1.5). It is proved under weak dependence of the shocks as Theorem **8.5**.

The next theorem, applying only to cases with $y_i \neq x_i$, deals with the distribution of the empirical covariance when the shocks are contemporaneously as well as serially independent.

4.3 Theorem Under Assumption **4.1**, if $\sigma_{uw} = 0$ and $d_x + d_y < \frac{1}{2}$,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i y_i \xrightarrow{\mathrm{d}} \mathrm{N}(0, V_{xy})$$

where $V_{xy} < \infty$.

Proof

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i y_i = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} b_k c_j\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_{i-k}w_{i-j}\right)$$

where for each j and k, random variables $Z(k,j)$ are defined by

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} u_{i-k}w_{i-j} \xrightarrow{\mathrm{d}} Z(k,j) \overset{\mathrm{d}}{\sim} \mathrm{N}(0,\sigma_u^2\sigma_w^2) \tag{4.7}$$

where the limit distributions hold by (say) the Lindeberg-Lévy central limit theorem. It can be verified that $Z(k,j) = Z(k',j')$ if $j-k = j'-k'$. Moreover, products of the form $u_{i-k}w_{i-j}u_{m-k'}w_{m-j'}$ have zero mean unless both $i-k = m-k'$ and $i-j = m-j'$, which imposes the same condition. In other words, if $j-j' \neq k-k'$ then $\mathrm{E}(Z(k,j)Z(k',j')) = 0$. Being Gaussian, the pairs are independent if they are not identical.

Define $\zeta = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} b_k c_j Z(k,j)$. Being a weighted sum of Gaussian random variables that are either identical or independent, ζ is itself Gaussian if its variance is finite. The pairs $b_k c_j Z(k,j)$ and $b_{k'}c_{j'}Z(k',j')$ are correlated if, for any $p \in \mathbb{N}$, both $k' = k+p$ and $j' = j+p$ and also if both $k = k'+p$ and $j = j'+p$. Summing the corresponding weights, the variance is calculated as

$$V_{xy} = \mathrm{E}(\zeta^2) = \sigma_u^2\sigma_w^2\left(\sum_{k=0}^{\infty} b_k^2\sum_{j=0}^{\infty} c_j^2 + 2\sum_{p=1}^{\infty}\left(\sum_{k=0}^{\infty} b_k b_{k+p}\sum_{j=0}^{\infty} c_j c_{j+p}\right)\right). \tag{4.8}$$

By (1.6), $\sum_{k=0}^{\infty} b_k b_{k+p} \sum_{j=0}^{\infty} c_j c_{j+p} = O(p^{2d_x+2d_y-2})$ and these terms are summable over p under the specified condition. ∎

Note that u and w are not required to be independent of each other in this result. If the products have zero mean and finite variance, serial independence is more than sufficient for the central limit theorem to operate. The corresponding result under weak dependence (but with zero cross-autocorrelations) is proved in Section 8.1 as Theorem **8.6**.

Some interesting implications of the condition $d_x + d_y < \frac{1}{2}$, placing a particular limitation on the scope of Gaussian inference, are discussed in Section 7.3. A natural example is the case where x exhibits stationary long memory and y is weakly dependent, with $d_y = 0$ so that $c_j = 0$ for $j > 0$. (This means serially independent under Assumption **4.1**.) In this case, simply enough, $V_{xy} = \sigma_w^2 \sigma_u^2 \sum_{k=0}^{\infty} b_k^2$. However, a limited degree of long memory in both processes is also compatible with a Gaussian limit.

4.2 The Covariance Decomposition

In the remainder of this chapter, and also in Chapters 5–6 to follow, the case of principal interest is going to be the limiting distribution of the covariance process

$$G_n = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} x_k y_{i+1} \tag{4.9}$$

where

$$K(n) = n^{d_x+d_y} L_x(n) L_y(n). \tag{4.10}$$

An equivalent notation would be $G_n = (nK(n))^{-1} \sum_{i=1}^{n-1} S_i y_{i+1}$ where S_i denotes the partial sum process. Strictly, it would be better to write (4.9) as G_n^{xy} so as to distinguish the case in which the roles of x and y are interchanged. This form of labelling becomes unavoidable in some later developments, but to avoid clutter the distinction is treated as implicit where this does not hinder sense.

The natural context for statistics such as (4.9) is regression analysis and specifically cointegrating regression. These applications are reviewed in Chapter 7. A leading example is the case $y_i = x_i$ which might arise in the context of testing for mean reversion in a partial sum process. In the conventional case of weakly dependent series, absence of such mean reversion is an indicator of a unit root. The aim is to show that the weak limit of G_n exists and for fBMs X and Y is a random variable $\int_0^1 X\mathrm{d}Y$, with known distribution. However, $\int_0^1 X\mathrm{d}Y$ here defined is not in general an Itô integral. In particular, it does not have a mean of zero and the main task undertaken in this chapter is to calculate its mean.

Serial uncorrelatedness of the shocks allows G_n in (4.9) to be split into components according to their expectations. Substituting from (4.1), the sum of products may be expanded as

$$G_n = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} b_j c_l u_{k-j} w_{i+1-l}. \tag{4.11}$$

Let this sum be decomposed as $G_n = G_{1n} + G_{2n} + G_{3n}$ where G_{1n} contains those terms in which $k - j \le i - l$, so that the time indices of w strictly exceed those of u, and in G_{3n} the indices of u lead those of w, so that $k - j > i + 1 - l$. Therefore, G_{2n} has the terms where $k - j = i + 1 - l$ such that the time indices of u and w match. The decomposition can be written out as

$$G_{1n} = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{j=0}^{\infty} \sum_{l=0}^{i+j-k} b_j c_l u_{k-j} w_{i+1-l} \tag{4.12}$$

$$G_{2n} = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{j=0}^{\infty} b_j c_{i+j-k+1} u_{k-j} w_{k-j} \tag{4.13}$$

and

$$G_{3n} = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{j=0}^{\infty} \sum_{l=i+j-k+2}^{\infty} b_j c_l u_{k-j} w_{i+1-l}. \tag{4.14}$$

With $\mathrm{E}(G_{1n}) = \mathrm{E}(G_{3n}) = 0$ under Assumption **4.1**, the first objective is to obtain a formula for the limiting mean of G_{2n}.

4.4 Theorem If Assumption **4.1** holds and $d_x + d_y > 0$, $\mathrm{E}(G_{2n}) \to \sigma_{uw}\lambda_{xy}$ as $n \to \infty$ where

$$\lambda_{xy} = \frac{1}{d_x + d_y}\Bigg(\frac{d_y}{1 + d_x + d_y} + \int_0^\infty \Big(d_y(1+\tau)^{d_x+d_y} + d_x\tau^{d_x+d_y} - (d_x + d_y)(1+\tau)^{d_y}\tau^{d_x}\Big)\mathrm{d}\tau\Bigg). \tag{4.15}$$

Proof It can be verified by inspection that

$$\sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{j=0}^{\infty} b_j c_{i+j-k+1} = \sum_{i=1}^{n-1} (n-i) \sum_{j=0}^{\infty} b_j c_{j+i} = \sum_{i=1}^{n-1} \sum_{k=0}^{\infty} \sum_{j=\max\{0,k-i+1\}}^{k} b_j c_{k+1}.$$

The sum of the b_j in the right-hand member is equal to $a_{n,i-k}(i/n, 0)$ as defined in (2.25), a sum containing either $k+1$ terms or i terms whichever is the smaller. Therefore, setting $\mathrm{E}(u_{k-j}w_{k-j}) = \sigma_{uw}$ for all j and k in (4.13), it is found that

$$\mathrm{E}(G_{2n}) = \frac{\sigma_{uw}}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=0}^{\infty} a_{n,i-k}(i/n, 0) c_{k+1}. \tag{4.16}$$

To obtain the limit of this sum, separate the terms into two blocks, for $k = 0, \ldots, i-1$ and $k \ge i$ respectively. Applying (2.29) and also (1.2) while noting that the slowly varying components cancel in the limit and that $d_x + d_y > 0$ by assumption, the first block gives

$$\frac{\sigma_{uw}}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=0}^{i-1} a_{n,i-k}(i/n, 0) c_{k+1} \sim \frac{\sigma_{uw} d_y}{n^2} \sum_{i=1}^{n-1} \sum_{k=0}^{i-1} \Big(\frac{k+1}{n}\Big)^{d_x+d_y-1}$$

$$\to \sigma_{uw} d_y \int_0^1 \int_0^\tau \xi^{d_x+d_y-1} \mathrm{d}\xi \mathrm{d}\tau$$
$$= \frac{\sigma_{uw} d_y}{(d_y+d_x)(1+d_y+d_x)}. \tag{4.17}$$

Applying (2.30), the second block gives

$$\frac{\sigma_{uw}}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=i}^{\infty} a_{n,i-k}(i/n,0) c_{k+1}$$
$$\sim \frac{\sigma_{uw} d_y}{n^2} \sum_{i=1}^{n-1} \sum_{k=0}^{\infty} \left(\left(\frac{i+k}{n} \right)^{d_x} - \left(\frac{k}{n} \right)^{d_x} \right) \left(\frac{i+k}{n} \right)^{d_y-1}$$
$$\to \sigma_{uw} d_y \int_0^\infty \int_0^1 ((\xi+\tau)^{d_x} - \tau^{d_x})(\xi+\tau)^{d_y-1} \mathrm{d}\xi \mathrm{d}\tau$$
$$= \frac{\sigma_{uw}}{(d_x+d_y)} \int_0^\infty \Big(d_y (1+\tau)^{d_x+d_y} + d_x \tau^{d_x+d_y}$$
$$- (d_x+d_y)(1+\tau)^{d_y} \tau^{d_x} \Big) \mathrm{d}\tau. \tag{4.18}$$

Combining these two limits completes the proof for the cases with $d_y \neq 0$.

If $d_y = 0$ then $\lambda_{xy} = 0$ and the explicit representations used in (4.17) and (4.18) do not hold. One possibility is $c_k = 0$ for $k > 0$, but to assume $c_k = O(k^{-1-\delta})$ for $\delta > 0$ is asymptotically equivalent. In the latter case,

$$\sum_{k=0}^{\infty} a_{n,i-k}(i/n,0) c_{k+1} = \left(\sum_{k=0}^{i-1} + \sum_{k=i}^{\infty} \right) \sum_{j=\max\{0,k-i+1\}}^{k} b_j c_{k+1}$$

where

$$\sum_{k=0}^{i-1} \sum_{j=0}^{k} b_j c_{k+1} = O(i^{d_x-\delta} L_x(i))$$

and also

$$\sum_{k=i}^{\infty} \sum_{j=k-i+1}^{k} b_j c_{k+1} = O(i^{d_x-\delta} L_x(i)).$$

Hence, with $K(n) = n^{d_x} L_x(n)$,

$$\frac{\sigma_{uw}}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=0}^{\infty} a_{n,i-k}(i/n,0) c_{k+1} = O(n^{-\delta})$$

confirming that $\mathrm{E}(G_{2n})$ vanishes in the limit. Expression (4.15) is therefore formally correct in all cases. ∎

This result does not impose the requirement that b_j and c_l depend exclusively on parameters d_x and d_y, but any additional parameters must enter via the functions L_x and L_y and hence are cancelled.

At this point, let λ_{yx} represent the counterpart of (4.15) with the roles of x and y interchanged so that d_x and d_y are swapped in the formula. Then $\lambda_{xy} + \lambda_{yx} = \Upsilon_{xy}$, where it is easily verified that

$$\Upsilon_{xy} = \frac{1}{1+d_x+d_y} + \int_0^\infty \left((1+\tau)^{d_x} - \tau^{d_x}\right)\left((1+\tau)^{d_y} - \tau^{d_y}\right) \mathrm{d}\tau \qquad (4.19)$$

and, needless to say, $\Upsilon_{yx} = \Upsilon_{xy}$. This expression exists whether or not $d_x+d_y > 0$, and reduces to 1 if $d_x = d_y = 0$. It is the bivariate case of Υ_{kl} in (2.49), the element of the long-run covariance matrix of the process (x_i, y_i), such that

$$\sigma_{uw}\Upsilon_{xy} = \lim_{n\to\infty} \frac{1}{nK(n)} \mathrm{E}\left(\sum_{i=1}^n x_i \sum_{i=1}^n y_i\right) \qquad (4.20)$$

where

$$\sum_{i=1}^n x_i \sum_{i=1}^n y_i = \sum_{i=1}^n x_i y_i + \sum_{i=1}^{n-1}\sum_{k=1}^{i} x_k y_{i+1} + \sum_{i=1}^{n-1}\sum_{k=1}^{i} y_k x_{i+1}. \qquad (4.21)$$

An important special case of these calculations is $y = x$, for which it is found that $\lambda_{xx} = \frac{1}{2}\Upsilon_{xx}$ and of course that $\Upsilon_{xx} = \Upsilon_{d_x}$ in the notation for the univariate case in (2.4). The expectation of the first right-hand-side term in (4.21) is $n\sigma_{xy}$ where σ_{xy} is finite according to (4.3), whereas after dividing by $nK(n)$ the expectations of the second and third terms converge respectively to $\sigma_{uw}\lambda_{xy}$ and $\sigma_{uw}\lambda_{yx}$. With $d_x + d_y > 0$ the contemporaneous component of the covariance therefore vanishes under the normalization by $nK(n)$, which explains why Υ_{xy} can be decomposed into two complementary terms. This is a key distinction between the distributions of long memory processes and the weakly dependent case.

4.3 Closed Forms

The expressions in (4.15) and (4.19) are intuitive and easily interpreted, but for computational purposes closed forms have an advantage. These can be obtained as follows, by an argument that closely parallels Theorem **2.2**.

4.5 Theorem Under Assumption **4.1**(a) and if $d_x + d_y > 0$,

$$\lambda_{xy} = \Gamma(d_x+1)\Gamma(d_y+1)\frac{\Gamma(1-d_x-d_y)\sin\pi d_y}{\pi\,(1+d_x+d_y)\,(d_x+d_y)}. \qquad (4.22)$$

Proof Let the integrand in (4.15) be written with a complex argument, as

$$\begin{aligned}\phi_{xy}(z) &= d_y(1+z)^{d_x+d_y} + d_x z^{d_x+d_y} - (d_x+d_y)(1+z)^{d_y} z^{d_x} \\ &= d_y(1+z)^{d_y}\left((1+z)^{d_x} - z^{d_x}\right) - d_x z^{d_x}\left((1+z)^{d_y} - z^{d_y}\right) \qquad (4.23)\end{aligned}$$

for $z \in \mathbb{C}$. Defining

$$z(\xi) = \begin{cases} \xi, & \xi \geq 0 \\ \mathrm{e}^{\mathrm{i}\pi}|\xi|, & \xi < 0 \end{cases}$$

similarly to (2.8), $f_{xy} = \phi_{xy} \circ z$ is a well-defined function of a real variable $\xi \in \mathbb{R}$. In (4.23) note that as z increases, $(1+z)^{d_x} - z^{d_x}$ approaches $d_x z^{d_x-1}$ and $(1+z)^{d_y} - z^{d_y}$ approaches $d_y z^{d_y-1}$, so that $f_{xy}(\xi) = O(\xi^{d_x+d_y-2})$ as $\xi \to \pm\infty$. Hence, f_{xy} is certainly integrable if $d_x > 0$ and $d_y > 0$. If $d_x < 0$, f_{xy} possesses a singularity at $\xi = 0$ but $f_{xy}(\xi)/z(\xi)^{d_x} \to -(d_x + d_y)$ as $\xi \to 0$. Similarly, if $d_y < 0$ then f_{xy} possesses a singularity at $\xi = -1$, but $f_{xy}(\xi)/(1 + z(\xi))^{d_y} \to -(d_x + d_y)\mathrm{e}^{\mathrm{i}\pi d_x}$ as $\xi \to -1$. Therefore f_{xy} is integrable for all $\xi \in \mathbb{R}$.

Let f_{yx} denote the formula complementary to f_{xy} with y and x interchanged. Introduce the change of variable $\tau = -1 - \xi$ so that similarly to (2.9), $z(\xi) = \mathrm{e}^{\mathrm{i}\pi}(1 + z(\tau))$ and $1 + z(\xi) = \mathrm{e}^{\mathrm{i}\pi} z(\tau)$. These relations imply according to (4.23) that

$$f_{yx}(\xi) = \mathrm{e}^{\mathrm{i}\pi(d_x+d_y)} f_{xy}(\tau). \tag{4.24}$$

By symmetry of the formulae, $f_{xy}(\tau) = \mathrm{e}^{\mathrm{i}\pi(d_x+d_y)} f_{yx}(\xi)$ likewise and it follows that $f_{xy}(\xi) = \mathrm{e}^{2\mathrm{i}\pi(d_x+d_y)} f_{xy}(\xi)$. (This makes sense noting $\mathrm{e}^{2\mathrm{i}\pi} = 1$.) Hence, it must be the case that

$$\int_{-\infty}^{\infty} f_{xy}(\xi)\mathrm{d}\xi = 0 \tag{4.25}$$

and equally that

$$\int_{-\infty}^{\infty} f_{yx}(\tau)\mathrm{d}\tau = 0. \tag{4.26}$$

Now integrate f_{yx} over the three regions of the line, $(0, \infty)$, $(-1, 0)$, and $(-\infty, -1)$, knowing from (4.26) that the sum of the segments is zero. Let the first segment be denoted

$$\mathcal{L}_{yx} = \int_{0}^{\infty} f_{yx}(\xi)\mathrm{d}\xi. \tag{4.27}$$

Let $\mathcal{L}_{xy} = \int_0^\infty f_{xy}(\xi)\mathrm{d}\xi$, the integral appearing in (4.15) whose evaluation is the objective. With $\tau = -\xi - 1$, in view of (4.24) the third segment is

$$\int_{-\infty}^{-1} f_{yx}(\xi)\mathrm{d}\xi = \int_{0}^{\infty} f_{yx}(\tau)\mathrm{d}\tau = \mathrm{e}^{\mathrm{i}\pi(d_x+d_y)}\mathcal{L}_{xy}. \tag{4.28}$$

The second segment rearranges in a manner closely similar to (2.10) as

$$\begin{aligned}\int_{-1}^{0} f_{yx}(\xi)\mathrm{d}\xi &= -d_x \int_0^1 (1-\xi)^{d_x+d_y}\mathrm{d}\xi + d_y \int_0^1 (\mathrm{e}^{\mathrm{i}\pi}\xi)^{d_x+d_y}\mathrm{d}\xi \\ &\qquad - (d_x + d_y)\int_0^1 (1-\xi)^{d_x}(\mathrm{e}^{\mathrm{i}\pi}\xi)^{d_y}\mathrm{d}\xi \\ &= \frac{d_x + \mathrm{e}^{\mathrm{i}\pi(d_x+d_y)}d_y}{1 + d_x + d_y} - (d_x + d_y)\mathrm{e}^{\mathrm{i}\pi d_y}\int_0^1 (1-\xi)^{d_x}\xi^{d_y}\mathrm{d}\xi \end{aligned} \tag{4.29}$$

where the integral in the final term is the Beta function $B(d_y+1, d_x+1)$ according to (B.14) of Appendix B. Adding together (4.27), (4.28), and (4.29) yields

$$\mathcal{L}_{yx} + \frac{d_x + \mathrm{e}^{\mathrm{i}\pi(d_x+d_y)}d_y}{1 + d_x + d_y} - (d_x+d_y)\mathrm{e}^{\mathrm{i}\pi d_y}B(d_x+1, d_y+1) + \mathrm{e}^{\mathrm{i}\pi(d_x+d_y)}\mathcal{L}_{xy} = 0. \tag{4.30}$$

The critical step is to note that this equality remains true if x and y are everywhere interchanged, the Beta function being symmetric in its arguments. Thus,

$$\mathcal{L}_{xy}+\frac{d_y+\mathrm{e}^{\mathrm{i}\pi(d_y+d_x)}d_x}{1+d_y+d_x}-(d_y+d_x)\mathrm{e}^{\mathrm{i}\pi d_x}B(d_y+1,d_x+1)+\mathrm{e}^{\mathrm{i}\pi(d_y+d_x)}\mathcal{L}_{yx}=0. \quad (4.31)$$

The procedure is to subtract (4.30) multiplied by $\mathrm{e}^{\mathrm{i}\pi(d_x+d_y)}$ from (4.31), so eliminating $\mathcal{L}_{yx}$, and then to solve the resulting equation for $\mathcal{L}_{xy}$. After cancellation and rearrangement this operation gives

$$\mathcal{L}_{xy}=\frac{\mathrm{e}^{\mathrm{i}\pi d_x}(1-\mathrm{e}^{2\mathrm{i}\pi d_y})}{1-\mathrm{e}^{2\mathrm{i}\pi(d_x+d_y)}}(d_x+d_y)\,B\,(d_x+1,d_y+1)-\frac{d_y}{1+d_x+d_y}. \quad (4.32)$$

It remains to simplify this expression. Successively, identities (B.2), (B.9), (B.11), and (B.15) give

$$\begin{aligned}\frac{\mathrm{e}^{\mathrm{i}\pi d_x}(1-\mathrm{e}^{2\mathrm{i}\pi d_y})}{1-\mathrm{e}^{2\mathrm{i}\pi(d_x+d_y)}}&=\frac{\mathrm{e}^{\mathrm{i}\pi d_x}(1-\mathrm{e}^{2\mathrm{i}\pi d_y})(1-\mathrm{e}^{-2\mathrm{i}\pi(d_x+d_y)})}{2(1-\cos 2\pi(d_x+d_y))}\\&=\frac{2\cos\pi d_x-2\cos\pi(d_x+2d_y)}{4\sin^2\pi(d_x+d_y)}\\&=\frac{\sin\pi d_y}{\sin\pi\,(d_x+d_y)}\\&=\frac{\Gamma(d_x+d_y)\Gamma(1-d_x-d_y)\sin\pi d_y}{\pi}.\end{aligned}$$

Also, by (B.14) and (B.13),

$$B\,(d_x+1,d_y+1)=\frac{\Gamma(d_x+1)\Gamma(d_y+1)}{(d_x+d_y)(1+d_x+d_y)\Gamma(d_x+d_y)}$$

Substituting these formulae into (4.32) and simplifying gives

$$\mathcal{L}_{xy}=\Gamma(d_x+1)\Gamma(d_y+1)\frac{\Gamma(1-d_x-d_y)\sin\pi d_y}{\pi(1+d_x+d_y)}-\frac{d_y}{1+d_x+d_y}.$$

Finally, substituting this expression into (4.15) gives (4.22). ∎

Directly from (4.22),

$$\begin{aligned}\Upsilon_{xy}&=\lambda_{xy}+\lambda_{yx}\\&=\Gamma(d_x+1)\Gamma(d_y+1)\frac{\Gamma(1-d_x-d_y)(\sin\pi d_y+\sin\pi d_x)}{\pi\,(1+d_x+d_y)\,(d_x+d_y)}.\end{aligned} \quad (4.33)$$

The scale factors $\Gamma(d_x+1)$ and $\Gamma(d_y+1)$ appearing in (4.22) and (4.33) are optional and as remarked on page 23, they would disappear from the formulae if the moving average coefficients were defined explicitly to include them as divisors. This alternative would be implemented in effect by replacing (4.10) with

$$K(n)=n^{d_x+d_y}\Gamma(d_x+1)\Gamma(d_y+1)L_x(n)L_y(n).$$

Economy of notation is the natural criterion for deciding which convention to prefer.

Formula (4.33) has been derived on the assumption $d_x + d_y > 0$ and returns an awkward 'zero over zero' result for the case $d_x = -d_y$, even though expression (4.19) is well defined for $d_x + d_y \le 0$. There is an equivalent and more robust version of the formula that has already been quoted as expression (2.53), in connection with Theorem **2.9**. This equivalence is shown as follows.

4.6 Theorem

$$\Upsilon_{xy} = \frac{\Gamma(d_x+1)\Gamma(d_y+1)\cos(\pi(d_x-d_y)/2)}{\Gamma(d_x+d_y+2)\cos(\pi(d_x+d_y)/2)}. \tag{4.34}$$

Proof Applying successively identities (B.8), (B.10), and (B.15) gives the relation

$$\begin{aligned}\sin\pi d_y + \sin\pi d_x &= \frac{\sin\pi(d_x+d_y)(\sin\pi d_y + \sin\pi d_x)}{2\sin(\pi(d_x+d_y)/2)\cos(\pi(d_x+d_y)/2)} \\ &= \sin\pi(d_x+d_y)\frac{\cos(\pi(d_x-d_y)/2)}{\cos(\pi(d_x+d_y)/2)} \\ &= \frac{\pi}{\Gamma(d_x+d_y)\Gamma(1-d_x-d_y)}\frac{\cos(\pi(d_x-d_y)/2)}{\cos(\pi(d_x+d_y)/2)}. \end{aligned} \tag{4.35}$$

Substituting (4.35) into (4.33), replacing $\Gamma(d_x+d_y)(1+d_x+d_y)(d_x+d_y)$ by $\Gamma(d_x+d_y+2)$ according to (B.13) and then cancelling the matching terms in the ratio gives finally (4.34). ∎

Formula (4.34) might also be obtained by operating directly on (4.19) in the manner of Theorem **4.5**. However, an alternative direct proof based on the harmonizable representation of the processes is given as Theorem **9.3** in Chapter 9.

4.4 Antipersistence

The condition $d_x + d_y > 0$ must hold to define λ_{xy} and λ_{yx}, without which at least one of the processes is antipersistent. The following results give an indication of what happens in this latter case. The cases $d_x + d_y < 0$ and $d_x + d_y = 0$ require separate treatment, while $d_x = -d_y \neq 0$ must also be distinguished from $d_x = d_y = 0$, which is the conventional unit root scenario.

4.7 Theorem

(i) If $d_x = -d_y \neq 0$ and $\sigma_{uw} \neq 0$ then $\mathrm{E}(G_{2n}) = O(\log n)$.

(ii) If $d_x + d_y < 0$ and $\sigma_{uw} \neq 0$ then $\mathrm{E}(G_{2n}) = O(n^{-d_x-d_y})$.

Proof For part (i), consider expression (4.16) and its decomposition into the terms with $k < i$ in (4.17) and those with $k \ge i$ in (4.18). With $d_x + d_y = 0$ the second member of (4.17) takes the form

$$\frac{\sigma_{uw}d_y}{n}\sum_{i=1}^{n-1}\sum_{k=1}^{i}k^{-1} = O\biggl(\frac{1}{n}\sum_{i=1}^{n-1}\log i\biggr) = O(\log n) \tag{4.36}$$

where the second equality applies Stirling's approximation (B.16). For the terms with $k \geq i$, consider the second member of (4.18). Applying a Taylor's expansion of first order around $(i+k)/n$ gives, with $d_x + d_y = 0$,

$$\begin{aligned}\frac{\sigma_{uw}d_y}{n^2}&\sum_{i=1}^{n-1}\sum_{k=0}^{\infty}\biggl(\Bigl(\frac{i+k}{n}\Bigr)^{d_x} - \Bigl(\frac{k}{n}\Bigr)^{d_x}\biggr)\Bigl(\frac{i+k}{n}\Bigr)^{d_y-1}\\ &\sim \frac{\sigma_{uw}d_yd_x}{n^2}\sum_{i=1}^{n-1}\sum_{k=0}^{\infty}\Bigl(\frac{i+k}{n}\Bigr)^{-2}\frac{i}{n}\\ &= \sigma_{uw}d_yd_x\sum_{i=1}^{n-1}i^{-1}\biggl(\frac{1}{n}\sum_{k=0}^{\infty}\Bigl(\frac{i}{i+k}\Bigr)^2\biggr) = O(\log n).\end{aligned} \tag{4.37}$$

The indicated order of magnitude in (4.37) is verified by showing that the term in parentheses in the penultimate member is $O(1)$ as $i \to n$ and $n \to \infty$. Since $i < n$,

$$\frac{1}{n}\sum_{k=0}^{\infty}\Bigl(\frac{i}{i+k}\Bigr)^2 \leq \sum_{k=0}^{\infty}k^{-1-\delta}\Bigl(\frac{i^{1/2}k^{(1+\delta)/2}}{i+k}\Bigr)^2$$

where $0 < \delta < 1$. The squared term in parentheses converges to zero as $k \to \infty$ for any fixed i, and also as $i \to \infty$.

For part (ii), the condition $d_x + d_y < 0$ means that for each $i < n$ the terms $a_{n,i-k}(i/n,0)c_{k+1} = O(k^{d_x+d_y-1})$ are summable over k. In view of (2.25), these sums have the form

$$\sum_{k=0}^{\infty}a_{n,i-k}(i/n,0)c_{k+1} = \sum_{k=0}^{i-1}\biggl(\sum_{j=0}^{k}b_j\biggr)c_{k+1} + \sum_{k=i}^{\infty}\biggl(\sum_{j=k-i+1}^{k}b_j\biggr)c_{k+1}.$$

The second block of terms vanishes as $i \to \infty$. The first block has a finite limit as $i \to \infty$ but does not vanish, given $b_0 = 1$ and $d_y \neq 0$. The conclusion is that

$$\frac{1}{n}\sum_{i=1}^{n-1}\sum_{k=0}^{\infty}a_{n,i-k}(i/n,0)c_{k+1} = O(1).$$

The result follows in view of the normalization $nK(n)$ defining G_n in (4.11). ∎

In the case $d_x + d_y < 0$, the summability means that all three terms of the decomposition (4.21) are $O_p(n)$. It can generally be assumed that there exists in this case a finite constant

$$\gamma_{xy} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n-1}\sum_{k=1}^{i}\mathrm{E}(x_ky_{i+1}) \tag{4.38}$$

and a complementary quantity γ_{yx}. However, there are no general formulae for γ_{xy} and γ_{yx} to compare with $\sigma_{uw}\lambda_{xy}$ and $\sigma_{uw}\lambda_{yx}$ because due to the summability of the terms in (4.38) they must depend on low-order lag coefficients which are not restricted under (1.2). The one thing known is that $\mathrm{E}(\sum_{i=1}^{n} x_i)^2 = O(n^{2d_x+1})$ and $\mathrm{E}(\sum_{i=1}^{n} y_i)^2 = O(n^{2d_y+1})$ according to Corollary **2.7**, so with $d_x + d_y < 0$ the expected left-hand side of (4.21) is $o(n)$ by the Cauchy-Schwarz inequality. That is to say,

$$\sigma_{xy} + \gamma_{xy} + \gamma_{yx} = 0. \tag{4.39}$$

In the particular case $y = x$, so that necessarily $d_x < 0$, $\gamma_{xx} = \sum_{k=1}^{\infty} \gamma_k$ where γ_k is defined for x in (1.6). The relation in (4.39) is then identical to that represented in equation (1.18). For the fractionally differenced model represented by (1.12) with $d_x < 0$, the γ_k are negative for every $k \geq 1$ according to (1.15). In this particular case, in view of (4.39) it follows from (1.17) that

$$\gamma_{xx} = -\frac{\sigma_u^2 \Gamma(1 - 2d_x)}{2\Gamma(1 - d_x)^2}.$$

The condition $\gamma_{xx} < 0$ reflects the fact that the limiting partial sum process (that is to say, $X(t)$ defined in (2.1)) is 'mean-reverting' in the sense that its increments tend to be negatively correlated and there exists a tendency towards zero when $1 + d_x < 1$. In the hierarchy of time dependence, the partial sum process with antipersistent increments falls short of the unit root process ($d_x = 0$). As noted, Υ_{xy} is well defined for $d_x + d_y \leq 0$ in spite of (4.39), the difference being the choice of normalization. Under the normalization by n the covariance vanishes, but under normalization by $nK(n)$ the limit in (4.20) exists, as shown in Theorem **2.9**.

However, all these conclusions depend on the assumption $\sigma_{uw} \neq 0$. If u_i and w_i are contemporaneously uncorrelated (implying under Assumption **4.1** that the cross-correlogram is zero at all orders) then of course each of the terms in (4.21) has zero expectation.

4.5 L_2 Convergence

An important fact is that G_{2n} is a consistent estimator of its mean, albeit not a feasible one. The following result implies that the limit distribution of $G_{1n} + G_{3n}$ matches that of the mean deviation of G_n, not overlooking that the mean diverges under the given normalization when $d_x + d_y < 0$ according to Theorem **4.7**. The present assumption of independent shocks simplifies the argument somewhat. The extension to the weak dependence case is Theorem **8.8**.

4.8 Theorem Under Assumption **4.1**, $G_{2n} - \mathrm{E}(G_{2n}) \to_{L_2} 0$.

Proof From (4.13),

$$G_{2n} - \mathrm{E}(G_{2n}) = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} P_{ik} \tag{4.40}$$

where

$$P_{ik} \equiv \sum_{j=0}^{\infty} b_j c_{i+1-k+j}(u_{k-j}w_{k-j} - \sigma_{uw}). \tag{4.41}$$

The square of $nK(n)(G_{2n} - \mathrm{E}(G_{2n}))$ is the summed elements of the outer product of the $\frac{1}{2}n(n-1)$-vector having elements $\{P_{ik}, k = 1, \dots i, i = 1, \dots, n-1\}$. It can be verified that

$$\mathrm{E}(G_{2n} - \mathrm{E}(G_{2n}))^2 \le \frac{2}{n^2K(n)^2} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{m=0}^{i-1} \sum_{p=k-i+m}^{k-1} |\mathrm{E}(P_{ik}P_{i-m,k-p})|. \tag{4.42}$$

If i and k in this sum count the rows of the matrix lower triangle, the indices $i-m$ and $k-p$ in the majorant of (4.42) fill in the column elements, with $i-m$ running from 1 to i and $k-p$ running from 1 to $i-m$. Note that the index p can take either sign subject to the inequality $i-m \ge k-p$. There is some double-counting of terms having $m = 0$, but to show that the majorant of (4.42) is of small order in n suffices for the proof.

The terms of the sum have a general bound of the form

$$\begin{aligned} |\mathrm{E}(P_{ik}P_{i-m,k-p})| \le \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \big| b_j b_l c_{i+1-k+j} c_{i-m+1-k+p+l} \\ \times \mathrm{E}(u_{k-j}u_{k-p-l}w_{k-j}w_{k-p-l} - \sigma_{uw}^2)\big|. \end{aligned} \tag{4.43}$$

Under serial independence of the shocks, the expectations in (4.43) vanish unless $j = p + l$ and so this expression reduces to

$$|\mathrm{E}(P_{ik}P_{i-m,k-p})| \le (\mu_{uw}^4 - \sigma_{uw}^2) \sum_{j=0}^{\infty} \big| b_j b_{j-p} c_{i+1-k+j} c_{i-m+1-k+j} \big|. \tag{4.44}$$

To bound (4.42), the first step is to divide the sum into two components according to the sign of p. In other words, write

$$\mathrm{E}(G_{2n} - \mathrm{E}(G_{2n}))^2 \le A_{1n} + A_{2n} \tag{4.45}$$

where A_{1n} has the terms in which $0 \le p \le k-1$ in the innermost sum of (4.42), while A_{2n} contains the remaining terms for the values of m having $k-i+m \le p < 0$, where these exist. Let A_{1n} be further decomposed by splitting the sum in (4.44) into the sums of the first k terms and of the rest so as to write for the case $p \ge 0$ (noting $b_{j-p} = 0$ when $j < p$),

$$\begin{aligned} |\mathrm{E}(P_{ik}P_{i-m,k-p})| &\le (\mu_{uw}^4 - \sigma_{uw}^2)\left(\sum_{j=p}^{k-1} + \sum_{j=k}^{\infty}\right) |b_j b_{j-p} c_{i+1-k+j} c_{i-m+1-k+j}| \\ &= B_{11} + B_{12}. \end{aligned} \tag{4.46}$$

This gives the decomposition

$$A_{1n} = \frac{2}{n^2K(n)^2}\sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{m=0}^{i-1}\sum_{p=\max\{0,k-i+m\}}^{k-1}(B_{11}+B_{12}) = A_{11n}+A_{12n}. \quad (4.47)$$

Similarly for A_{2n}, for those cases of (4.44) contributing to the sum with $k < i-m$ and hence $p < 0$, decompose the sum as

$$\begin{aligned}
&|\mathrm{E}(P_{ik}P_{i-m,k-p})| \\
&\quad = (\mu_{uw}^4 - \sigma_{uw}^2)\left(\sum_{j=0}^{i-m-k} + \sum_{j=i-m-k+1}^{\infty}\right)|b_jb_{j-p}c_{i+1-k+j}c_{i-m+1-k+j}| \\
&\quad = B_{21}+B_{22}. \qquad (4.48)
\end{aligned}$$

The sum over m in (4.42) in this case takes the upper limit of $i-k-1$, so that

$$A_{2n} = \frac{2}{n^2K(n)^2}\sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{m=0}^{i-k-1}\sum_{p=k-i+m}^{-1}(B_{21}+B_{22}) = A_{21n}+A_{22n}. \quad (4.49)$$

This is an empty sum in the cases with $k = i$ and then is assigned the value 0.

The argument now proceeds by bounding each of the terms, A_{11n}, A_{12n}, A_{21n}, and A_{22n}, as functions of n. Substituting from (1.2) but omitting the slowly varying components, which will be cancelled in the limit of (4.47), also assigning the value $b_0 = 1$ to 0^{d_x-1},

$$\begin{aligned}
B_{11} &\ll \sum_{j=p}^{k-1} j^{d_x+2d_y-3}(j-p)^{d_x-1}\left(1+\frac{i+1-k}{j}\right)^{d_y-1} \\
&\qquad\qquad \times\left(1+\frac{i-m+1-k}{j}\right)^{d_y-1} \\
&\ll (k-p)^{2d_x+2d_y-3}. \qquad (4.50)
\end{aligned}$$

The second inequality of (4.50) is valid since $i+1-k > 0$ and $i-m+1-k > 0$, so the terms in parentheses in the second member with exponents $d_y - 1 < 0$ are both smaller than 1. Similarly, j can be replaced by $j-p$ without diminishing the sum since its exponent is negative. Inserting the bound in (4.50) in A_{11n} in place of B_{11} and bounding the sums by integral approximation gives

$$\begin{aligned}
n^2K(n)^2A_{11n} &\ll \sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{m=0}^{i-1}\sum_{p=0}^{k-1}(k-p)^{2d_x+2d_y-3} \\
&\ll \sum_{i=1}^{n-1} i\sum_{k=1}^{i} k^{2d_x+2d_y-2} \\
&\ll \sum_{i=1}^{n-1} i^{2d_x+2d_y} = O(n^{2d_x+2d_y+1}). \qquad (4.51)
\end{aligned}$$

In B_{12} on the other hand, $j+1-k>0$ and $i>m\geq 0$ and so, similarly,

$$\begin{aligned} B_{12} &\ll \sum_{j=k}^{\infty} j^{d_x-1}(j-p)^{d_x-1} i^{d_y-1}\left(1+\frac{1-k+j}{i}\right)^{d_y-1} \\ &\qquad\qquad \times (i-m)^{d_y-1}\left(1+\frac{1-k+j}{i-m}\right)^{d_y-1} \\ &\ll (k-p)^{2d_x-1} i^{d_y-1}(i-m)^{d_y-1} \end{aligned} \tag{4.52}$$

where in this case the terms in large parentheses with negative exponents are vanishing as $j\to\infty$. Substituting the bounding terms of (4.52) into A_{12n} gives the bound

$$\begin{aligned} n^2K(n)^2A_{12n} &\ll \sum_{i=1}^{n-1}\sum_{k=1}^{i} i^{d_y-1}\sum_{m=0}^{i-1}(i-m)^{d_y-1}\sum_{p=0}^{k-1}(k-p)^{2d_x-1} \\ &\ll \sum_{i=1}^{n-1} i^{d_y-1}\sum_{k=1}^{i} i^{d_y}k^{2d_x} = O(n^{2d_x+2d_y+1}). \end{aligned} \tag{4.53}$$

According to (4.47) it follows that $A_{1n}=O(n^{-1})$.

Next, consider A_{2n} for those cases with $i>k$. The term corresponding to (4.50) is

$$\begin{aligned} B_{21} &\ll \sum_{j=0}^{i-m-k} j^{d_x+2d_y-3}(j-p)^{d_x-1}\left(1+\frac{i+1-k}{j}\right)^{d_y-1} \\ &\qquad\qquad \times\left(1+\frac{i-m+1-k}{j}\right)^{d_y-1} \\ &\ll (i-m-k)^{2d_x+2d_y-3}. \end{aligned} \tag{4.54}$$

The terms in large parentheses with negative exponents don't exceed 1, as in B_{11}, and in this case $p<0$ so replacing $j-p$ by j does not decrease the bound. Thus,

$$\begin{aligned} n^2K(n)^2A_{21n} &\ll \sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{m=0}^{i-k-1}\sum_{p=k-i+m}^{-1}(i-m-k)^{2d_x+2d_y-3} \\ &\ll \sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{m=0}^{i-k-1}(i-m-k)^{2d_x+2d_y-2} \\ &\ll \sum_{i=1}^{n-1}\sum_{k=1}^{i}(i-k)^{2d_x+2d_y-1} = O(n^{2d_x+2d_y+1}). \end{aligned} \tag{4.55}$$

The term corresponding to (4.52) by similar reasoning is

$$B_{22} \ll \sum_{j=i-m-k+1}^{\infty} j^{d_x-1}(j-p)^{d_x-1} i^{d_y-1}\left(1+\frac{1-k+j}{i}\right)^{d_y-1}$$

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad \times (i-m)^{d_y-1}\left(1+\frac{1-k+j}{i-m}\right)^{d_y-1} \\
&\ll (i-m-k)^{2d_x-1} i^{d_y-1} (i-m)^{d_y-1} \qquad (4.56)
\end{aligned}
$$

and so

$$
\begin{aligned}
n^2 K(n)^2 A_{22n} &\ll \sum_{i=1}^{n-1} i^{d_y-1} \sum_{k=1}^{i} \sum_{m=0}^{i-k-1} (i-m)^{d_y-1} \sum_{p=k-i+m}^{-1} (i-m-k)^{2d_x-1} \\
&\ll \sum_{i=1}^{n-1} i^{d_y-1} \sum_{k=1}^{i} \sum_{m=0}^{i-k-1} (i-m-k)^{2d_x+d_y-1} \\
&\ll \sum_{i=1}^{n-1} i^{d_y-1} \sum_{k=1}^{i} (i-k)^{2d_x+d_y} = O(n^{2d_x+2d_y+1}). \qquad (4.57)
\end{aligned}
$$

The second inequality of (4.57) is valid since the bound is not decreased by replacing $(i-m)^{d_y-1}$ by $(i-m-k)^{d_y-1}$. In view of (4.55) and (4.57), $A_{2n} = O(n^{-1})$ according to (4.49). Hence by (4.45) the proof is complete. ∎

Chapter 5

Stochastic Integrals

5.1 Mean Deviations

The next major objective is to study the asymptotic behaviour of the terms G_{1n} in (4.12) and G_{3n} in (4.14). The limiting forms of these terms will be denoted by Ξ_1 and Ξ_3 respectively and their sum by Ξ. The task of the present chapter is to elucidate the forms of the random variables Ξ_1 and Ξ_3, while Chapter 6 tackles the main business of proving that $G_n - \mathrm{E}(G_n) \to_{\mathrm{d}} \Xi$, subject to Assumption **4.1** and some additional conditions, collected as Assumption **6.9**, which include $d_x + d_y > -\frac{1}{2}$.

The essential first step is to express G_{1n} and G_{3n}, by some judicious rearrangement of terms, as $\mathcal{F}_n$-adapted stochastic processes. For the case of G_{1n} consider the expression in (4.12). For a given k and i, imagine the terms of this sum as entries in a rectangular table with infinitely many rows, with j the row index and l the column index. The first row has $i-k+1$ entries, the second row has $i-k+2$ entries, and so forth. In (4.12) the sum is first by columns giving the inner sum of row elements and then by rows. Interchanging the order of summation over j and l, so that now summation is first by rows over column elements and then by columns, gives

$$G_{1n} = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{l=0}^{\infty} \sum_{j=\max\{0,l+k-i\}}^{\infty} b_j c_l u_{k-j} w_{i+1-l}. \tag{5.1}$$

Next, let $p = k - j$ and $m = i - l$. Making these substitutions to eliminate j and l, note that (5.1) is the same as

$$G_{1n} = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{m=-\infty}^{i} c_{i-m} w_{m+1} \sum_{k=1}^{i} \sum_{p=-\infty}^{\min\{k,m\}} b_{k-p} u_p. \tag{5.2}$$

The final step is again to interchange orders of summation, this time over m and

i, and also over k and p. The result of this rearrangement is

$$G_{1n} = \frac{1}{nK(n)} \sum_{m=-\infty}^{n-1} w_{m+1} \left(\sum_{p=-\infty}^{m} u_p \sum_{i=\max\{1,m\}}^{n-1} c_{i-m} \sum_{k=\max\{1,p\}}^{i} b_{k-p} \right)$$

$$= \frac{1}{\sqrt{n}} \sum_{m=-\infty}^{n-1} q_{nm} w_{m+1} \tag{5.3}$$

where

$$q_{nm} = \frac{1}{\sqrt{n}K(n)} \sum_{p=-\infty}^{m} a_{nmp} u_p \tag{5.4}$$

and, restoring the original subscripts $l = i - m$ and $j = k - p$,

$$a_{nmp} = \sum_{l=\max\{0,1-m\}}^{n-1-m} c_l \left(\sum_{j=\max\{0,1-p\}}^{l+m-p} b_j \right). \tag{5.5}$$

Take care to distinguish this usage of the symbol a from that in (2.25), noting the three subscripts in place of the earlier two.

G_{3n} can be rearranged in much the same manner. Defining $p = k - j$ and $m = i + 1 - l$ (noting that now $m < p < n$) and identifying these subscripts with the variables w_m and u_p gives

$$G_{3n} = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \sum_{p=-\infty}^{k} \sum_{m=-\infty}^{p-1} b_{k-p} c_{i+1-m} u_p w_m.$$

Next, interchanging orders of summation, first over i and k and then over k and p, has the result

$$G_{3n} = \frac{1}{nK(n)} \sum_{p=-\infty}^{n-1} u_p \left(\sum_{m=-\infty}^{p-1} w_m \sum_{k=\max\{p,1\}}^{n-1} b_{k-p} \sum_{i=k}^{n-1} c_{i+1-m} \right)$$

$$= \frac{1}{\sqrt{n}} \sum_{p=-\infty}^{n-1} h_{np} u_p \tag{5.6}$$

where

$$h_{np} = \frac{1}{\sqrt{n}K(n)} \sum_{m=-\infty}^{p-1} e_{npm} w_m \tag{5.7}$$

and (restoring original subscripts)

$$e_{npm} = \sum_{j=\max\{0,1-p\}}^{n-1-p} b_j \left(\sum_{l=j+p+1-m}^{n-m} c_l \right). \tag{5.8}$$

Equations (5.3) and (5.4), and (5.6) and (5.7), show G_{1n} and G_{3n} to have the appearance of sample covariances of a relatively familiar type, apart from the infinite order of the sum. Specifically, these are the covariances between independent processes, w_i and u_i respectively, and moving average processes lagged one period. That their means are zero has been arranged by construction, with the contemporaneous pairs removed to G_{2n}. What is unusual is the construction of the moving average weights. These are in general non-summable and hence affiliated with unit root processes, although here the weights a_{nmp} and e_{npm} are dependent on their position in the sequence rather than being just unity or zero.

To get a feel for the formula for G_{1n}, as the exemplar case, it is of interest to see what happens to (5.5) in the respective cases $d_x = 0$, so that specifically $b_j = 0$ for $j > 0$, and also $d_y = 0$ so that $c_l = 0$ for $l > 0$. In the first case, noting that $p \leq m$, $a_{nmp} = 1 + c_1 + \cdots + c_{n-1-m}$ if $p \geq 1$ and $a_{nmp} = 0$ otherwise. In the second case, $a_{nmp} = b_{\max\{0,1-p\}} + \cdots + b_{m-p}$. In the two cases combined, $a_{nmp} = 1$ for $p = 1, \ldots, m$ and $1 \leq m \leq n-1$, and $a_{nmp} = 0$ otherwise so q_{nm} reduces, as expected, to a random walk initialized at $p = 1$.

5.2 Integral Approximations

The next step is to evaluate the limiting forms as $n \to \infty$ of the arrays $K(n)^{-1}a_{nmp}$ and $K(n)^{-1}e_{npm}$. The following theorem is helpful for connecting the limiting behaviour of the partial sums appearing in (5.5) and (5.8) with the assumptions about the sequences $\{b_j\}$ and $\{c_l\}$.

5.1 Theorem If $b > a \geq 0$, $\alpha > 0$ and $L(n)$ is slowly varying at ∞,

$$\frac{1}{n^\alpha L(n)} \sum_{j=[na]}^{[nb]-1} j^{\alpha-1} L(j) = \frac{b^\alpha - a^\alpha}{\alpha} + o(1). \tag{5.9}$$

Proof $L(j)/L(n) = 1 + o(1)$ for $[na] \leq j \leq [nb]$ and $a > 0$. In the case $a = 0$ there are terms in the sum for which j is finite in the limit and so $L(j)/L(n) \not\to 1$ as $n \to \infty$. However, since the sum in (5.9) diverges these terms are of small order in n and omitting them does not affect the limit. It is therefore valid to conclude that for $\alpha \geq 0$,

$$\frac{1}{n^\alpha L(n)} \sum_{j=[na]}^{[nb]-1} j^{\alpha-1} L(j) = \frac{1}{n} \sum_{j=[na]}^{[nb]-1} \left(\frac{j}{n}\right)^{\alpha-1} + o(1).$$

Taylor's expansions of the function $(\cdot)^\alpha/\alpha$ around j/n specify the relations

$$\frac{1}{\alpha}\left(\frac{j+1}{n}\right)^\alpha = \frac{1}{\alpha}\left(\frac{j}{n}\right)^\alpha + \frac{1}{n}\left(\frac{j}{n}\right)^{\alpha-1} + O(n^{-2}).$$

The proof is completed in view of the telescoping sum,

$$\sum_{j=[na]}^{[nb]-1} \frac{1}{\alpha}\left(\left(\frac{j+1}{n}\right)^\alpha - \left(\frac{j}{n}\right)^\alpha\right) = \frac{1}{\alpha}\left(\left(\frac{[nb]}{n}\right)^\alpha - \left(\frac{[na]}{n}\right)^\alpha\right). \quad \blacksquare$$

A fundamental assumption to be carried through the development in this section is that $d_y \geq 0$ in the discussion of G_{1n}, and also that $d_x \geq 0$ in the treatment of G_{3n}. This is because antipersistence calls for a different treatment of the limit formulae. It is easiest from the point of view of exposition to deal with these two analyses separately and the required variations are to be found in §5.4.

Consider first the case of G_{1n}. A preliminary is to define the functions

$$Z_1^A(t,s) = d_y \int_0^1 \tau^{d_y-1}\Big(1 - \frac{t-1}{t-s}\tau\Big)^{d_x} \mathrm{d}\tau \tag{5.10}$$

for $-\infty < t < 1$ and $-\infty < s < t$, and

$$Z_2^A(t,s) = d_y \int_0^1 \tau^{d_y-1}\Big(1 - \frac{t}{t-s}\tau\Big)^{d_x} \mathrm{d}\tau \tag{5.11}$$

for $-\infty < t < 0$ and $-\infty < s < t$. These are the integral forms of the hypergeometric functions $F(a,b;c;z)$ where $a = -d_x$, $b = d_y$, and $c = d_y + 1$, with $z = (t-1)/(t-s)$ in the case of Z_1^A and $z = t/(t-s)$ in the case of Z_2^A. These functions are represented by the Gauss hypergeometric series for $|z| \leq 1$ (see (B.24)) and are defined by analytic continuation elsewhere. The integrals exist with real arguments if $b > 0$ and $c > b$ and since $z < 0$ in both cases, with $d_y > 0$ the integrals are everywhere well-defined.

5.2 Lemma If $d_y > 0$ and $d_x + d_y > 0$, $K(n)^{-1}a_{n[nt][ns]} = A(t,s) + o(1)$ as $n \to \infty$ for real-valued indices t, s with $-\infty < s \leq t < 1$, where

$$\begin{aligned} A(t,s) = d_y &\int_{\max\{0,-t\}}^{1-t} v^{d_y-1}(v+t-s)^{d_x}\mathrm{d}v \\ &- 1_{\{s<0\}}(-s)^{d_x}\big((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\big). \end{aligned} \tag{5.12}$$

If $d_y > 0$ and $d_x + d_y \leq 0$, the same conclusion holds for $-\infty < s < t < 1$. □

The connection between (5.5) and (5.12) can be made by writing $m = [nt]$ and $p = [ns]$. The pair of symbols s and t, which in Chapters 2 and 3 played the roles of the upper and lower bounds of an interval so that $s > t$, are here re-used in a different context in which $s < t$ is the rule. What might be viewed as connecting the usages is the fact that in both cases s moves about more, while t has the nature of an anchor. In any event, they will change places in the analysis of G_{3n}. The expression in (5.12) vanishes at the point $t = 1$ noting that $a_{nnp} = 0$ identically, involving an empty sum.

Proof of 5.2 Considering the components of equation (5.5), define u and v by $j = [nu]$ and $l = [nv]$. If $d_x > 0$, recalling $b_j \sim d_x j^{d_x-1}L_x(j)$ by (1.2) write

$$\frac{b_{[nu]}}{n^{d_x}L_x(n)} = d_x\frac{u^{d_x-1}}{n} + o(1).$$

By Theorem **5.1**,

$$\frac{1}{n^{d_x}L_x(n)}\sum_{j=\max\{0,1-[ns]\}}^{[nv]+[nt]-[ns]} b_j = (v+t-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x} + o(1). \tag{5.13}$$

For the anti-persistent case in which $-\frac{1}{2} < d_x < 0$, in view of (1.19) it is possible to write

$$\frac{b_{[nu]}}{n^{d_x}L_x(n)} = (u+1/n)^{d_x} - u^{d_x} + o(1)$$

and in this case equality (5.13) follows directly. In the case $d_x = 0$, the assumption of summability gives

$$\sum_{j=\max\{0,1-[ns]\}}^{[nv]+[nt]-[ns]} b_j = \begin{cases} O(1), & s \geq 0 \\ o(1), & s < 0. \end{cases}$$

This is formally equivalent to (5.13) when $L_x(n)$ is a constant not depending on n. The sum could be assigned the limiting value $1-1_{\{s<0\}}$ by choice of normalization.

Similarly,

$$\frac{c_{[nv]}}{n^{d_y}L_y(n)} = d_y\frac{v^{d_y-1}}{n} + o(1)$$

and (5.5) therefore implies $K(n)^{-1}a_{n[nt][ns]} = A(t,s) + o(1)$ with

$$A(t,s) = d_y\int_{\max\{0,-t\}}^{1-t} v^{d_y-1}\Big((v+t-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\Big)\mathrm{d}v \tag{5.14}$$

which matches the formula in (5.12) provided the integral in question exists.

There are three cases to consider. If $d_x > 0$ then $v^{d_y-1}(v+t-s)^{d_x} \leq v^{d_y-1}(1-s)^{d_x}$ since $v \leq 1-t$ and v^{d_y-1} is integrable by assumption. If $d_x < 0$ but at the same time $d_x + d_y > 0$, the integrand is bounded above at the point $t = s$ and $v^{d_x+d_y-1}$ is integrable. In these cases the integral exists for all $-\infty < s \leq t < 1$. If $d_x + d_y \leq 0$, in which case $d_x < 0$ by assumption, substitute (5.10) and (5.11) into (5.14) so as to write the integral as

$$\begin{aligned} d_y\int_{\max\{0,-t\}}^{1-t} & v^{d_y-1}(v+t-s)^{d_x}\mathrm{d}v \\ &= (t-s)^{d_x}\Big((1-t)^{d_y}Z_1^A(t,s) - 1_{\{t<0\}}(-t)^{d_y}Z_2^A(t,s)\Big) \end{aligned} \tag{5.15}$$

making the changes of variable $\tau = v/(1-t)$ in the first term and $\tau = v/(-t)$ in the second term. This shows that the integral exists for all $-\infty < s < t < 1$, but that there is a singularity at the point $s = t$. ∎

Since $Z_1^A(t,s)$ and $Z_2^A(t,s)$ converge to zero as $s \to t$ when $d_x < 0$, it is not correct to conclude that (5.15) diverges at $s = t$ in that case. It is simply that it becomes impossible to assign the integral a value there, although a solution does exist for all $s < t$. To interpret the singularity, it is helpful to note that after normalization and going to the limit, the gap $t-s$ relates to the terms of the series $\{a_{nmp}, p \le m\}$ in (5.5) as p declines. When $d_x + d_y > 0$ these terms diverge at the rate $n^{d_x+d_y}$ whereas with $d_x + d_y < 0$ they converge to zero at rate $n^{d_x+d_y}$. The practical difficulty is that when the series is convergent, the first few coordinates are not vanishing and so are ill-adapted to the normalization by $n^{d_x+d_y}$. As can be seen in the definition of q_{nm} in (5.4), the point $p = m$ is where the dates of the shock variables differ by only a single period.

This phenomenon links directly to the differing orders of magnitude of the mean and the mean deviation processes examined in Theorem **4.7**. The question also arises under the heading of autocorrelated shocks in Chapter 8, where it is treated more formally, but the simplest solution in the immediate run is to impose the restriction $s < t$ on the integral and by this means move the excluded terms (at most finite in number) from G_{1n} to G_{2n} where, having measure zero in the limit, their effect on the limiting expectation can be neglected.

Now consider the case of G_{3n}. Define, analogously to (5.10) and (5.11), the hypergeometric integrals

$$Z_1^E(s,t) = d_x \int_0^1 \tau^{d_x-1}\Big(1 - \frac{s-1}{s-t}\tau\Big)^{d_y} \mathrm{d}\tau$$

for $-\infty < s < 1$ and $-\infty < t < s$, and

$$Z_2^E(s,t) = d_x \int_0^1 \tau^{d_x-1}\Big(1 - \frac{s}{s-t}\tau\Big)^{d_y} \mathrm{d}\tau$$

for $-\infty < s < 0$ and $-\infty < t < s$, noting that these are well defined when $d_x > 0$. The following result closely parallels Lemma **5.2** except that d_x and d_y switch roles and $s \ge t$, reflecting the switch of roles of p and m.

5.3 Lemma If $d_x > 0$ and $d_x + d_y > 0$, $K(n)^{-1}e_{n[ns][nt]} = E(s,t) + o(1)$ as $n \to \infty$ for real-valued indices t, s with $-\infty < t \le s < 1$, where

$$\begin{aligned} E(s,t) = (1-t)^{d_y}\big((1-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\big) \\ - d_x \int_{\max\{0,-s\}}^{1-s} v^{d_x-1}(v+s-t)^{d_y}\mathrm{d}v. \end{aligned} \tag{5.16}$$

If $d_x > 0$ and $d_x + d_y \le 0$, the same conclusion holds for $-\infty < t < s < 1$.

Proof Define v by $j = [nv]$. Applying Theorem **5.1** to (5.8), similarly to (5.13),

$$\frac{1}{n^{d_y}L_y(n)} \sum_{l=[nv]+[ns]+1-[nt]}^{n-[nt]} c_l = (1-t)^{d_y} - (v+s-t)^{d_y} + o(1)$$

so that

$$E(s,t) = d_x \int_{\max\{0,-s\}}^{1-s} v^{d_x-1}\big((1-t)^{d_y} - (v+s-t)^{d_y}\big)\mathrm{d}v.$$

After rearrangement this formula matches (5.16). The integral exists with $d_x + d_y > 0$ by reasoning similar to **5.2**. If $d_x + d_y \leq 0$, so that necessarily $d_y < 0$, substitutions with respective changes of variable $\tau = v/(1-s)$ and $\tau = v/(-s)$ yield the form

$$d_x \int_{\max\{0,-s\}}^{1-s} v^{d_x - 1}(v + s - t)^{d_y} \mathrm{d}v$$
$$= (s-t)^{d_y}\big((1-s)^{d_x} Z_1^E(s,t) - 1_{\{s<0\}}(-s)^{d_x} Z_2^E(s,t)\big). \tag{5.17}$$

which exists except at the point $s = t$. ∎

5.3 Heuristic Representation

Let U and W denote Brownian motions on the real line segment $(-\infty, 1]$, setting $U(0) = W(0) = 0$ (which involves no loss of generality) and with variances $\mathrm{E}(U(1)^2) = \sigma_u^2$ and $\mathrm{E}(W(1)^2) = \sigma_w^2$ and covariance $\mathrm{E}(U(1)W(1)) = \sigma_{uw}$. The substitutions $\mathrm{d}U(s)$ for $u_{[ns]}/\sqrt{n}$ and $\mathrm{d}W(t)$ for $w_{[nt]}/\sqrt{n}$ can then be made to develop an asymptotic approximation to (5.3). Writing $\mathcal{F}_n(t) = \mathcal{F}_{[nt]} \in \boldsymbol{F}$, as defined following Assumption **4.1**, let $\{\mathcal{F}(t), t \in \mathbb{R}\}$ denote the filtration to which U and W are adapted, where $\mathcal{F}(t)$ is the limiting case of $\mathcal{F}_n(t)$ as $n \to \infty$.

The limit of the normalized random variable G_{1n} in (5.3) can be expressed heuristically in the form

$$\Xi_1 = \int_{-\infty}^{1} Q(t)\mathrm{d}W(t) \tag{5.18}$$

where with $A(s,t)$ defined in (5.12),

$$Q(t) = \int_{-\infty}^{t} A(t,s)\mathrm{d}U(s) \tag{5.19}$$

is a $\mathcal{F}(t)$-adapted stochastic process in continuous time. Note the remark on page 18 concerning the interpretation of infinite-lag processes, which also applies to (5.19) and (5.18). Notwithstanding the possible normalization issue at the point $s = t$ in (5.19) when $d_x + d_y < 0$, bear in mind that the contribution of this point is of measure zero under continuity of the Brownian motion. Since the variations of U are controlled its contribution to the distribution of $Q(t)$ is negligible.

For future reference, useful implications of formula (5.19) include the variance of $Q(t)$ for $-\infty < t \leq 1$,

$$\mathrm{E}(Q(t)^2) = \sigma_u^2 \int_{-\infty}^{t} A^2(t,s)\mathrm{d}s \tag{5.20}$$

and also the variance of an increment,

$$\mathrm{E}(Q(t+\delta) - Q(t))^2$$

$$= \sigma_u^2 \int_t^{t+\delta} A(t+\delta, s)^2 \mathrm{d}s + \sigma_u^2 \int_{-\infty}^{t} \big(A(t+\delta, s) - A(t, s)\big)^2 \mathrm{d}s. \tag{5.21}$$

If $d_y = 0$ and $c_l = 0$ for $l > 0$, (5.5) specializes to $a_{nmp} = a_{np}(m/n, 0)$ where a_{np} is defined in (2.25) and (5.4) becomes

$$q_{nm} = \frac{1}{n^{d_x+1/2} L_x(n)} \sum_{k=1}^{m} x_k \sim X_n(m/n) \tag{5.22}$$

from (2.27). In other words, $Q(t) = X(t)$ for $t \geq 0$ and 0 for $t < 0$ where X is fBM as defined by (2.1) with $d = d_x$. Equation (5.19) is then redundant and

$$\Xi_1 = \int_0^1 X(t) \mathrm{d}W(t) \tag{5.23}$$

which is the Itô integral with respect to W of a fBM integrand.

In the same way, the limit of G_{3n} can be expressed in the form

$$\Xi_3 = \int_{-\infty}^{1} H(s) \mathrm{d}U(s) \tag{5.24}$$

where the $\mathcal{F}(s)$-adapted integrand process is

$$H(s) = \int_{-\infty}^{s} E(s, t) \mathrm{d}W(t). \tag{5.25}$$

Observe in particular that when $d_y = 0$, $E(s, t) = 0$ for all s and t. The term Ξ_3 arises only in the case of a fractional integrator function. Otherwise, $\Xi = \Xi_1$ and this term has the form shown in (5.23).

Notice the important fact that in all cases, both Ξ_1 and (where it exists) Ξ_3 are stochastic integrals of $\mathcal{F}$-adapted integrand processes with respect to $\mathcal{F}$-adapted Brownian motions. These integrals are of the Itô type. Subject to sufficient regularity conditions on the integrands, essentially those of finite variances and almost sure continuity, plus the validity of mean-squared approximations by integrals with finite domain of integration, they may be analyzed in the conventional fashion, familiar from the unit root analysis.

5.4 Antipersistent Integrators

When $d_y < 0$, while the formula in (5.5) is valid the integral approximation in (5.12) fails and another way of constructing an asymptotic approximation must be found. In the case $d_x < 0$, the same is true of (5.8) and (5.16).

Focusing first on G_{1n}, an alternative form of expression (5.5) exists whenever it is possible to write $c_l = c_l^* - c_{l-1}^*$ for $l > 0$ for a suitably defined sequence $\{c_l^*\}$, with $c_0^* = c_0 = 1$. Then, if $c_l \sim d_y l^{d_y - 1} L_y(l)$ for $l > 0$ with $-\frac{1}{2} < d_y < 0$ it must be the case that $c_l^* \sim l^{d_y} L_y(l)$, as explained in the discussion leading to (1.19). With this substitution the double sum of (5.5) can be written as

$$\begin{aligned} a_{nmp} &= 1_{\{m>0\}} c_0^* \sum_{j=\max\{0,1-p\}}^{m-p} b_j + \sum_{l=\max\{1,1-m\}}^{n-1-m} (c_l^* - c_{l-1}^*) \sum_{j=\max\{0,1-p\}}^{l+m-p} b_j \\ &= c_{n-1-m}^* \sum_{j=\max\{0,1-p\}}^{n-1-p} b_j - \sum_{l=\max\{0,-m\}}^{n-2-m} c_l^* b_{l+1+m-p} \end{aligned} \tag{5.26}$$

where the second equality is obtained by cancelling the pairs of matching terms with opposite signs.

A simple application of the formula in (5.26), for comparison with (5.22), is the case $d_y = 0$. Specifically, if $c_l = 0$ for $l > 0$ then $c_l^* = 1$ for all $l > 0$. Formula (5.26) therefore yields $a_{nmp} = b_{\max\{0,1-p\}} + \cdots + b_{m-p}$ for $1 \leq m \leq n-1$ and 0 otherwise, which matches the corresponding case of (5.5).

The question of interest is whether there is a valid counterpart of Lemma **5.2**. While (5.26) is identical to (5.5), the same set of terms simply being ordered in a different way, the asymptotic approximation needs to be constructed differently. Therefore, although the notation a_{nmp} applies in each case, the expression $A(t,s)$ in (5.12) has to be replaced a different one, to be denoted $A^*(t,s)$. The same integral approximation arguments apply. Similarly to (5.10) and (5.11), define hypergeometric functions

$$Z_1^{A*}(t,s) = (1+d_y) \int_0^1 \tau^{d_y} \left(1 - \frac{t-1}{t-s}\tau\right)^{d_x - 1} \mathrm{d}\tau \tag{5.27}$$

for $-\infty < t < 1$ and $-\infty < s < t$, and

$$Z_2^{A*}(t,s) = (1+d_y) \int_0^1 \tau^{d_y} \left(1 - \frac{t}{t-s}\tau\right)^{d_x - 1} \mathrm{d}\tau. \tag{5.28}$$

for $-\infty < t < 0$ and $-\infty < s < t$. In these cases of $F(a,b;c;z)$, $b = 1 + d_y > 0$ and $c = 2 + d_y$ and the integrals are well defined as before.

5.4 Lemma If $-\frac{1}{2} < d_y < 0$, and $d_x + d_y > 0$, $K(n)^{-1} a_{n[nt][ns]} = A^*(t,s) + o(1)$ as $n \to \infty$ for real-valued indices t, s with $-\infty < s \leq t < 1$, where

$$\begin{aligned} A^*(t,s) &= (1-t)^{d_y} \left((1-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\right) \\ &\quad - d_x \int_{\max\{0,-t\}}^{1-t} v^{d_y} (v + t - s)^{d_x - 1} \mathrm{d}v. \end{aligned} \tag{5.29}$$

If $-\frac{1}{2} < d_y < 0$ and $d_x + d_y \leq 0$, the same conclusion holds for $-\infty < s < t < 1$.

Proof The first term of (5.29) is easily constructed from that of (5.26) since

$$\frac{c_{n-1-[nt]}^*}{n^{d_y} L_y(n)} = (1-t)^{d_y} + o(1).$$

The second term is found by the argument analogous to that of Lemma **5.2** since

$$\frac{c^*_{[nv]}}{n^{d_y}L_y(n)} = v^{d_y} + o(1).$$

Integrability holds in this case with $d_x + d_y > 0$ because with $d_x - 1 < 0$ the integrand is bounded above by $v^{d_y+d_x-1}$. If $d_x + d_y \leq 0$, substitute (5.27) with change of variable $\tau = v/(1-t)$ and (5.28) with change of variable $\tau = v/(-t)$ to obtain

$$\begin{aligned} &d_x \int_{\max\{0,-t\}}^{1-t} v^{d_y}(v+t-s)^{d_x-1}\mathrm{d}v \\ &= \frac{d_x}{1+d_y}(t-s)^{d_x-1}\Big((1-t)^{d_y+1}Z_1^{A*}(t,s) - 1_{\{t<0\}}(-t)^{d_y+1}Z_2^{A*}(t,s)\Big) \quad (5.30) \end{aligned}$$

which is well defined for $-\infty < s < t < 1$. ∎

In the case of G_{3n}, the modification corresponding to (5.26) can be performed on the formula in (5.8) for the case $d_x < 0$. Define the sequence b^*_j by $b^*_0 = 1$ and $b_j = b^*_j - b^*_{j-1}$ for $j > 0$, with $b^*_j \sim j^{d_x}L_x(j)$. Noting $p > m$ in this case, the double sum in (5.8) becomes

$$\begin{aligned} e_{nmp} &= 1_{\{p>0\}}b^*_0 \sum_{l=p+1-m}^{n-m} c_l + \sum_{j=\max\{1,1-p\}}^{n-1-p} (b^*_j - b^*_{j-1}) \sum_{l=j+p+1-m}^{n-m} c_l \\ &= \sum_{j=\max\{0,1-p\}}^{n-1-p} b^*_j c_{j+p+1-m} - 1_{\{p\leq 0\}}b^*_{-p} \sum_{l=2-m}^{n-m} c_l \quad (5.31) \end{aligned}$$

where the second equality is obtained by cancelling equal and oppositely signed terms. Arguments closely paralleling **5.4** now give the following.

5.5 Lemma If $-\frac{1}{2} < d_x < 0$ and $d_x + d_y > 0$ then $e_{n[nt][ns]} = E^*(s,t) + o(1)$ as $n \to \infty$ for real-valued indices t, s with $-\infty < t \leq s < 1$, where

$$\begin{aligned} E^*(s,t) = d_y &\int_{\max\{0,-s\}}^{1-s} u^{d_x}(u+s-t)^{d_y-1}\mathrm{d}u \\ &- 1_{\{s\leq 0\}}(-s)^{d_x}\big((1-t)^{d_y} - (-t)^{d_y}\big). \quad (5.32) \end{aligned}$$

If $-\frac{1}{2} < d_x < 0$ and $d_x + d_y \leq 0$, the same conclusion holds for $-\infty < t < s < 1$. □

This being the fourth instance of this type of result with essentially similar characteristics to the preceding three, the details don't really require further repetition. These are left for the reader to fill in as may be appropriate.

In the approximation arguments to appear in Chapter 6, the expressions A^* and E^* are used in place of A and E to deal with antipersistent integrators in Ξ_1 and Ξ_3 respectively. With this exception, the arguments are very similar in the two cases.

5.5 Integration by Parts

At this point, recall from §4.2 the suggested notation $G_n^{xy} = G_{1n}^{xy} + G_{2n}^{xy} + G_{3n}^{xy}$ in place of G_n, when the distinction needs to be made. In this framework, it is shown by Theorem **4.8** that $G_{2n}^{xy} \to_{L_2} \sigma_{uw}\lambda_{xy}$ and it is thus far conjectured (and to be proved formally in Chapter 6) that G_{1n}^{xy} and G_{3n}^{xy} converge in distribution to limits Ξ_{1xy} and Ξ_{3xy} with sum Ξ_{xy} where these symbols now stand in for the Ξ_1, Ξ_3 and Ξ used previously. The complementary cases are defined as Ξ_{1yx}, Ξ_{3yx} and Ξ_{yx} in which the roles of the variables x and y are interchanged, with λ_{yx} already defined on page 64. In a similar fashion, notations A_{xy}, A_{yx} and A^*_{xy}, A^*_{yx}, and also E_{xy}, E_{yx} and E^*_{xy}, E^*_{yx} will be used to distinguish the integrals in (5.12), (5.29), (5.16), and (5.32) from their complementary cases.

With these changes in place, let X and Y denote the fBMs having the form of (2.1) with respective parameters d_x and d_y and respective driving processes U and W, as defined in §5.3. In the case $d_x + d_y > 0$, it appears natural to equate the random variable $\Xi_{xy} + \sigma_{uw}\lambda_{xy}$ with the stochastic integral $\int_0^1 X\mathrm{d}Y$ and likewise $\Xi_{yx} + \sigma_{uw}\lambda_{yx}$ with $\int_0^1 Y\mathrm{d}X$. However, for the designation 'integral' here to be appropriate it should be the case that the integration by parts rule

$$\int_0^1 X\mathrm{d}Y + \int_0^1 Y\mathrm{d}X = X(1)Y(1) \tag{5.33}$$

is obeyed. Matters are less clear cut when $d_x + d_y \le 0$ since then λ_{xy} and λ_{yx} don't exist but an unambiguous requirement for a valid formulation, in every case with $|d_x| < \frac{1}{2}$ and $|d_y| < \frac{1}{2}$, is

$$\Xi_{xy} + \Xi_{yx} + \sigma_{uw}\Upsilon_{xy} = X(1)Y(1). \tag{5.34}$$

To explore these questions, some interesting facts about the complementary expressions can be brought to light. First, from (5.18) and (5.24) it is evident that

$$\Xi_{xy} = \int_{-\infty}^1 \int_{-\infty}^t A_{xy}(t,s)\mathrm{d}U(s)\mathrm{d}W(t) + \int_{-\infty}^1 \int_{-\infty}^s E_{xy}(s,t)\mathrm{d}W(t)\mathrm{d}U(s).$$

Swapping the variables, the arguments are all completely symmetric so it is also immediate that

$$\Xi_{yx} = \int_{-\infty}^1 \int_{-\infty}^s A_{yx}(s,t)\mathrm{d}W(t)\mathrm{d}U(s) + \int_{-\infty}^1 \int_{-\infty}^t E_{yx}(t,s)\mathrm{d}U(s)\mathrm{d}W(t).$$

Adding these terms together gives

$$\begin{aligned}\Xi_{xy} + \Xi_{yx} = &\int_{-\infty}^1 \int_{-\infty}^t (A_{xy}(t,s) + E_{yx}(t,s))\mathrm{d}U(s)\mathrm{d}W(t) \\ &+ \int_{-\infty}^1 \int_{-\infty}^s (A_{yx}(s,t) + E_{xy}(s,t))\mathrm{d}W(t)\mathrm{d}U(s).\end{aligned} \tag{5.35}$$

Examine (5.12) and also the version of (5.16) that is obtained by swapping x and y and also t and s. The integral terms are equal and oppositely signed and so cancel in the sum, and after some rearrangement it is found that

$$\begin{aligned} &A_{xy}(t,s) + E_{yx}(t,s) \\ &\quad = \big((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\big)\big((1-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\big). \end{aligned} \tag{5.36}$$

Given the symmetry of the two cases, it is just as easily seen that

$$\begin{aligned} &A_{yx}(s,t){+}E_{xy}(s,t) \\ &\quad = \big((1-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\big)\big((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\big) \end{aligned} \tag{5.37}$$

so that expressions (5.36) and (5.37) are actually identical.

These identities provide a key step in the proof of the next theorem confirming the equality in (5.34). This is not an asymptotic result, U and W being assumed to be regular Brownian motions and the integrals otherwise depending solely on the d_x and d_y parameters and σ_{uw}. Notice that σ_{uw} is the symbol appearing in the limit in Theorem **4.4**, because there the limit is derived under Assumption **4.1** in which the finite-sample shocks are serially independent. If these are in fact autocorrelated, which is the case to be developed in Chapter 8 under Assumption **8.1**, then σ_{uw} should be replaced by ω_{uw} standing for the long-run covariance. In the present context, nothing depends on these assumptions since the constant is merely an attribute of the joint distribution of (U, W) and the switch is a formality.

5.6 Theorem If $|d_x| < \frac{1}{2}$ and $|d_y| < \frac{1}{2}$ then (5.34) holds.

Proof Assume initially that $d_y \geq 0$ and $d_x \geq 0$, these being the cases to which formulae (5.36), (5.37) and (5.35) are relevant.

It will be helpful to introduce a compact notation. For a function $F_Y : (-\infty, 1] \mapsto \mathbb{R}$ and fractional process X (as defined in (2.1), with parameter d_x) define $\int F_Y \delta X$ by the formula

$$\int F_Y \delta X = \int_{-\infty}^{1} F_Y(\tau)\big((1-\tau)^{d_x} - 1_{\{\tau<0\}}(-\tau)^{d_x}\big)\mathrm{d}U(\tau). \tag{5.38}$$

In particular, note on putting $F_Y = 1$ that $\int \delta X = X(1)$ from (2.1). Similarly, for a function F_X and fractional process Y with parameter d_y define

$$\int F_X \delta Y = \int_{-\infty}^{1} F_X(t)\big((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\big)\mathrm{d}W(t). \tag{5.39}$$

such that $\int \delta Y = Y(1)$. Let the roles of F_Y and F_X be played initially by the respective processes $\widetilde{Y} : (-\infty, 1] \mapsto \mathbb{R}$ and $\widetilde{X} : (-\infty, 1] \mapsto \mathbb{R}$, where

$$\widetilde{Y}(\tau) = \int_{-\infty}^{\tau} \big((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\big)\mathrm{d}W(t)$$

and

$$\widetilde{X}(t) = \int_{-\infty}^{t} \left((1-\tau)^{d_x} - 1_{\{\tau<0\}}(-\tau)^{dx}\right)\mathrm{d}U(\tau).$$

Thus, note that $\widetilde{Y}(\tau) = Y(\tau)$ for $\tau \geq 0$ and $\widetilde{X}(t) = X(t)$ for $t \geq 0$, the difference being that these processes are defined over the entire domain indicated. It can be verified that $\int \widetilde{Y}\delta X$ and $\int \widetilde{X}\delta Y$ are Itô integrals with respective integrator processes U and W, with the important implication that they have zero means. Further, substituting the matching formulae from (5.36) and (5.37) into (5.35) yields the representation

$$\Xi_{xy} + \Xi_{yx} = \textstyle\int \widetilde{X}\delta Y + \int \widetilde{Y}\delta X. \tag{5.40}$$

Next, define functions

$$\widehat{X}(t) = X(1) - \widetilde{X}(t) = \int_{t}^{1} \left((1-\tau)^{d_x} - 1_{\{\tau<0\}}(-\tau)^{d_x}\right)\mathrm{d}U(\tau) \tag{5.41}$$

and

$$\widehat{Y}(\tau) = Y(1) - \widetilde{Y}(\tau) = \int_{\tau}^{1} \left((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right)\mathrm{d}W(t). \tag{5.42}$$

Setting $F_X = \widehat{X}$, for example, consider the following alternative representations of the iterated integral:

$$\begin{aligned}\textstyle\int \widehat{X}\delta Y &= \int_{-\infty}^{1}\int_{-\infty}^{1} 1_{\{\tau\geq t\}}\left((1-\tau)^{d_x} - 1_{\{\tau<0\}}(-\tau)^{d_x}\right) \\ &\qquad \times \left((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right)\mathrm{d}U(\tau)\mathrm{d}W(t) \\ &= \textstyle\int \widetilde{Y}\delta X + \mathrm{E}\left(\int \widehat{X}\delta Y\right).\end{aligned} \tag{5.43}$$

The order of iteration is being swapped here, the indicator function in the second member setting the domain of integration to be the upper triangle of the plane $(-\infty, 1] \times (-\infty, 1]$. Since $\mathrm{E}\left(\int \widetilde{X}\delta Y\right) = 0$, the additional term in the third member of (5.43) equates the means of each side and can be viewed as representing the diagonal contribution to the double integral. Direct calculation using

$$\mathrm{E}\left(\mathrm{d}U(\tau)\mathrm{d}W(t)\right) = \begin{cases} \sigma_{uw}\mathrm{d}t, & t = \tau \\ 0, & \text{otherwise} \end{cases}$$

gives

$$\begin{aligned}\mathrm{E}\left(\textstyle\int \widehat{X}\delta Y\right) &= \sigma_{uw}\int_{-\infty}^{1}\left((1-t)^{d_x} - 1_{\{t<0\}}(-t)^{d_x}\right)\left((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right)\mathrm{d}t \\ &= \sigma_{uw}\left(\int_0^1 (1-t)^{d_x+d_y}\,\mathrm{d}t + \int_0^\infty \left((1+t)^{d_x} - t^{d_x}\right)\left((1+t)^{d_y} - t^{d_y}\right)\mathrm{d}t\right) \\ &= \sigma_{uw}\Upsilon_{xy}\end{aligned} \tag{5.44}$$

where the second equality is by (4.19). Given the first equality in (5.41) and the fact that $\mathrm{E}\big(\int \widetilde{X}\delta Y\big) = 0$, it can also be verified that $\mathrm{E}\big(\int \widehat{X}\delta Y\big) = \mathrm{E}(X(1)Y(1))$ so that (5.44) can be paired with the result already given in (4.20).

The final step in the argument is to add $\int \widetilde{X}\delta Y$ to each side of (5.43). Applying (5.40), (5.44), and the fact that $\int \widehat{X}\delta Y + \int \widetilde{X}\delta Y = X(1)Y(1)$ yields (5.34). By symmetry, to match (5.43) it is equally the case that

$$\int \widehat{Y}\delta X = \int \widetilde{X}\delta Y + \sigma_{uw}\Upsilon_{xy} \tag{5.45}$$

which gives the same result by addition of $\int \widetilde{Y}\delta X$ to both sides.

This completes the proof for the case $d_x \geq 0$ and $d_y \geq 0$. The argument is extended to allow antipersistence in either or both of the variables by using the formulae developed in §5.4. Similarly to what was done previously let $A^*_{yx}(s,t)$ and $E^*_{yx}(t,s)$ be the cases complementary to $A^*_{xy}(t,s)$ and $E^*_{xy}(s,t)$ from (5.29) and (5.32), with d_y and d_x and also s and t interchanged. It can then be verified by direct inspection of (5.36) that

$$A^*_{xy}(t,s) + E^*_{yx}(t,s) = A_{xy}(t,s) + E_{yx}(t,s) \tag{5.46}$$

and also

$$A^*_{yx}(s,t) + E^*_{xy}(s,t) = A_{yx}(s,t) + E_{xy}(s,t). \tag{5.47}$$

The integral terms cancel as before and all four of these sums match the identity in (5.36). If $d_y < 0$ then let $A^*_{xy}(t,s) + E^*_{yx}(t,s)$ replace $A_{xy}(t,s) + E_{yx}(t,s)$ in (5.35). If $d_x < 0$, let $A^*_{yx}(s,t) + E^*_{xy}(s,t)$ replace $A_{yx}(s,t) + E_{xy}(s,t)$. It is only the sum of these terms that appear in the preceding argument and it follows from (5.46) and (5.47) that the proof does not depend in the signs of d_x and d_y. ∎

Chapter 6

Weak Convergence of Integrals

The objective of this chapter is to explore the conditions under which

$$(X_n, Y_n, G_n - \mathrm{E}(G_n)) \xrightarrow{\mathrm{d}} \big(X, Y, \Xi\big) \tag{6.1}$$

denoting joint weak convergence in $D^2_{[0,1]} \times \mathbb{R}$ where $D^2_{[0,1]}$ denotes the space of càdlàg pairs on the unit interval equipped with the Skorokhod topology, as explained on page 53. In (6.1), X_n and Y_n are the normalized partial sums of fractional processes $\{x_i\}$ and $\{y_i\}$ defined for the exemplar case by (2.27) with respective shock variances σ_u^2 and σ_w^2 and covariance σ_{uw}, while X and Y are fractional Brownian motions as in (2.1). It is understood that G_n and Ξ are the objects that were written in §5.5 as G_n^{xy} and Ξ_{xy}, with the additional decorations now being omitted to minimize clutter. As discussed in §5.5, there is also the understanding here that $\Xi = \int_0^1 X \mathrm{d}Y - \sigma_{uw}\lambda_{xy}$.

The joint convergence of the triple specified in (6.1) is a requisite for arguments depending on application of the continuous mapping theorem to functionals of the components, such those to be explored in Chapter 7. It follows from **3.19** that arbitrary linear combinations of $(X_n, Y_n,\ G_n - \mathrm{E}(G_n))$ must converge to the corresponding combinations of the limit processes. The marginal weak limits for the first two members of (6.1) follow in effect from Theorem **3.20** under the conditions specified. The scalar $G_n - \mathrm{E}(G_n) = G_{1n} + G_{3n}$ can be embedded in $D_{[0,1]}$ by the simple expedient of defining a process on $[0, 1]$ to equal it at every point. If it can be shown that $G_{1n} \to_{\mathrm{d}} \Xi_1$ and $G_{3n} \to_{\mathrm{d}} \Xi_3$ where the limit random variables Ξ_1 and Ξ_3 can be identified with the stochastic integrals in (5.18) and (5.24), and $\Xi = \Xi_1 + \Xi_3$, the continuous mapping theorem would then yield the weak limit for the third element of (6.1).

6.1 More Fractional Asymptotics

In Chapters 4 and 5, formulae have been derived that intuitively correspond with the limiting form of the fractional covariance, respectively the mean process and the mean deviations, as the sample size n increases. It now has to be shown that these components do converge weakly to the identified limit distributions. To address the fact that the partial sums in question have an infinite number of terms, the approach as in Chapter 3 is to split each sum into a block of order $n(N+1)$ and a remainder, where the remainder is of small order in L_2-norm as $N \to \infty$; not as $n \to \infty$, note, because N in fact characterizes the limit distribution.

Because of the form of the sums, there are in practice two remainder terms to consider. For the chosen $N \in \mathbb{N}$, first define

$$G_{1n}^N = \frac{1}{\sqrt{n}} \sum_{m=-nN}^{n-1} q_{nm}^N w_{m+1} \tag{6.2}$$

where

$$q_{nm}^N = \frac{1}{\sqrt{n}K(n)} \sum_{p=-nN}^{m} a_{nmp} u_p. \tag{6.3}$$

Then (5.3) may be rearranged as

$$G_{1n} = G_{1n}^N + \frac{1}{\sqrt{n}} \sum_{m=-nN}^{n-1} (q_{nm} - q_{nm}^N) w_{m+1} + \frac{1}{\sqrt{n}} \sum_{m=-\infty}^{-nN-1} q_{nm} w_{m+1}. \tag{6.4}$$

Similarly, define

$$G_{3n}^N = \frac{1}{\sqrt{n}} \sum_{p=-nN}^{n-1} h_{np}^N u_p \tag{6.5}$$

and

$$h_{np}^N = \frac{1}{\sqrt{n}K(n)} \sum_{m=-nN}^{p-1} e_{npm} w_m \tag{6.6}$$

and so for G_{3n} in (5.6) write

$$G_{3n} = G_{3n}^N + \frac{1}{\sqrt{n}} \sum_{p=-nN}^{n-1} (h_{np} - h_{np}^N) u_p + \frac{1}{\sqrt{n}} \sum_{p=-\infty}^{-nN-1} h_{np} u_p. \tag{6.7}$$

It has to be shown that the limits of the last two terms on the right-hand sides of (6.4) and (6.7) as $n \to \infty$ can each be made as small as desired in L_2-norm by taking N large enough. The following pair of lemmas provide bounds for the limit formulae.

6.1 Lemma $K(n)^{-1}\left|a_{n[nt][ns]}\right| \leq \bar{A}(t,s) + o(1)$ as $n \to \infty$, where

$$\bar{A}(t,s) = \left|(1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right|\left|(g-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\right| \tag{6.8}$$

with $g = t$ if $d_x < 0$ and $s \geq 0$, and $g = 1$ otherwise.

Proof Putting $[nt]$ for m and $[ns]$ in place of p for $-\infty < s \le t \le 1$ in formula (5.5),

$$\begin{aligned}
\frac{\left|a_{n[nt][ns]}\right|}{K(n)} &= \frac{1}{K(n)}\left|\sum_{l=\max\{0,1-[nt]\}}^{n-1-[nt]} c_l\Bigg(\sum_{j=\max\{0,1-[ns]\}}^{l+[nt]-[ns]} b_j\Bigg)\right| \\
&\le \frac{1}{K(n)}\left|\sum_{l=\max\{0,1-[nt]\}}^{n-1-[nt]} c_l\right| \max_{\max\{0,1-[nt]\}\le l\le n-1-[nt]}\left|\sum_{j=\max\{0,1-[ns]\}}^{l+[nt]-[ns]} b_j\right| \\
&= \left|(1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right| \\
&\qquad\times \max_{\max\{0,t\}\le g\le 1}\left|(g-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\right| + o(1)
\end{aligned} \tag{6.9}$$

where g is defined by $[ng] = l + [nt]$. When $d_x \ge 0$, $(g-s)^{d_x}$ is monotone nondecreasing in g, and is maximized over $[0,1]$ at $g = 1$. When $d_x < 0$, $(g-s)^{d_x}$ is monotone decreasing in g, and if $s \ge 0$ so that $t \ge 0$, the maximum in (6.9) is achieved at $g = t$. On the other hand, if $d_x < 0$ and $s < 0$ then $\left|(g-s)^{d_x} - (-s)^{d_x}\right|$ is maximized at $g = 1$. ∎

Since $\bar{A}(t,s)$ bounds both $A(t,s)$ in (5.12) and $A^*(t,s)$ in (5.29), results that depend on bounding $\bar{A}(t,s)$ hold both for $|d_x| < \frac{1}{2}$ and for $|d_y| < \frac{1}{2}$.

6.2 Lemma $K(n)^{-1}|e_{n[ns][nt]}| \le \bar{E}(s,t) + o(1)$ where

$$\bar{E}(s,t) = \left|(1-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\right|\left|(1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right|. \tag{6.10}$$

Proof From (5.8), for $s \ge t$,

$$\begin{aligned}
\frac{\left|e_{n[ns][nt]}\right|}{K(n)} &= \frac{1}{K(n)}\left|\sum_{j=\max\{0,1-[ns]\}}^{n-1-[ns]} b_j\Bigg(\sum_{l=j+[ns]+1-[nt]}^{n-[nt]} c_l\Bigg)\right| \\
&\le \frac{1}{K(n)}\left|\sum_{j=\max\{0,1-[ns]\}}^{n-1-[ns]} b_j\right| \max_{\max\{0,1-[ns]\}\le j\le n-1-[ns]}\left|\sum_{l=j+[ns]+1-[nt]}^{n-[nt]} c_l\right| \\
&= \left|(1-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\right| \\
&\qquad\times \max_{\max\{0,s\}\le g\le 1}\left|(1-t)^{d_y} - (g-t)^{d_y}\right| + o(1)
\end{aligned}$$

where $[ng] = j + [ns]$. First, suppose $s \ge 0$. Whether $d_y \ge 0$ or $d_y < 0$, $|(1-t)^{d_y} - (g-t)^{d_y}|$ is maximized over $[s,1]$ at $g = s$. In the case $s < 0$ the same considerations apply, but the extremum over $[0,1]$ is at $g = 0$ in each case. The proof is completed by noting that for any $s \in [t,1]$ and d_y of either sign,

$$\left|(1-t)^{d_y} - (\max\{0,s\} - t)^{d_y}\right| \le \left|(1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right|. \quad ∎$$

In the same way as for $\bar{A}(t,s)$, results that work by bounding $\bar{E}(s,t)$ hold for $|d_x| < \frac{1}{2}$ and $|d_y| < \frac{1}{2}$.

The next pair of theorems bound the remainders $G_{1n} - G_{1n}^N$ and $G_{3n} - G_{3n}^N$, defined respectively in (6.4) and (6.7).

6.3 Theorem Under Assumption **4.1**,

(i) $\lim_{n\to\infty} n^{-1}\mathrm{E}\big(\sum_{m=-nN}^{n-1}(q_{nm}-q_{nm}^N)w_{m+1}\big)^2 = O(N^{2d_x-1})$

(ii) $\lim_{n\to\infty} n^{-1}\mathrm{E}\big(\sum_{m=-\infty}^{-nN-1} q_{nm}w_{m+1}\big)^2 = O(N^{2(d_x+d_y-1)})$.

Proof For part (i),

$$\begin{aligned}\frac{1}{n}\mathrm{E}\bigg(\sum_{m=-nN}^{n-1}(q_{nm}-q_{nm}^N)w_{m+1}\bigg)^2 &= \frac{1}{n^2K(n)^2}\mathrm{E}\bigg(\sum_{m=-nN}^{n-1}\sum_{p=-\infty}^{-nN-1}a_{nmp}u_pw_{m+1}\bigg)^2 \\ &= \sigma_u^2\sigma_w^2\frac{1}{n^2K(n)^2}\sum_{m=-nN}^{n-1}\sum_{p=-\infty}^{-nN-1}a_{nmp}^2 \\ &\le \sigma_u^2\sigma_w^2\int_{-N}^{1}\int_{-\infty}^{-N}\bar{A}^2(t,s)\mathrm{d}s\mathrm{d}t + o(1) \qquad (6.11)\end{aligned}$$

as $n\to\infty$. Applying Lemma **6.1**, with $g=1$ since $s<0$,

$$\begin{aligned}\int_{-N}^{1}\int_{-\infty}^{-N}\bar{A}^2(t,s)\mathrm{d}s\mathrm{d}t &= \int_0^{N+1}\big((t+1)^{d_y}-t^{d_y}\big)^2\mathrm{d}t\int_N^\infty\big((1+s)^{d_x}-s^{d_x}\big)^2\mathrm{d}s \\ &= O(N^{2d_x-1}). \qquad (6.12)\end{aligned}$$

For part (ii), similarly,

$$\begin{aligned}\frac{1}{n}\mathrm{E}\bigg(\sum_{m=-\infty}^{-nN-1}q_{nm}w_{m+1}\bigg)^2 &= \sigma_u^2\sigma_w^2\frac{1}{n^2K(n)^2}\sum_{m=-\infty}^{-nN-1}\sum_{p=-\infty}^{m}a_{nmp}^2 \\ &\le \sigma_u^2\sigma_w^2\int_{-\infty}^{-N}\int_{-\infty}^{t}\bar{A}^2(t,s)\mathrm{d}s\mathrm{d}t + o(1)\end{aligned}$$

where

$$\begin{aligned}\int_{-\infty}^{-N}\int_{-\infty}^{t}\bar{A}^2(t,s)\mathrm{d}s\mathrm{d}t &= \int_N^\infty\big((1+t)^{d_y}-t^{d_y}\big)^2\int_t^\infty\big((1+s)^{d_x}-s^{d_x}\big)^2\mathrm{d}s\mathrm{d}t \\ &\ll \int_N^\infty\big((1+t)^{d_y}-t^{d_y}\big)^2t^{2d_x-1}\mathrm{d}t \\ &= O(N^{2(d_y+d_x-1)}). \quad \blacksquare \qquad (6.13)\end{aligned}$$

6.4 Theorem Under Assumption **4.1**,

(i) $\lim_{n\to\infty} n^{-1}\mathrm{E}\big(\sum_{p=-nN}^{n-1}(h_{np}-h_{np}^N)u_p\big)^2 = O(N^{2d_y-1})$

(ii) $\lim_{n\to\infty} n^{-1}\mathrm{E}\big(\sum_{p=-\infty}^{-nN-1} h_{np}u_p\big)^2 = O(N^{2(d_x+d_y-1)})$.

Proof For part (i), following the pattern of (6.11) and applying Lemma **6.2**,

$$\frac{1}{n}\mathrm{E}\biggl(\sum_{p=-nN}^{n-1}(h_{np}-h_{np}^N)u_p\biggr)^2 \le \sigma_w^2\sigma_u^2\int_{-N}^{1}\int_{-\infty}^{-N}\bar{E}^2(s,t)\mathrm{d}t\mathrm{d}s + o(1)$$
$$= O(N^{2d_y-1}). \tag{6.14}$$

For part (ii), as in (6.13),

$$\frac{1}{n}\mathrm{E}\biggl(\sum_{p=-\infty}^{-nN-1}h_{np}u_p\biggr)^2 \le \sigma_u^2\sigma_w^2\int_{-\infty}^{-N}\int_{-\infty}^{s}\bar{E}^2(s,t)\mathrm{d}t\mathrm{d}s + o(1)$$
$$= O(N^{2(d_x+d_y-1)}). \quad \blacksquare \tag{6.15}$$

For the subsequent analysis the normalized sums $a_{nmp}/K(n)$ and $e_{npm}/K(n)$ must be shown to satisfy a collection of asymptotic boundedness and continuity conditions. The following two theorems follow essentially the same lines so that the proofs can be read in conjunction, but there are some minor differences between the formulae. Since (6.10) is not dependent on a parameter like (6.8), some complications are avoided in that case.

6.5 Theorem Defining a_{nmp} as in (5.5),

(i) $\limsup_n n^{-1}K(n)^{-2}\sum_{p=-\infty}^{[nt]} a_{n[nt]p}^2 < \infty$ for each fixed $t \in (-\infty, 1)$.

(ii) $\limsup_n n^{-1}K(n)^{-2}\sum_{p=-\infty}^{[nt]} a_{n[nt]p}^2 = O((-t)^{2d_y+2d_x-3})$ as $t \to -\infty$.

(iii) $\sup_{t\in(-\infty,1-\delta]}\limsup_n n^{-1}K(n)^{-2}\sum_{p=[nt]+1}^{[n(t+\delta)]} a_{n[n(t+\delta)]p}^2 = O(\delta^{\min\{1,2d_x+1\}})$.

(iv) $\sup_{t\in(-\infty,1-\delta]}\limsup_n n^{-1}K(n)^{-2}\sum_{p=-\infty}^{[nt]} (a_{n[n(t+\delta)]p} - a_{n[nt]p})^2 = O(\delta^{2d_x+1})$.

Proof To prove (i), break the indicated sum into components $\sum_{p=0}^{[nt]} a_{n[nt]p}^2$ and $\sum_{p=-\infty}^{\min\{-1,[nt]\}} a_{n[nt]p}^2$ where the first sum is 0 in the case $t < 0$. With $p = [ns]$, applying Lemma **6.1** to the first component noting the definition of g in (6.8) gives

$$\limsup_n \frac{1}{nK(n)^2}\sum_{p=0}^{[nt]} a_{n[nt]p}^2 \le \int_0^t \bar{A}^2(t,s)\mathrm{d}s$$
$$= (1-t)^{2d_y}\int_0^t (g-s)^{2d_x}\mathrm{d}s$$
$$= \frac{(1-t)^{2d_y}}{2d_x+1}\left\{\begin{array}{cc} 1-(1-t)^{2d_x+1}, & d_x \ge 0 \\ t^{2d_x+1}, & d_x < 0 \end{array}\right\}$$

$$< \infty. \tag{6.16}$$

For the second component,

$$\begin{aligned}
\limsup_n \frac{1}{nK(n)^2} \sum_{p=-\infty}^{\min\{-1,[nt]\}} a^2_{n[nt]p} &\leq \int_{-\infty}^{\min\{0,t\}} \bar{A}^2(t,s)\mathrm{d}s \\
&\leq \left((1-t)^{d_y} - 1_{\{t<0\}}(-t)^{d_y}\right)^2 \int_{\max\{0,-t\}}^{\infty} \left((1+s)^{d_x} - s^{d_x}\right)^2 \mathrm{d}s \\
&< \infty. \qquad (6.17)
\end{aligned}$$

To prove (ii) it is sufficient to note that with $t < 0$ the majorant of (6.17) is of $O((-t)^{2d_y+2d_x-3})$.

To show (iii), write

$$\begin{aligned}
\limsup_n \frac{1}{nK(n)^2} \sum_{p=[nt]+1}^{[n(t+\delta)]} a^2_{n[n(t+\delta)]p} &\leq \int_t^{t+\delta} \bar{A}^2(t+\delta,s)\mathrm{d}s \\
&= \left((1-t-\delta)^{d_y} - 1_{\{t+\delta<0\}}(-t-\delta)^{d_y}\right)^2 \\
&\qquad \times \int_t^{t+\delta} \left((g-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\right)^2 \mathrm{d}s.
\end{aligned}$$

If $t \geq 0$ and $d_x < 0$, Lemma **6.1** sets $g = t + \delta$ and

$$\int_t^{t+\delta} (t+\delta-s)^{2d_x}\mathrm{d}s = \frac{\delta^{2d_x+1}}{2d_x+1}$$

which vanishes with δ since $d_x > -\frac{1}{2}$. Otherwise, $g = 1$. If $t \geq 0$ then according to the mean value theorem,

$$\begin{aligned}
\int_t^{t+\delta} (1-s)^{2d_x}\mathrm{d}s &= \frac{(1-t)^{2d_x+1} - (1-t-\delta)^{2d_x+1}}{2d_x+1} \\
&= O(\delta). \qquad (6.18)
\end{aligned}$$

If $t < 0$ then with δ sufficiently small, $t + \delta \leq 0$. Similarly to (6.18),

$$\begin{aligned}
\int_t^{t+\delta} \left((1-s)^{d_x} - (-s)^{d_x}\right)^2 \mathrm{d}s &\leq 2\left(\int_t^{t+\delta} (1-s)^{2d_x}\mathrm{d}s + \int_t^{t+\delta} (-s)^{2d_x}\mathrm{d}s\right) \\
&= O(\delta).
\end{aligned}$$

These bounds are independent of t and hold uniformly with respect to $t \in (-\infty, 1)$.

Finally, to prove (iv), consider the formula in (5.5) for the case where m is replaced by $m+q$ where $0 < q < n-m$. The formula can be decomposed as

$$a_{n,m+q,p} = \left(\sum_{l=\max\{1-m-q,0\}}^{\max\{1-m,0\}-1} + \sum_{l=\max\{1-m,0\}}^{n-m-q-1}\right) c_l \left(\sum_{j=\max\{1-p,0\}}^{l+m-p} + \sum_{j=l+m-p+1}^{l+m+q-p}\right) b_j$$

whereas

$$a_{nmp} = \Bigg(\sum_{l=\max\{1-m,0\}}^{n-m-q-1} + \sum_{l=n-m-q}^{n-m-1} \Bigg) c_l \sum_{j=\max\{1-p,0\}}^{l+m-p} b_j.$$

After cancelling equal and opposite-signed terms,

$$a_{n,m+q,p} - a_{nmp} = D_{1n}(q,m,p) + D_{2n}(q,m,p) - D_{3n}(q,m,p) \tag{6.19}$$

where

$$D_{1n}(q,m,p) = \sum_{l=\max\{1-m-q,0\}}^{\max\{1-m,0\}-1} c_l \sum_{j=\max\{1-p,0\}}^{l+m-p} b_j$$

$$D_{2n}(q,m,p) = \sum_{l=\max\{1-m-q,0\}}^{n-m-q-1} c_l \sum_{j=l+m-p+1}^{l+m+q-p} b_j$$

$$D_{3n}(q,m,p) = \sum_{l=n-m-q}^{n-m-1} c_l \sum_{j=\max\{1-p,0\}}^{l+m-p} b_j.$$

D_{1n} vanishes unless $m \le 0$. Using arguments analogous to (6.9), writing $q = [n\delta]$ for $0 < \delta < 1-t$ the inequalities

$$\frac{|D_{jn}([n\delta],[nt],[ns])|}{K(n)} \le \bar{D}_j(\delta,t,s) + o(1)$$

hold for $j = 1, 2$ and 3, where

$$\bar{D}_1(\delta,s,t) = 1_{\{t<0\}} |(-t)^{d_y} - 1_{\{\delta+t<0\}}(-\delta-t)^{d_y}| \times \max_{\max\{-t-\delta,0\} \le g_1 \le \max\{-t,0\}} |(g_1+t-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}| \tag{6.20}$$

$$\bar{D}_2(\delta,t,s) = |(1-t-\delta)^{d_y} - 1_{\{t+\delta<0\}}(-t-\delta)^{d_y}| \times \max_{\max\{t,-\delta\} \le g_2 \le 1-t-\delta} |(g_2+\delta-s)^{d_x} - (g_2-s)^{d_x}| \tag{6.21}$$

$$\bar{D}_3(\delta,t,s) = |(1-t)^{d_y} - (1-t-\delta)^{d_y}| \times \max_{1-t-\delta \le g_3 \le 1-t} |(g_3+t-s)^{d_x} - (-s)^{d_x}|. \tag{6.22}$$

In view of (6.19),

$$\limsup_{n\to\infty} \frac{1}{nK(n)^2} \sum_{p=-\infty}^{[nt]} (a_{n[n(t+\delta)]p} - a_{n[nt]p})^2 \le 3 \sum_{j=1}^{3} \int_{-\infty}^{t} \bar{D}_j^2(\delta,t,s)\mathrm{d}s.$$

The squares of the second factors of $\bar{D}_1$ in (6.20) and $\bar{D}_3$ in (6.22), are integrable with respect to s. Considering the squared first factors, applying the mean value theorem shows that $\int_{-\infty}^{t}\bar{D}_j^2(\delta,t,s)\mathrm{d}s = O(\delta^2)$ for $j=1$ and $j=3$. In the case of $\bar{D}_1$, $1_{\{t+\delta<0\}}=0$ at the same time as $1_{\{t<0\}}=1$ only if $-\delta<t<0$, so this does not happen whenever δ is small enough. In the case of (6.21), the integral of the squared second factor of $\bar{D}_2$ needs to be bounded. To do this put g_2^* for the maximizing value of g_2 and then make the change of variable $x=(g_2^*-s)/\delta$ in the integral, and so verify that $\int_{-\infty}^{t}\bar{D}_2^2(\delta,t,s)\mathrm{d}s = O(\delta^{2d_x+1})$. These orders of magnitude are all independent of t, and (iv) follows. ∎

6.6 Theorem Defining e_{npm} as in (5.8),

(i) $\limsup_n n^{-1}K(n)^{-2}\sum_{m=-\infty}^{[ns]} e_{n[ns]m}^2 < \infty$ for each fixed $s\in(-\infty,1]$.

(ii) $\limsup_n n^{-1}K(n)^{-2}\sum_{m=-\infty}^{[ns]} e_{n[ns]m}^2 = O((-s)^{2d_y+2d_x-3})$ as $s\to-\infty$.

(iii) $\sup_{s\in(-\infty,1-\delta]}\limsup_n n^{-1}K(n)^{-2}\sum_{m=[ns]+1}^{[n(s+\delta)]} e_{n[ns]m}^2 = O(\delta)$.

(iv) $\sup_{s\in(-\infty,1-\delta]}\limsup_n n^{-1}K(n)^{-2}\sum_{m=-\infty}^{[ns]} (e_{n[n(s+\delta)]m}-e_{n[ns]m})^2$

$$= O(\delta^{2d_y+1}).$$

Proof To show (i) and (ii), similarly to (6.16),

$$\limsup_n \frac{1}{nK(n)^2}\sum_{m=0}^{[ns]} e_{n[ns]m}^2 \le \int_0^s \bar{E}^2(s,t)\mathrm{d}t < \infty$$

and similarly to (6.17),

$$\limsup_n \frac{1}{nK(n)^2}\sum_{m=-\infty}^{\min\{-1,[ns]\}} e_{n[ns]m}^2 \le \int_{-\infty}^{\min\{0,s\}} \bar{E}^2(s,t)\mathrm{d}t = O((-s)^{2d_x+2d_y-3}).$$

The proof of (iii) is simpler than in Theorem **6.5**(iii) in view of (6.10). The result

$$\limsup_n \frac{1}{nK(n)^2}\sum_{m=[ns]+1}^{[n(s+\delta)]} e_{n[n(s+\delta)]m}^2 \le \int_s^{s+\delta} \bar{E}^2(s+\delta,t)\mathrm{d}t = O(\delta)$$

holds similarly to (6.18).

For part (iv), the decomposition analogous to (6.19) for $0<q<n-p$ is

$$e_{n,p+q,m} = \left(\sum_{j=\max\{0,1-p-q\}}^{\max\{0,1-p\}-1} + \sum_{j=\max\{0,1-p\}}^{n-1-p-q}\right) b_j \sum_{l=j+p+q+1-m}^{n-m} c_l$$

$$e_{npm} = \Bigg(\sum_{j=\max\{0,1-p\}}^{n-1-p-q} + \sum_{n-p-q}^{n-1-p} \Bigg) b_j \Bigg(\sum_{l=j+p-m}^{j+p+q-m} + \sum_{l=j+p+q+1-m}^{n-m} \Bigg) c_l.$$

It can be verified that in this case

$$e_{n,p+q,m} - e_{npm} = D_{1n}(q,p,m) - D_{2n}(q,p,m) - D_{3n}(q,p,m)$$

where

$$D_{1n}(q,p,m) = \sum_{j=\max\{0,1-p-q\}}^{\max\{0,1-p\}-1} b_j \sum_{l=j+p+q+1-m}^{n-m} c_l$$

$$D_{2n}(q,p,m) = \sum_{n-p-q}^{n-1-p} b_j \sum_{l=j+p+q+1-m}^{n-m} c_l$$

$$D_{3n}(q,p,m) = \sum_{j=\max\{0,1-p\}}^{n-1-p} b_j \sum_{l=j+p-m}^{j+p+q-m} c_l.$$

D_{1n} vanishes unless $p \leq 0$. For $0 < \delta < 1-s$ the inequalities

$$\frac{|D_{jn}([n\delta],[ns],[nt])|}{K(n)} \leq \bar{D}_j(\delta,s,t) + o(1)$$

hold for $j = 1, 2$ and 3, where

$$\begin{aligned} \bar{D}_1(\delta,s,t) = 1_{\{s<0\}}&|(-s)^{d_x} - 1_{\{s+\delta<0\}}(-s-\delta)^{d_x}| \\ &\times \max_{\max\{-s-\delta,0\}\leq g_1 \leq \max\{-s,0\}} |(1-t)^{d_y} - (g_1+s+\delta-t)^{d_y}| \end{aligned} \tag{6.23}$$

$$\begin{aligned} \bar{D}_2(\delta,s,t) = |(1-s)^{d_x} &- (1-s-\delta)^{d_x}| \\ &\times \max_{1-s-\delta\leq g_2\leq 1-s} |(1-t)^{d_y} - (g_2+s+\delta-t)^{d_y}| \end{aligned} \tag{6.24}$$

$$\begin{aligned} \bar{D}_3(\delta,s,t) = |(1-s)^{d_x} &- 1_{\{s<0\}}(-s)^{d_x}| \\ &\times \max_{1-s\leq g_3\leq \max\{0,-s\}} |(g_3+t+\delta-s)^{d_y} - (g_3+t-s)^{d_y}|. \end{aligned} \tag{6.25}$$

and

$$\limsup_{n\to\infty} \frac{1}{nK(n)^2} \sum_{p=-\infty}^{[nt]} (e_{n[n(s+\delta)]m} - e_{n[ns]m})^2 \leq 3\sum_{j=1}^{3} \int_{-\infty}^{t} \bar{D}_j^2(\delta,s,t)\mathrm{d}t.$$

The analysis is now very similar to the proof of **6.5**(iv). The squared second factors of $\bar{D}_1$ in (6.23) and $\bar{D}_2$ in (6.24), are integrable with respect to t. The mean value theorem applied to the first factors therefore gives $\int_{-\infty}^{t} \bar{D}_j^2(\delta,s,t)\mathrm{d}s = O(\delta^2)$ for $j = 1$ and $j = 2$ where in the case of $\bar{D}_1$, $1_{\{s+\delta<0\}} = 0$ and $1_{\{s<0\}} = 1$ only if $\delta > -s$. In the case of (6.25), put g_3^* for the maximizing value of g_3, make the change of variable $x = (g_3^* + t - s)/\delta$ and so verify that $\int_{-\infty}^{t} \bar{D}_3^2(\delta,s,t)\mathrm{d}s = O(\delta^{2d_y+1})$. ∎

6.2 The Main Result

According to (6.4) and (6.7), the finite sums G_{1n}^N and G_{3n}^N differ from G_{1n} and G_{3n} by tail components that are shown, respectively by Theorems **6.3** and **6.4**, to be negligible in L_2-norm when N is large enough. The sequences $\{w_m\}$ and $\{u_p\}$ as specified in Assumption **4.1** are mapped to càdlàg processes on the real line segment by writing $w_n^N(t) = w_{[nt]}$, $t \in [-N, 1]$ and $u_n^N(s) = u_{[ns]}$, $s \in [-N, 1]$. In the same way, the discrete arrays $\{q_{nm}^N\}$ from (6.3) and $\{h_{np}^N\}$ from (6.6), are mapped to $D_{[-N,1]}$ by the assignments

$$q_n^N(t) = q_{n[nt]}^N,\ t \in [-N, 1] \tag{6.26}$$

and

$$h_n^N(s) = h_{n[ns]}^N,\ s \in [-N, 1]. \tag{6.27}$$

The counterpart truncations for the putative limit processes (5.18) and (5.24) are $\Xi_1^N = \int_{-N}^1 Q^N(t)\mathrm{d}W(t)$, where $Q^N(t) = \int_{-N}^t A(t,s)\mathrm{d}U(s)$; and also $\Xi_3^N = \int_{-N}^1 H^N(s)\mathrm{d}U(s)$ where $H^N(s) = \int_{-N}^s E(s,t)\mathrm{d}W(t)$. By analogy with (2.18), Theorems **6.3** and **6.4** provide the bounds required to show the orders of magnitude of the respective remainders, which are as follows.

6.7 Corollary Under Assumption **4.1**,

$$\mathrm{E}(\Xi_1 - \Xi_1^N)^2 = O(N^{\max\{2d_y-1, 2(d_x+d_y-1)\}}).$$

Proof Lemma **6.1**, gives $A^2(t,s) \le \bar{A}^2(t,s)$ and $A^{*2}(t,s) \le \bar{A}^2(t,s)$ and hence

$$\begin{aligned}\mathrm{E}(\Xi_1 - \Xi_1^N)^2 &= \mathrm{E}\left(\int_{-\infty}^{-N} Q(t)\mathrm{d}W(t) + \int_{-N}^{1}(Q(t) - Q^N(t))\mathrm{d}W(t)\right)^2 \\ &\le 2\sigma_u^2\sigma_w^2\left(\int_{-\infty}^{-N}\int_{-\infty}^{t}\bar{A}^2(t,s)\mathrm{d}s\mathrm{d}t + \int_{-N}^{1}\int_{-\infty}^{-N}\bar{A}^2(t,s)\mathrm{d}s\mathrm{d}t\right).\end{aligned}$$

The result follows by (6.12) and (6.13). ∎

6.8 Corollary Under Assumption **4.1**,

$$\mathrm{E}(\Xi_3 - \Xi_3^N)^2 = O(N^{\max\{2d_y-1, 2(d_x+d_y-1)\}}).$$

Proof Similarly, by Lemma **6.2**,

$$\begin{aligned}\mathrm{E}(\Xi_3 - \Xi_3^N)^2 &= \mathrm{E}\left(\int_{-\infty}^{-N} H(s)\mathrm{d}U(s) + \int_{-N}^{1}(H(s) - H^N(s))\mathrm{d}U(s)\right)^2 \\ &\le 2\sigma_u^2\sigma_w^2\left(\int_{-\infty}^{-N}\int_{-\infty}^{s}\bar{E}^2(s,t)\mathrm{d}t\mathrm{d}s + \int_{-N}^{1}\int_{-\infty}^{-N}\bar{E}^2(s,t)\mathrm{d}t\mathrm{d}s\right).\end{aligned}$$

The result follows by (6.14) and (6.15). ∎

The weak convergence results of this chapter are given for the finite-domain processes that are indicated with superscript N. Corollaries **6.7** and **6.8** provide the means to link these results with limit processes Ξ_1 and Ξ_3, on the basis that by taking N sufficiently large the remainders can be made as small as desired in L_2 norm. The assertions that $G_{1n} \to_{\mathrm{d}} \Xi_1$ and $G_{3n} \to_{\mathrm{d}} \Xi_3$ and hence in particular the convergence to the third of the triplet of limits in (6.1) are to be understood in this light.

In addition to Assumption **4.1**, which has sufficed for the results up to date, a further assumption is needed for the weak convergence of q_n^N and h_n^N.

6.9 Assumption $d_x + d_y > -\frac{1}{2}$ and

(a) $\{u_i\}_{i=-\infty}^{\infty}$ is L_r-bounded for $r \geq \max\{2,\, 1/(\frac{1}{2}+d_x),\, 1/(\frac{1}{2}+d_x+d_y)\}$

(b) $\{w_i\}_{i=-\infty}^{\infty}$ is L_r-bounded for $r \geq \max\{2,\, 1/(\frac{1}{2}+d_y),\, 1/(\frac{1}{2}+d_x+d_y)\}$. □

This assumption incorporates Assumption **3.1** for u_i and the counterpart condition for w_i, so that Theorem **3.2** applies for both variables, but there are also new conditions to handle the additional weak convergences arising in this chapter. These can be seen to reflect the same concerns that motivated Assumption **3.1** and are needed for the results appearing in §6.4.

The convergence of the third element of (6.1) is a consequence of the following joint weak convergence in the space $D^4_{[-N,1]} \times \mathbb{R}^2$, where $D^4_{[-N,1]}$ is endowed with the Skorokhod topology. Pages 20 and 53 provide some related details here, noting that the same considerations arise when the domain of the functions is generalized beyond $[0,1]$.

6.10 Theorem Under Assumptions **4.1** and **6.9**,

$$\left(w_n^N, q_n^N, u_n^N, h_n^N, G_{1n}^N, G_{3n}^N\right) \xrightarrow{\mathrm{d}} \left(W^N, Q^N, U^N, H^N, \Xi_1^N, \Xi_3^N\right) \tag{6.28}$$

where (W^N, Q^N, U^N, H^N) are Gaussian, almost surely continuous, and adapted to a common filtration $\{\mathcal{F}(t), t \in [-N,1]\}$ with respect to which W^N and U^N are martingales. □

Of the six instances of weak convergence indicated in (6.28), the cases of w_n^N and u_n^N are standard results. The conditions of Theorem **3.2** are satisfied by these directly, with $d = 0$. W^N and U^N are Brownian motions on $[-N,1]$ with $W^N(0) = U^N(0) = 0$ and variances $\mathrm{E}(W^N(1)^2) = \sigma_w^2$ and $\mathrm{E}(U^N(1)^2) = \sigma_u^2$. These processes contribute to the covariances only in increment form, so their location is arbitrary and can be assumed as such without loss of generality.

Requiring special attention by contrast are the weak limits of q_n^N in (6.26) and h_n^N in (6.27). Although these processes have independent increments by Assumption **4.1**, the arrays of moving average coefficients a_{nmp} and e_{npm} are messy functions of their arguments. Unlike Brownian motions, q_n^N and h_n^N evolve by re-weighting existing increments as well as by appending new ones. Considering as the exemplar case (5.4) and its limit process (5.19), the coefficients of u_p in the partial sum q_{nm} as m proceeds depend on both m and p. This is not an eventuality allowed for even in FCLTs for processes with heteroscedastic increments. The one familiar case is $d_y = 0$, where q_n^N reduces to a truncated fractional Brownian motion according to (5.22) and Theorem **3.2** provides the requisite result.

Otherwise, the argument proceeds in two stages. The first stage makes use of the fact that for each t, $q_n^N(t)$ and $h_n^N(t)$ and the corresponding limit processes can be viewed similarly as the terminal points of partial sums defined on $[-N, t]$, having heteroscedastic but independent increments. Taking the limit process Q^N as the exemplar case, consider

$$\widetilde{Q}^N(t,\tau) = \int_{-N}^{\tau} A(t,s)\mathrm{d}U^N(s),\ \ \tau \in (-N, t].$$

This is a variance-transformed (heteroscedastic) Brownian motion, depending on a constant t, having increment weights that otherwise depend only the time s. Taking the terminal point $\tau = t$ in particular, $\widetilde{Q}^N(t,t) = Q^N(t)$ is then known, by a conventional FCLT allowing heteroscedasticity, to have a Gaussian distribution with known variance. The process $Q^N(t)$, $t \in (-N, 1]$ constructed from these limit points for each t and driven by the Brownian U^N has independent increments by construction.

What this argument does not show is the continuity of the limit processes, and the second stage of the proof, given in §6.4, is to establish stochastic equicontinuity. This can be done by some relatively minor variations of previously established arguments.

6.3 The Integrand Processes

The proofs for the cases of q_n^N and h_n^N are closely parallel and only that for q_n^N is given in full. The required result is the following.

6.11 Theorem Under Assumptions **4.1** and **6.9**(a), $q_n^N \to_{\mathrm{d}} Q^N$ where the limit is an almost surely continuous Gaussian process on the interval $[-N, 1]$. □

The proof of **6.11** is obtained by the combination of a pair of lemmas establishing, respectively, the Gaussianity and the almost surely continuity. Here is the first.

6.12 Lemma Under Assumption **4.1**, $Q^N(t)$ for $-N < t \leq 1$ is a Gaussian process having independent increments.

Proof The argument follows the lines of Theorem **3.2** except that the form of the moving average weights is different. Similarly to the decomposition of (3.2), it is convenient to split the sums into two components. Recalling the representations in (6.26) and (6.3), set

$$q_n^N(t) = q_{1n}(t) + q_{2n}^N(t) \tag{6.29}$$

where $q_{1n}(t)$ has the increments labelled $p = 1, \ldots, [nt]$ in the cases with $t > 0$ and $q_{1n}(t) = 0$ when $t \leq 0$, while $q_{2n}^N(t)$ has the increments labelled $p = 1 -$

$nN, \ldots, \min\{0, [nt]\}$. Let the postulated limits of $q_{1n}(t)$ and $q_{2n}^N(t)$ be denoted $Q_1(t)$ and $Q_2^N(t)$, summing to $Q^N(t)$. The cases $d_y > 0$ and $d_y < 0$ are treated separately, the logic being virtually the same but the formulae different.

Case 1: $d_y > 0$.

Decompose the function $A(t,s)$ in (5.12) as

$$A(t,s) = 1_{\{s>0\}} A_1(t,s) + 1_{\{s \le 0\}} A_2(t,s) \tag{6.30}$$

where A_1 is the first term of (5.12). According to Lemma **5.2**, since $t \ge s$ the moving average weights in $q_{1n}(t)$ have the form $K(n)^{-1} a_{n[nt][ns]} = A_1(t,s) + o(1)$, where, using the representation in (5.15) for $s < t$,

$$A_1(t,s) = (1-t)^{d_y} (t-s)^{d_x} Z_1^A(t,s). \tag{6.31}$$

For fixed $t \in (0,1)$ set $p = [ns]$ and so write the array of moving average coefficients from (6.3) as $\{c_{np}\}_{p=1}^{[nt]}$ where according to (6.31) for $s < t$,

$$c_{np} = \frac{|a_{n,[nt],[nt]-p+1}|}{\sqrt{n}K(n)} \sim (1-t)^{d_y} Z_1^A(t,s) \frac{([nt]-p)^{d_x}}{n^{d_x+1/2}}. \tag{6.32}$$

The slowly varying components cancel in the limit. It is legitimate to write $(1-t)/(t-s) \sim (n-[nt])/([nt]-p)$ in (5.10) and hence to treat Z_1^A as a function of p and n, but $Z_1^A = O(1)$ as $n \to \infty$ for all $s < t$. The point $s = t$ must be excluded from the equivalence in (6.32), noting that Z_1^A diverges at that point if $d_x > 0$ and equals zero if $d_x < 0$.

To verify limiting Gaussianity, apply the Lindeberg condition test of Lemma **3.4** directly with c_{np}^2 defined in (6.32) replacing $a_{ni}^2/\kappa(n)^2$ in (3.8). Similarly to (3.9),

$$\sum_{p=1}^{[nt]} c_{np}^2 \mathrm{E}(u_p^2 1_{\{|c_{np}u_p|>\varepsilon\}}) \le \max_{1\le p\le [nt]} \mathrm{E}(u_p^2 1_{\{|u_p|>\varepsilon/c_{np}\}}) \sum_{p=1}^{[nt]} c_{np}^2 \tag{6.33}$$

where $\sum_{p=1}^{[nt]} c_{np}^2 = O(1)$ according to (6.32). In the case $d_x > 0$ the maximum in question is at $p = 1$ and $c_{n1} = O(n^{-1/2})$. If $d_x < 0$ on the other hand, the maximum in (6.33) is at $p = [nt]$. The representation (6.32) is not useful here but it is convenient to refer directly to (5.5) which shows that

$$\begin{aligned} a_{n[nt][nt]} &= c_0 b_0 + c_1(b_0 + b_1) + \cdots + c_{n-1-[nt]}(b_0 + \cdots + b_{n-1-[nt]}) \\ &= O(n^{d_y} L_x(n) L_y(n)) \end{aligned}$$

so $c_{n[nt]} = O(n^{-d_x-1/2})$. Theorem **A.4** of Appendix A therefore gives, similarly to (3.10),

$$\max_{1\le p\le [nt]} \mathrm{E}(u_p^2 1_{\{|u_p|>\varepsilon/c_{np}\}}) = \begin{cases} o(n^{1-r/2}), & d_x > 0 \\ o(n^{(d_x+1/2)(2-r)}), & d_x < 0. \end{cases} \tag{6.34}$$

These conditions match those of (3.10) and (3.11), and confirm that the Lindeberg condition holds in both cases. Also note the limiting variance function,

$$\mathrm{E}\big(q_{1n}(t)^2\big) \to \sigma_u^2 \int_0^t A_1^2(t,s)\mathrm{d}s. \tag{6.35}$$

Now consider $q_{2n}^N(t)$, where $s \leq 0$ and according to Lemma **5.2** the moving average weights have the form $K(n)^{-1}a_{n[nt][ns]} = A_2(t,s) + o(1)$ where by (5.12) and (5.15), for $t > s$,

$$\begin{aligned} A_2(t,s) = (1-t)^{d_y}\big((t-s)^{d_x} Z_1^A(t,s) - (-s)^{d_x}\big) \\ - 1_{\{t<0\}}(-t)^{d_y}\big((t-s)^{d_x} Z_2^A(t,s) - (-s)^{d_x}\big). \end{aligned} \tag{6.36}$$

As s ranges over $[-N, \min\{0,t\}]$ with t fixed, a salient feature of the function Z_1^A is that

$$Z_1^A(t,s) - 1 = d_y \int_0^1 \tau^{d_y - 1}\left(\left(\frac{1-t}{t-s}\tau + 1\right)^{d_x} - 1\right)\mathrm{d}\tau = O((-s)^{-1}) \tag{6.37}$$

and $Z_2^A(t,s) = 1 + O((-s)^{-1})$ similarly.

In the case $t > 0$, (6.36) gives

$$|A_2(t,s)| = \left|(1-t)^{d_y}\left((t-s)^{d_x} - (-s)^{d_x} + (t-s)^{d_x}(Z_1^A(t,s) - 1)\right)\right| \tag{6.38}$$

and as $s \to -\infty$,

$$(t-s)^{d_x} - (-s)^{d_x} \sim td_x(-s)^{d_x - 1}. \tag{6.39}$$

In view of (6.37), this implies $|A_2(t,s)| = O((-s)^{d_x-1})$. Setting $p = [ns]$ as before, but this time in the range $-nN \leq p \leq 0$, in view of (6.39) the constant array has the form

$$c_{np} = \frac{|a_{n,[nt],[nt]-p}|}{\sqrt{n}K(n)} \backsimeq \frac{|p|^{d_x-1}}{n^{d_x+1/2}} \tag{6.40}$$

where the notation '$\backsimeq$' is here to be interpreted specifically to mean that the approximating expression contains a factor that is not written explicitly but can be treated as $O(1)$ as $n \to \infty$. The parallel is with (6.32) but the more complicated form of the extra factor is left implicit here, to be extracted from (6.38) if desired.

If $t < 0$ then the second term of (6.36) must also be included. However, this has the form

$$\begin{aligned} &(-t)^{d_y}\big((t-s)^{d_x} Z_2(t,s) - (-s)^{d_x}\big) \\ &\quad = (-t)^{d_y}\big(((t-s)^{d_x} - (-s)^{d_x}) + (t-s)^{d_x}(Z_2^A(t,s) - 1)\big) \\ &\quad = O((-s)^{d_x-1}) \end{aligned} \tag{6.41}$$

where the indicated order of magnitude follows by the foregoing arguments. Hence, $|A_2(t,s)| = O((-s)^{d_x-1})$ holds for all finite t and (6.40) is validated for all t, likewise.

The trick introduced on page 35 can be adopted here. Condition (3.12) was verified by dividing the negative domain of the process into N blocks of length n observations, labelled $k = 0, \dots, N-1$ so that $1 - n(k+1) \le p \le -nk$ in the k^{th} block. Applying the same arrangement, consider first the case $t > 0$. If $k = 0$, then

$$\sum_{p=1-n}^{0} c_{np}^2 \mathrm{E}(u_p^2 1_{\{|c_{np}u_p|>\varepsilon\}}) \le \max_{1-n\le p\le 0} \mathrm{E}(u_p^2 1_{\{|u_p|>\varepsilon/c_{np}\}}) \sum_{p=1-n}^{0} c_{np}^2$$

and the maximum of c_{np} is at $p = 0$, whether $d_x > 0$ or $d_x < 0$. According to formula (5.5), $c_{n0} = O(n^{-1/2})$ when $d_x > 0$, but $c_{n0} = O(n^{-d_x-1/2})$ when $d_x < 0$. Therefore, similarly to (6.34),

$$\max_{1-n\le p\le 0} \mathrm{E}(u_p^2 1_{\{|u_p|>\varepsilon/c_{np}\}}) = \begin{cases} o(n^{1-r/2}), & d_x > 0 \\ o(n^{(d_x+1/2)(2-r)}), & d_x < 0. \end{cases}$$

When $k \ge 1$, the maximum of c_{np} in each of these blocks is found at $p = -nk$ and according to (6.40),

$$c_{n,-nk}^2 \backsimeq \frac{|-nk|^{2d_x-2}}{n^{2d_x+1}} = O(n^{-3}k^{2d_x-2}).$$

Theorem **A.4** gives after simplification (compare (3.15))

$$\begin{aligned}\sum_{p=1-n(k+1)}^{-nk} c_{np}^2 \mathrm{E}(u_p^2 1_{\{|c_{np}u_p|>\varepsilon\}}) &\le n \max_{1-n(k+1)\le p\le -nk} c_{np}^2 \mathrm{E}(u_p^2 1_{\{|u_p|>\varepsilon/c_{np}\}}) \\ &= o(n^{1-3r/2}k^{(d_x-1)r}).\end{aligned}$$

These bounds are of small order in n and also summable over k for $d_x < \frac{1}{2}$. Since the N blocks are independent by assumption, their sum $q_{2n}^N(t)$ converges to a Gaussian limit $Q_2^N(t)$.

The last case to be considered is $t \le 0$. This requires $A_2(t,s)$ in (6.36) to include the second term, but in view of (6.41), the same arguments operate in just the same way in this case, apart from the sum of terms being initialized at the point $p = [nt]$ instead of $p = 0$.

The limiting variance is

$$\mathrm{E}\big(q_{2n}^N(t)^2\big) \to \sigma_u^2 \int_{-N}^{\min\{t,0\}} A_2^2(t,s)\mathrm{d}s \tag{6.42}$$

which, since the integrand is $O((-s)^{2d_x-2})$, is finite in the limit as $N \to \infty$.

Case 2: $d_y < 0$.

This is dealt with a similar manner to Case 1, but in place of A_1 and A_2 the decomposition is applied to $A^*(t,s)$ from (5.29). Using the substitutions in (5.30),

$$A^*(t,s) = 1_{\{s>0\}}A_1^*(t,s) + 1_{\{s\le 0\}}A_2^*(t,s) \tag{6.43}$$

where for $s < t$,

$$A_1^*(t,s) = (1-t)^{d_y}(1-s)^{d_x} - \frac{d_x}{d_y+1}(1-t)^{d_y+1}(t-s)^{d_x-1}Z_1^{A*}(t,s) \quad (6.44)$$

with $Z_1^{A*}(t,s)$ from (5.27) and also, with $Z_2^{A*}(t,s)$ from (5.28),

$$\begin{aligned} A_2^*(t,s) = (1-t)^{d_y}\big((1-s)^{d_x} - (-s)^{d_x}\big) - \frac{d_x}{d_y+1}(t-s)^{d_x-1} \\ \times \Big((1-t)^{d_y+1}Z_1^{A*}(t,s) - 1_{\{t<0\}}(-t)^{d_y+1}Z_2^{A*}(t,s)\Big). \end{aligned} \quad (6.45)$$

It may be verified in the same manner as before that $Z_1^{A*}(t,s) = 1 + O((-s)^{-1})$ and $Z_2^{A*}(t,s) = 1 + O((-s)^{-1})$.

Considering the case $q_{1n}(t)$, with $0 < s < t$, the same sequence of arguments can be followed up to (6.32). However, now rewrite (6.44) in the form

$$A_1^*(t,s) = (1-t)^{d_y}(t-s)^{d_x}\left(\left(\frac{1-s}{t-s}\right)^{d_x} - \frac{d_x}{d_y+1}\frac{1-t}{t-s}Z_1^{A*}(t,s)\right). \quad (6.46)$$

As before, set $p = [ns]$ with t fixed and so similarly to the case in (6.32) write

$$c_{np} \backsimeq \frac{([nt]-p)^{d_x}}{n^{d_x+1/2}}$$

where the symbol '$\backsimeq$' denotes limiting proportionality with the bracketed term in (6.46). While depending on s, this is $O(1)$ as $s \to 0$. The Lindeberg condition is verified as in (6.33).

For the case $s \leq 0$, observe that (6.45) can be rearranged according to the mean value theorem as

$$\begin{aligned} A_2^*(t,s) = (t-s)^{d_x-1}\Bigg((1-t)^{d_y}d_x\left(\frac{\lambda-s}{t-s}\right)^{d_x-1} \\ - \frac{d_x}{d_y+1}\Big(\big((1-t)^{d_y+1} - 1_{\{t<0\}}(-t)^{d_y+1}\big) + (1-t)^{d_y+1}\big(Z_1^{A*}(t,s)-1\big) \\ - 1_{\{t<0\}}(-t)^{d_y+1}\big(Z_2^{A*}(t,s)-1\big)\Big)\Bigg) \end{aligned} \quad (6.47)$$

with $\lambda \in [0,1]$ depending on s, but nonetheless the expression in large parentheses is $O(1)$ as $s \to -\infty$ with t fixed, with the last two terms being of $O((-s)^{-1})$. In this range, therefore,

$$c_{np} \backsimeq \frac{|p|^{d_x-1}}{n^{d_x+1/2}}$$

as in the matching result in (6.40). With (6.47) in place of (6.45), the Lindeberg condition can be checked in the same way as for $d_y > 0$.

This completes the demonstration for the case $d_y < 0$ and hence the proof. ∎

Combining the asymptotic variances appearing in (6.35) and (6.42), assuming N is taken large enough that the tail component is negligible, the variance process of

$Q(t)$ is given in (5.20) for the case $d_y > 0$, but for full generality should be written as

$$\mathrm{E}(Q(t)^2) = \begin{cases} \sigma_u^2 \int_{-\infty}^t A^2(t,s)\mathrm{d}s, & d_y > 0 \\ \sigma_u^2 \int_{-\infty}^t A^{*2}(t,s)\mathrm{d}s, & d_y < 0 \end{cases}. \tag{6.48}$$

6.4 Almost Sure Continuity

Lemma **6.12** has shown that the finite-dimensional distributions of the partial sum processes $Q^N(t)$ for each t are Gaussian. The process Q^N is composed of the terminal points of these processes, and hence is Gaussian with variance (6.48) at each point t. It remains to show that Q^N is almost surely continuous. Stochastic equicontinuity can be shown by adapting the type of argument used in **3.2**, although the two cases differ in important ways. Lemmas **3.6** and **3.8** were applied to establish the asymptotic continuity of the fractional process X_n on the interval $[0,1]$, while q_n^N is a process with independent increments on the interval $[-N,1]$. Nonetheless, thanks to the linear structure there are features that the two cases conveniently have in common.

Consider the decomposition

$$q_n^N(t+\delta) - q_n^N(t) = Y_{1n}(t+\delta,t) + Y_{2n}(t+\delta,t)$$

where

$$Y_{1n}(s,t) = \frac{1}{\sqrt{n}K(n)} \sum_{p=[nt]+1}^{[ns]} a_{n[ns]p}u_p \tag{6.49a}$$

$$Y_{2n}(s,t) = \frac{1}{\sqrt{n}K(n)} \sum_{p=1-nN}^{[nt]} (a_{n[ns]p} - a_{n[nt]p})u_p. \tag{6.49b}$$

Also define $\nu_n^2(t,\delta) = \nu_{1n}^2(t,\delta) + \nu_{2n}^2(t,\delta)$ where

$$\nu_{1n}^2(t,\delta) = \frac{\sigma_u^2}{nK(n)^2} \sum_{p=[nt]+1}^{[n(t+\delta)]} a_{n[n(t+\delta)]p}^2 \tag{6.50a}$$

$$\nu_{2n}^2(t,\delta) = \frac{\sigma_u^2}{nK(n)^2} \sum_{p=1-nN}^{[nt]} (a_{n[n(t+\delta)]p} - a_{n[nt]p})^2. \tag{6.50b}$$

Since Y_{1n} and Y_{2n} are non-overlapping sums,

$$\mathrm{E}(q_n^N(t+\delta) - q_n^N(t))^2 = \nu_n^2(t,\delta)$$

under Assumption **4.1**. There are parallels here with Theorem **3.5**, as the following result shows.

6.13 Theorem Under Assumption **4.1**, $\sup_{t\le s\le t+\delta}|Y_{1n}(s,t)|$ is bounded in probability if and only if **6.9**(a) holds.

Proof According to (5.5),

$$\begin{aligned} Y_{1n}(s,t) = \frac{1}{\sqrt{n}K(n)}\Big(&\big(c_0b_0 + c_1(b_0+b_1) + c_2(b_0+b_1+b_2)+\cdots\big)u_{[ns]} \\ &+\big((c_0(b_0+b_1)+c_1(b_0+b_1+b_2) \\ &\quad + c_2(b_0+b_1+b_2+b_3)+\cdots\big)u_{[ns]-1}+\cdots\Big). \end{aligned} \tag{6.51}$$

When $d_x + d_y < 0$ the coefficient of $u_{[ns]}$ in (6.51) is $O(n^{-(1/2+d_x+d_y)})$ as $n\to\infty$, since prior to normalization the terms are summable. The argument is identical to that of Theorem **3.5** except that $\kappa(n)$ in (3.20) is replaced with $\sqrt{n}K(n)$. Condition **6.9**(a) is sufficient for $(\sqrt{n}K(n))^{-r} = O(n^{-1})$. ∎

Similarly to Theorem **3.5** and (3.24) in the remark following, Theorem **6.13** shows that Assumption **6.9**(a) is necessary for uniform integrability of the q_n^N increments. Subject to this requirement, the main argument leading to stochastic equicontinuity is shown, just as in the univariate case, by a combination of Theorem **A.5** of Appendix A and a modified version of Lemma **3.6**. The moving average coefficients are constructed differently here from the univariate case, but the basic idea of finding a uniformly integrable dominating sequence for the supremum is the same.

Adopting the shorthand notation introduced for (3.18) on page 36, let

$$\widetilde{T}_n^2(t,\delta) = \sup_{\{t\le s\le t+\delta\}} \frac{(q_n^N(s)-q_n^N(t))^2}{\nu_n^2(t,\delta)} \tag{6.52}$$

and similarly define

$$\widetilde{Y}_{1n}(t,\delta) = \sup_{\{t\le s\le t+\delta\}}\frac{|Y_{1n}(s,t)|}{\nu_{1n}(t,\delta)}, \quad \widetilde{Y}_{2n}(t,\delta) = \sup_{\{t\le s\le t+\delta\}}\frac{|Y_{2n}(s,t)|}{\nu_{2n}(t,\delta)}$$

taking care to note the matching denominators.

6.14 Lemma Under Assumptions **4.1** and **6.9**, the collection $\{\widetilde{T}_n^2(t,\delta), n\in\mathbb{N}\}$ is uniformly integrable for any t in $(-N,1)$ and $0<\delta<1-t$, with

$$\mathrm{E}\big(\widetilde{T}_n^2(t,\delta)1_{\{\widetilde{T}_n^2(t,\delta)>\eta\}}\big) = o(\eta^{2-r}).$$

Proof Begin by noting that

$$\widetilde{T}_n^2(t,\delta) \le \big(\widetilde{Y}_{1n}(t,\delta)+\widetilde{Y}_{2n}(t,\delta)\big)^2. \tag{6.53}$$

The approach is to show that each term in the majorant of (6.53) satisfies the conditions of Theorem **A.5** and then to invoke Theorems **A.7** and **A.6**.

Starting with Y_{1n}, consider a modified version of formula (5.5). For fixed integer $b \in \{[nt], \dots, [n(t+\delta)]\}$ and likewise any $p \in \{[nt], \dots, [n(t+\delta)]\}$, define

$$\bar{a}_{nbp} = \sum_{l=\max\{0,1-b,p-b\}}^{n-1-b} c_l \Bigg(\sum_{j=\max\{0,1-p\}}^{l+b-p} b_j \Bigg). \tag{6.54}$$

The main difference from (5.5) is that p is allowed to exceed b. Defining r by $b = [nr]$, the analysis paralleling Lemma **5.2** gives the result

$$\frac{\bar{a}_{n[nr][ns]}}{K(n)} \to d_y \int_{\max\{0,-r,s-r\}}^{1-r} v^{d_y-1}\big((v+r-s)^{d_x} - 1_{\{s<0\}}(-s)^{d_x}\big)\mathrm{d}v. \tag{6.55}$$

Defining

$$\bar{v}_{1nb}^2(t,\delta) = \frac{\sigma_u^2}{nK(n)^2} \sum_{p=[nt]+1}^{[n(t+\delta)]} \bar{a}_{nbp}^2 \tag{6.56}$$

note the implication of (6.55) that $\bar{v}_{1nb}^2(t,\delta)/v_{1n}^2(t,\delta) = O(1)$ as $n \to \infty$.

In the partial sum sequence

$$\bar{Y}_{1nb}(s,t) = \frac{1}{\sqrt{n}K(n)} \sum_{p=[nt]+1}^{[ns]} \bar{a}_{nbp}u_p \tag{6.57}$$

the moving average weights do not depend on s, and this is also true of the squared weights $\bar{a}_{nbp}^2/v_{1n}^2(t,\delta)$ whose sum over p is noted to be $O(1)$. Similarly to (6.52) let

$$\tilde{\bar{Y}}{}_{1nb}^2(t,\delta) = \sup_{\{t\le s\le t+\delta\}} \frac{\bar{Y}_{1nb}^2(s,t)}{\nu_{1n}^2(t,\delta)} \tag{6.58}$$

taking care to note that the denominator is $\nu_{1n}^2(t,\delta)$, not $\bar{v}_{1nb}^2(t,\delta)$. According to Theorem **A.5** and the assumptions the sequence $\big\{\tilde{\bar{Y}}{}_{1nb}^2(t,\delta), n \in \mathbb{N}\big\}$ is uniformly integrable and satisfies the relation

$$\mathrm{E}\big(\tilde{\bar{Y}}{}_{1nb}^2(t,\delta)1_{\{\tilde{\bar{Y}}_{1nb}(t,\delta)>\eta\}}\big) = o(\eta^{2-r}).$$

This is true for any choice of b in the indicated range. Now set $b = [ns^*]$ where s^* is the solution to $\sup_{\{t\le s\le t+\delta\}} Y_{1n}^2(s,t)/\nu_{1n}^2(t,\delta)$, so that

$$\sup_{\{t\le s\le t+\delta\}} \frac{Y_{1n}^2(s,t)^2}{\nu_{1n}^2(t,\delta)} = \frac{\bar{Y}_{1n[ns^*]}^2(s^*,t)}{\nu_{1n}^2(t,\delta)} \le \sup_{\{t\le s\le t+\delta\}} \frac{\bar{Y}_{1n[ns^*]}^2(s,t)}{\nu_{1n}^2(t,\delta)}. \tag{6.59}$$

The collection of the majorants of (6.59) for each $n \in \mathbb{N}$ is uniformly integrable, since each member is drawn from one of the uniformly integrable collections defined for (6.58). The result that the collection $\big\{\widetilde{Y}_{1n}^2(t,\delta), n \in \mathbb{N}\big\}$ is uniformly integrable, with the property $\mathrm{E}\big(\widetilde{Y}_{1n}^2(t,\delta)1_{\{\widetilde{Y}_{1n}(t,\delta)>\eta\}}\big) = o(\eta^{2-r})$ for each $n \in \mathbb{N}$, follows by Theorem **A.6**.

The second term of (6.53) is not a partial sum and depends on s only through the definition of the moving average weights. For this sum, argue as follows. First, for each fixed $s \in [t, t+\delta]$, consider the square of the sum in (6.49b). After normalization by $\nu_{2n}^2(t, s-t)$ for the chosen s, all these sequences, including the supremum over the closed interval $[t, t+\delta]$, satisfy the conditions of Theorem **A.5** directly. To be precise, this conclusion follows by Theorem **A.6** in view of the fact that the squared terminal sum, having the full set of $[nt] + nN - 1$ terms, cannot exceed the supremum of the squared partial sums which Theorem **A.5** says is uniformly integrable. Similarly to before, let

$$\widetilde{Y}_{2n}^2(t,\delta) = \sup_{\{t \le s \le t+\delta\}} \frac{Y_{2n}^2(s,t)}{\nu_{2n}^2(t,\delta)}.$$

Since $\nu_{2n}^2(t, s-t) \le \nu_{2n}^2(t,\delta)$ when $s \le t+\delta$ and n is large enough, by a minor extension of Theorem **6.5**(iv), the uniformly integrability of $\{\widetilde{Y}_{2n}^2(t,\delta), n \in \mathbb{N}\}$ follows, with $\mathrm{E}\big(\widetilde{Y}_{2n}^2(t,\delta)1_{\{\widetilde{Y}_{2n}(t,\delta) > \eta\}}\big) = o(\eta^{2-r})$ for each $n \in \mathbb{N}$. The theorem now follows by application of Theorems **A.7** and **A.6 to** (6.53). ∎

At this point, it is possible to set out the formal proof of the weak convergence result from §6.3 for the process $q_n^N = \{q_n^N(t), t \in [-N, 1]\}$.

Proof of Theorem 6.11 Given Lemmas **6.12** and **6.14**, the proof for q_n^N aligns closely with that of Theorem **3.2**. To show stochastic equicontinuity, apply Theorems **3.7** and **3.8** with Y_n equated to q_n^N. In this application of the theorems, $L = -N$ and $U = 1$ and also $d = d_x$, with the sum of (6.50a) and (6.50b) playing the role of $\nu_n^2(t,\delta)$ in (3.17). Condition (a) of Theorem **3.7** is not problematic, given the composition of q_n^N as a linear process with independent L_r-bounded increments. Condition (b) of Theorem **3.7** is shown by confirming the conditions of Theorem **3.8**. Since $r \ge 1/(\frac{1}{2} + \min\{0, d_x, d_y + d_x\})$ by Assumption **6.9**(a), the exponent of δ in condition (3.42) with the substitution $d = d_x$ is nonnegative at worst, confirming that Theorem **3.8** holds in this case subject to confirmation of the stated conditions. Of these, condition **3.8**(b) is shown in Lemma **6.14** while condition **3.8**(a) holds by parts (iii) and (iv) of Lemma **6.5**. ∎

The limit process Q^N is a variance-transformed Brownian motion whose variance function as $N \to \infty$ is the function defined in (6.48).

The parallel result for h_n^N follows closely similar lines and is stated formally as follows.

6.15 Theorem Under Assumptions **4.1** and **6.9**(b), $h_n^N \to_{\mathrm{d}} H^N$ where the limit is an almost surely continuous Gaussian process on the interval $[-N, 1]$. □

The proof differs from that of Theorem **6.11** only because the functional forms of e_{npm} and hence of $E(s,t)$ and $E^*(s,t)$ replace those of a_{nmp}, $A(t,s)$ and $A^*(t,s)$. Making such substitutions as

$$c_{nm} = \frac{|e_{n,[ns],[ns]-m+1}|}{\sqrt{n}K(n)} \sim (1-s)^{d_x} Z_1^E(s,t) \frac{([ns]-m)^{d_y}}{n^{d_y+1/2}}$$

to replace (6.32), for example, the convergence criteria are derived in a wholly similar manner. The stochastic equicontinuity also follows by the argument paralleling Lemma **6.14**. Letting h_n^N replace q_n^N in (6.52) the same notation and text can be recycled, with the only significant variations being the replacement of $\bar{a}_{nbp}$ in (6.54) by

$$\bar{e}_{nbm} = \sum_{j=\max\{0,1-b,m-b\}}^{n-1-b} b_j \Bigg(\sum_{l=j+b+1-m}^{n-m} c_l \Bigg)$$

and the citation of Lemma **6.6** place of Lemma **6.5**. The variance function, in parallel with (6.48), has the form under Assumption **4.1** of

$$\mathrm{E}(H(s)^2) = \begin{cases} \sigma_w^2 \int_{-\infty}^{s} E^2(s,t)\mathrm{d}t & d_x > 0 \\ \sigma_w^2 \int_{-\infty}^{s} E^{*2}(s,t)\mathrm{d}t & d_x < 0. \end{cases} \tag{6.60}$$

6.5 Stochastic Integral Convergence

The next and final phase of the analysis is to show the convergence of the last two elements of (6.28). The following argument applies almost identically to G_{1n}^N and G_{3n}^N, with G_{1n}^N being taken as usual as the exemplar case.

6.16 Theorem Under Assumptions **4.1** and **6.9**, $G_{1n}^N \to_{\mathrm{d}} \Xi_1^N$.

Proof From (6.2) and (6.3),

$$G_{1n}^N = \frac{1}{nK(n)} \sum_{m=-nN}^{n-1} \sum_{p=-nN}^{m} a_{nmp} u_p w_{m+1}. \tag{6.61}$$

Choose an integer subsequence k_n, with $k_n \to \infty$ but $k_n/n \to 0$ as $n \to \infty$. Then, for $j = 0, \ldots, k_n$ let $n_j = [n(N+1)j/k_n] - nN$ so that $n_0 = -nN$, $n_{k_n} = n$, and $n_j - n_{j-1} \to \infty$ for every $j \geq 1$. Finally, define $t_j = n_j/n$ so that $\{t_0, \ldots, t_{k_n}\}$ is a partition of $[-N, 1]$ with the property

$$\max_{1 \leq j \leq k_n} (t_j - t_{j-1}) = O(1/k_n) \tag{6.62}$$

as $n \to \infty$.

With this notation, define a partially aggregated version of the covariance function,

$$\begin{aligned} G_{1n}^{N*} &= \sum_{j=1}^{k_n} q_n^N(t_{j-1})\big(w_n^N(t_j) - w_n^N(t_{j-1})\big) \\ &= \frac{1}{nK(n)} \sum_{j=1}^{k_n} \Bigg(\sum_{p=n_0}^{n_{j-1}} a_{nn_{j-1}p} u_p \sum_{m=n_{j-1}+1}^{n_j} w_m \Bigg). \end{aligned} \tag{6.63}$$

Consider the normalized increments in (6.63), which in the case of q_n^N are for each j the partial sums of the elements $q_{n,-nN}^N, \ldots, q_{n,n_{j-1}}^N$ as defined in (6.3). On the assumptions, both the inner sums in (6.63) converge to a.s. continuous Gaussian limits, respectively Q^N by Theorem **6.11** and W^N by Theorem **3.2** for the case $d = 0$. Replacing q_n^N and w_n^N in (6.63) by these limit processes, define

$$P_n = \sum_{j=1}^{k_n} Q^N(t_{j-1})\big(W^N(t_j) - W^N(t_{j-1})\big) = \sum_{j=1}^{k_n} Q^N(t_{j-1}) \int_{t_{j-1}}^{t_j} \mathrm{d}W^N(t). \quad (6.64)$$

The tasks of the proof are to show the connections between G_{1n}^N and G_{1n}^{N*}, P_n and Ξ_1^N, and finally between G_{1n}^{N*} and P_n. The sequences of the differences between these pairs of objects are shown in each case to converge to zero in mean square as $n \to \infty$.

The first of these remainders has the form

$$G_{1n}^N - G_{1n}^{N*} = \frac{1}{nK(n)} \sum_{j=1}^{k_n} \sum_{m=n_{j-1}+1}^{n_j-1} \Bigg(\sum_{p=n_{j-1}+1}^{m} a_{nmp} u_p + \sum_{p=n_0}^{n_{j-1}} (a_{nmp} - a_{nn_{j-1}p}) u_p \Bigg) w_{m+1}. \quad (6.65)$$

Since the shocks are L_2-bounded and serially independent and nowhere overlap,

$$\mathrm{E}(G_{1n}^N - G_{1n}^{N*})^2 = \frac{\sigma_u^2 \sigma_w^2}{n^2 K(n)^2} \sum_{j=1}^{k_n} \sum_{m=n_{j-1}+1}^{n_j-1} \Bigg(\sum_{p=n_{j-1}+1}^{m} a_{nmp}^2 + \sum_{p=n_0}^{n_{j-1}} (a_{nmp} - a_{nn_{j-1}p})^2 \Bigg). \quad (6.66)$$

Consider the two sums of squares in (6.66). By Theorem **6.5**(iii),

$$\begin{aligned} \frac{1}{n^2 K(n)^2} \sum_{j=1}^{k_n} \sum_{m=n_{j-1}+1}^{n_j-1} \sum_{p=n_{j-1}+1}^{m} a_{nmp}^2 &\ll \frac{1}{n} \sum_{j=1}^{k_n} \sum_{m=n_{j-1}+1}^{n_j-1} \Big(\frac{m - n_{j-1}}{n} \Big)^{\min\{1,1+2d_x\}} \\ &\ll \frac{1}{n^{\min\{2,2+2d_x\}}} \sum_{j=1}^{k_n} (n_j - n_{j-1})^{\min\{2,2+2d_x\}} \\ &= O\big(k_n^{-\min\{1,1+2d_x\}}\big). \end{aligned}$$

For the second block similarly, by Theorem **6.5**(iv),

$$\frac{1}{n^2 K(n)^2} \sum_{j=2}^{k_n} \sum_{m=n_{j-1}+1}^{n_j-1} \sum_{p=n_0}^{n_{j-1}} (a_{nmp} - a_{nn_{j-1}p})^2$$

$$\ll \frac{1}{n}\sum_{j=2}^{k_n}\sum_{m=n_{j-1}+1}^{n_j-1}\Big(\frac{m-n_{j-1}}{n}\Big)^{1+2d_x}$$
$$= O\big(k_n^{-1-2d_x}\big).$$

In other words,

$$\mathrm{E}(G_{1n}^N - G_{1n}^{N*})^2 = O\big(\max\{k_n^{-1}, k_n^{-1-2d_x}\}\big) \tag{6.67}$$

which with $d_x > -\frac{1}{2}$ implies $|G_{1n}^N - G_{1n}^{N*}| \to_{L_2} 0$ in all cases.

Next, compare P_n in (6.64) with $\Xi_1^N = \int_{-N}^1 Q^N(t)\mathrm{d}W^N(t)$. The fact that $\mathrm{E}\big(Q^N(t+\delta) - Q^N(t)\big)^2 = O(\delta^{2d_x+1})$ can be deduced from the formula in (5.21) combined with Lemma **6.1** and the calculations in parts (iii) and (iv) of Theorem **6.5**. Using this bound and the fact that W^N is a Brownian motion gives

$$\mathrm{E}\big(P_n - \Xi_1^N\big)^2 = \mathrm{E}\bigg(\sum_{j=1}^{k_n}\int_{t_{j-1}}^{t_j}\big(Q^N(t_{j-1}) - Q^N(t)\big)\mathrm{d}W^N(t)\bigg)^2$$
$$\ll \sum_{j=1}^{k_n}\int_{t_{j-1}}^{t_j}(t-t_{j-1})^{2d_x+1}\mathrm{d}t$$
$$\ll \sum_{j=1}^{k_n}(t_j - t_{j-1})^{2d_x+2} = O(k_n^{-2d_x-1}) \tag{6.68}$$

where the final order of magnitude is by (6.62). In other words, $|P_n - \Xi_1^N| \to_{L_2} 0$.

The remaining step is to connect G_{1n}^{N*} with P_n in (6.64), and this is done by applying the Skorokhod representation theorem (see [64]), according to which the joint weak convergence of q_n^N and w_n^N to Q^N and W^N implies the existence of processes that are distributed like q_n^N and w_n^N and converge almost surely to limits distributed like Q^N and W^N.[1] These so-called Skorokhod processes will be denoted $\widehat{q}_n^N$ and $\widehat{w}_n^N$.

If $\widehat{G}_{1n}^{N*}$ is the counterpart of G_{1n}^{N*} evaluated with $\widehat{q}_n^N$ and $\widehat{w}_n^N$ in place of q_n^N and w_n^N then $\widehat{G}_{1n}^{N*}$ and G_{1n}^{N*} have the same distribution. To complete the proof it therefore suffices to show that $|\widehat{G}_{1n}^{N*} - P_n| \to_{L_2} 0$. The first step in this demonstration is by way of the easily verifiable identity

$$\widehat{G}_{1n}^{N*} - P_n$$
$$= \sum_{j=1}^{k_n}\widehat{q}_n^N(t_{j-1})\big(\widehat{w}_n^N(t_j) - \widehat{w}_n^N(t_{j-1})\big) - \sum_{j=1}^{k_n}Q^N(t_{j-1})\big(W^N(t_j) - W^N(t_{j-1})\big)$$
$$= \sum_{j=1}^{k_n}\big(\widehat{q}_n^N(t_{j-1}) - Q^N(t_{j-1})\big)\big(\widehat{w}_n^N(t_j) - \widehat{w}_n^N(t_{j-1})\big)$$

[1] The Skorokhod theorem is proved as SLT Theorem 29.6.

$$- \sum_{j=1}^{k_n} \big(Q^N(t_j) - Q^N(t_{j-1})\big)\big(\widehat{w}_n^N(t_j) - W^N(t_j)\big)$$
$$+ Q^N(1)\big(\widehat{w}_n^N(1) - W^N(1)\big) - Q^N(-N)\big(\widehat{w}_n^N(-N) - W^N(-N)\big). \qquad (6.69)$$

It is sufficient to show that each of the four right-hand side terms of (6.69) converges to zero in mean square. Starting with the first one, the Cauchy-Schwarz inequality for sums gives

$$\begin{aligned}
&\left(\sum_{j=1}^{k_n} \big(\widehat{q}_n^N(t_{j-1}) - Q^N(t_{j-1})\big)\big(\widehat{w}_n^N(t_j) - \widehat{w}_n^N(t_{j-1})\big)\right)^2 \\
&\qquad \le \sum_{j=1}^{k_n} \big(\widehat{q}_n^N(t_{j-1}) - Q^N(t_{j-1})\big)^2 \sum_{j=1}^{k_n} \big(\widehat{w}_n^N(t_j) - \widehat{w}_n^N(t_{j-1})\big)^2 \\
&\qquad \le k_n \max_{1 \le j \le k_n} \big(\widehat{q}_n^N(t_{j-1}) - Q^N(t_{j-1})\big)^2 \sum_{j=1}^{k_n} \big(\widehat{w}_n^N(t_j) - \widehat{w}_n^N(t_{j-1})\big)^2. \qquad (6.70)
\end{aligned}$$

The object is to bound the expected value of the majorant of (6.70).

The process $\widehat{q}_n^N$ in (6.70) is càdlàg, a step function with discontinuities (see page 20 for details), although its almost sure limit Q^N is almost surely continuous according to Theorem **6.11**. A gentle digression into probability theory is necessary at this point. Let $(\Omega, \mathcal{F}, P)$ be the probability space where these random processes reside and let $\omega \in \Omega$ denote an outcome. By Egorov's theorem,[2] almost sure convergence implies uniform convergence on a set of outcomes $A_\varepsilon \in \mathcal{F}$, where $P(A_\varepsilon) \ge 1 - \varepsilon$ and $\varepsilon > 0$ is arbitrary. For càdlàg functions, which inhabit a space of processes endowed with the Skorokhod topology, what this amounts to is that if $\widehat{q}_n^N(\omega) \to Q^N(\omega)$ almost surely then, for A_ε so defined,

$$\sup_{\omega \in A_\varepsilon} d_S\big(\widehat{q}_n^N(\omega), Q^N(\omega)\big) \to 0$$

where d_S is the Skorokhod distance discussed on page 20. Since Q^N is almost surely continuous, there also exists $E_Q \in \mathcal{F}$ with $P(E_Q) = 1$ such that each $\omega \in E_Q$ has the following property: for any $\eta > 0$, there exists a constant $\delta > 0$ such that if

$$\sup_{t \in [-N,1]} \big|\widehat{q}_n^N(\omega, t) - Q^N(\omega, \lambda(t))\big| \le \delta$$

where $\lambda : [-N, 1] \mapsto [-N, 1]$ represents the homeomorphism defining the Skorokhod distance, then

$$\begin{aligned}
&\sup_{t \in [-N,1]} \big|\widehat{q}_n^N(\omega, t) - Q^N(\omega, t)\big| \\
&\qquad \le \sup_{t \in [-N,1]} \big|\widehat{q}_n^N(\omega, t) - Q^N(\omega, \lambda(t))\big| + \sup_{t \in [-N,1]} \big|Q^N(\omega, \lambda(t)) - Q^N(\omega, t)\big|
\end{aligned}$$

[2] Proved as SLT Theorem 19.4.

$$\leq \delta + \eta.$$

In other words, this says that when a càdlàg function is close to a continuous function the Skorokhod distance must be correspondingly close to the uniform distance. Letting $A_\varepsilon^* = A_\varepsilon \cap E_Q$ so that $P(A_\varepsilon^*) = P(A_\varepsilon)$, it follows that $\delta_n^q \to 0$ as $n \to \infty$ where

$$\delta_n^q = \sup_{\omega \in A_\varepsilon^*} \sup_{t \in [-N,1]} \left| \widehat{q}_n^N(\omega, t) - Q^N(\omega, t) \right|.$$

Noting that the distributions of w_n^N and $\widehat{w}_n^N$ are the same, (6.70) implies that

$$\begin{aligned} \mathrm{E}\bigg(1_{A_\varepsilon^*} \sum_{j=1}^{k_n} \big(\widehat{q}_n^N(t_{j-1}) - Q^N(t_{j-1})\big)\big(\widehat{w}_n^N(t_j) - \widehat{w}_n^N(t_{j-1})\big)\bigg)^2 \\ \leq k_n \delta_n^{q\,2} \sum_{j=1}^{k_n} \mathrm{E}\big(w_n^N(t_j) - w_n^N(t_{j-1})\big)^2 \\ = k_n \delta_n^{q\,2} \sum_{j=1}^{k_n} (t_j - t_{j-1})^2 = O(\delta_n^{q\,2}) \end{aligned}$$

where the final equality is by (6.62). Letting $\varepsilon \to 0$ and hence $P(A_\varepsilon^*) \to 1$, it follows that the first term of (6.69) converges to zero in mean square.

The same type of argument may be applied to the second term of (6.69). $\widehat{w}_n^N$ is also a càdlàg process converging with probability 1 to an almost surely continuous limit, and there exists a set A_ε^* with $P(A_\varepsilon^*) \geq 1 - \varepsilon$ such if

$$\delta_n^w = \sup_{\omega \in A_\varepsilon^*} \sup_{t \in [-N,1]} \left| \widehat{w}_n^N(\omega, t) - W^N(\omega, t) \right|$$

then $\delta_n^w \to 0$ as $n \to \infty$. It follows in this case that

$$\begin{aligned} \mathrm{E}\bigg(1_{A_\varepsilon^*} \sum_{j=1}^{k_n} \big(Q^N(t_j) - Q^N(t_{j-1})\big)\big(\widehat{w}_n^N(t_j) - W^N(t_j)\big)\bigg)^2 \\ \leq k_n \delta_n^{w2} \sum_{j=1}^{k_n} \mathrm{E}\big(Q^N(t_j) - Q^N(t_{j-1})\big)^2 \\ = O(k_n^{-2d_x} \delta_n^{w2}) \end{aligned}$$

where the order of magnitude is by reasoning similar to (6.68). When $d_x < 0$ it is necessary that k_n diverge slowly enough that $k_n^{-2d_x} \delta_n^{w2} \to 0$, which since $d_x > -\frac{1}{2}$ is possible under the assumptions. The expected squares of the third and fourth terms of (6.69) are of $O(\delta_n^{w2})$ and so vanish under the same assumptions. ∎

Under Assumption **4.1**, it appears reasonable to conjecture that $\delta_n^w = O(n^{-1/2})$. If that were the case, any $k_n = o(n)$ would satisfy the indicated convergence.

The counterpart result for G_{3n}^N is stated for completeness, variations in the argument of Theorem **6.16** being left for interested readers to supply.

6.17 Theorem Under Assumptions **4.1** and **6.9**, $G_{3n}^N \to_{\mathrm{d}} \Xi_3^N$. □

It remains to gather together the various results of this chapter to address the proposition initially posed.

Proof of Theorem 6.10 The different elements of (6.28) are accounted for, in the first instance, by Theorem **3.2** applied with $d = 0$ to give regular Brownian motion limits to both w_n^N and u_n^N. Theorems **6.11**, **6.15**, **6.16**, and **6.17** yield the remaining elementwise limits. The six elements are all adapted to the same filtration and by defining functions equated everywhere to the fixed values G_{1n}^N and G_{3n}^N they can be embedded in $D_{[-N,1]}^6$, converging to a limit lying in $C_{[-N,1]}^6$ almost surely. Theorem **3.19** gives the required joint convergence. ∎

Chapter 7

Fractional Cointegration

7.1 Stationary Regression

This chapter reviews some implications of the analysis of the preceding chapters for the interpretation of linear regressions. Consider first the stationary regression

$$p_i = \alpha + \beta x_i + y_i \tag{7.1}$$

where x_i and y_i are stationary zero-mean processes, α and β are parameters, and p_i is defined by the equation. The assumption to be maintained is that either or both of x_i and y_i are fractional processes, and that Assumption **4.1** applies. The leading case on which to fix ideas is where y is weakly dependent, with $d_y = 0$, although a long memory or antipersistent residual is permitted under suitable conditions. In this chapter it is convenient to omit the slowly varying components in (1.2) since these would complicate the notation while adding no useful insights. To incorporate them would basically involve modifying the normalization factors so that they disappear from limit expressions.

The OLS error-of-estimate for the slope coefficient β in (7.1), after the conventional normalization by $\sqrt{n}$, has the well-known formula

$$\sqrt{n}(\hat{\beta} - \beta) = \frac{n^{-1/2}\sum_i x_i y_i - n^{-3/2}\sum_i x_i \sum_i y_i}{n^{-1}\sum_i x_i^2 - \left(n^{-1}\sum_i x_i\right)^2}. \tag{7.2}$$

Consider what can be said about each of the terms in (7.2). If $\sigma_{uw} = 0$ and $d_x + d_y < \frac{1}{2}$, the first term of the numerator is asymptotically Gaussian by Theorem **4.3**, with variance $V_{xy} < \infty$ from (4.8). In the second term of the numerator, the two sums are respectively $O(n^{1/2+d_x})$ and $O(n^{1/2+d_y})$ by Corollary **2.7**, so under the same assumption, $1 + d_x + d_y < \frac{3}{2}$ and this term is of small order relative to the first. In the denominator, a case of Theorem **4.2** gives $n^{-1}\sum_i x_i^2 \to_{L_2} \sigma_x^2$ with the formula in (1.5), whereas $n^{-1}\sum_i x_i = o_p(1)$ by Corollary **2.7**, so here too the second term is of relatively small order. The conclusion is therefore that

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathrm{d}} \mathrm{N}\left(0, \frac{V_{xy}}{\sigma_x^4}\right). \tag{7.3}$$

In the case with $d_y = 0$ so that $y_i = w_i$ with variance σ_w^2, the remark following the proof of **4.3** on page 61 shows that $V_{xy} = \sigma_w^2\sigma_x^2$ and hence the variance of $\sqrt{n}(\hat{\beta} - \beta)$ has the usual limit of σ_w^2/σ_x^2. If the regressor in (7.1) has a nonzero mean, with the form $x_i + \mu$ where x_i has the specification of (1.1)+(1.2), since the regression formula expresses the data in the form of sample mean deviations it is a simple exercise to show that (7.3) continues to hold. The terms that would contain μ in (7.2) all cancel identically in the formula.

If $\sigma_{uw} \neq 0$ on the other hand, the estimator is well-known to be biased and inconsistent with

$$\hat{\beta} - \beta \xrightarrow{L_2} \frac{\sigma_{xy}}{\sigma_x^2}. \tag{7.4}$$

Also consider the intercept, where several different scenarios are possible since in this case the presence of a regressor mean μ does make a difference. If y is weakly dependent with $d_y = 0$, the normalized error-of-estimate formula is

$$\sqrt{n}(\hat{\alpha} - \alpha) = \frac{n^{-1}\sum_i x_i^2 n^{-1/2}\sum_i y_i - n^{-1}\sum_i x_i n^{-1/2}\sum_i x_i y_i}{n^{-1}\sum_i x_i^2 - \left(n^{-1}\sum_i x_i\right)^2}. \tag{7.5}$$

If x_i is replaced by $x_i + \mu$ in (7.5), a simple calculation gives

$$\sqrt{n}(\hat{\alpha} - \alpha) \sim \frac{1}{\sqrt{n}}\sum_{i=1}^{n} y_i - \mu\sqrt{n}(\hat{\beta} - \beta)$$

which, whether $\mu = 0$ or otherwise, has the same normalization as in the weakly dependent regressor case. However, if y is a long memory fractional process with $d_y > 0$ the required normalizing factor becomes $n^{1/2-d_y}$. The term containing μ, if present, is now of small order and

$$n^{1/2-d_y}(\hat{\alpha} - \alpha) \sim \frac{1}{n^{1/2+d_y}}\sum_{i=1}^{n} y_i \xrightarrow{\mathrm{d}} \mathrm{N}\left(0, \sigma_w^2\Upsilon_{d_y}\right). \tag{7.6}$$

Matters become more complicated if y is antipersistent with $d_y < 0$, although if $\mu = 0$ the equivalence in (7.6) continues to hold. The further ramifications of this case are left for the interested reader to explore.

7.2 Cointegrating Regression

The leading applications of regression to fractional processes involve cointegrating relations between processes featuring stochastic trends. Long memory and nonstationarity are closely connected in time series analysis and while nonstationarity is conventionally due to a unit root, this becomes a special case in the present analysis. The important questions to be investigated are sufficiently answered by consideration of the bivariate model

$$P_i = \alpha + \beta S_i + y_i \tag{7.7}$$

where now the regressor is the partial sum process $S_i = \sum_{j=1}^{i} x_j$, and x and y are stationary fractional processes with fractional parameters d_x and d_y. P_i is defined by (7.7), so that the model embodies cointegration of nonstationary processes P and S with cointegrating parameter β. Unlike the stationary case, it must be assumed for present purposes that $\mu = 0$ so that there is no drift in the stochastic trend. The relaxation of this restriction is treated in §7.4.

Write the OLS error-of-estimate formula for (7.7), similarly to (7.2) but this time in unadorned form, as

$$\hat{\beta} - \beta = \frac{\sum_i S_i y_i - n^{-1} \sum_i S_i \sum_i y_i}{\sum_i S_i^2 - n^{-1} \left(\sum_i S_i\right)^2}. \tag{7.8}$$

The data moments involved in this calculation, together with the weak limit of each under suitable normalization, are as follows. Define fBM processes X and Y, having the form of (2.1). Then, under Assumption **4.1**, Theorem **3.2** and the continuous mapping theorem give

$$\frac{1}{n^{d_x+3/2}} \sum_{i=1}^{n} S_i \xrightarrow{\mathrm{d}} \int_0^1 X(\xi)\mathrm{d}\xi \tag{7.9a}$$

$$\frac{1}{n^{2d_x+2}} \sum_{i=1}^{n} S_i^2 \xrightarrow{\mathrm{d}} \int_0^1 X^2(\xi)\mathrm{d}\xi \tag{7.9b}$$

$$\frac{1}{n^{d_y+1/2}} \sum_{i=1}^{n} y_i \xrightarrow{\mathrm{d}} Y(1). \tag{7.9c}$$

Similarly, by the various steps leading to (6.1) in Chapter 6 as well as Theorems **4.7** and **4.8** (see pages 62, 68, and 84 for the various symbol definitions)

$$\frac{1}{n^{d_x+d_y+1}} \sum_{i=1}^{n} S_i y_i \xrightarrow{\mathrm{d}} \Xi_{xy} + \sigma_{uw}\lambda_{xy} \text{ if } d_x + d_y > 0 \tag{7.10a}$$

$$\frac{1}{n^{d_x+d_y+1}} \sum_{i=1}^{n} S_i y_i \xrightarrow{\mathrm{d}} \Xi_{xy} \quad \text{if } -\tfrac{1}{2} < d_x + d_y \le 0 \text{ and } \sigma_{uw} = 0 \tag{7.10b}$$

$$\frac{1}{n\log n} \sum_{i=1}^{n} S_i y_i \xrightarrow{L_2} \gamma_{xy} \text{ if } d_x + d_y = 0 \text{ and } \sigma_{uw} \ne 0 \tag{7.10c}$$

$$\frac{1}{n} \sum_{i=1}^{n} S_i y_i \xrightarrow{L_2} \gamma_{xy} + \sigma_{xy} \text{ if } d_x + d_y < 0 \text{ and } \sigma_{uw} \ne 0. \tag{7.10d}$$

The restriction $d_x + d_y > 0$ in (7.10a) is imposed by the conditions of Theorem **4.4** while the restriction $d_x + d_y > -\frac{1}{2}$ in (7.10b) is due to Assumption **6.9**. Since $\mathrm{E}(G_n) = \mathrm{E}(G_{2n})$ where $G_n = (nK(n))^{-1} \sum_{i=1}^{n} (S_i y_i - x_i y_i)$, Theorem **4.7** provides the rationale for (7.10c) and (7.10d). In (7.10c), according to **4.7**(i) the expectation of the sum dominates the mean deviation by the factor $\log n$ and the contemporaneous covariance is also of smaller order and so doesn't appear.

The difference with case (7.10d) is that σ_{xy} appears in view of Theorem **4.2**. By Theorem **4.7**(ii) the correct normalization is now $1/n$ for both parts of the sum.

The second term in the numerator of (7.8) after normalization has the limit

$$\frac{1}{n^{d_x+d_y+2}}\sum_{i=1}^{n}S_i\sum_{j=1}^{n}y_j \xrightarrow{\mathrm{d}} \int_0^1 X\mathrm{d}\xi Y(1).$$

The mean of this random variable is also of interest and might be written as $\sigma_{uw}\kappa_{xy}$. To calculate κ_{xy}, let the companion function for y corresponding to $a_{ni}(s,t)$ in (2.25) for x be defined as

$$g_{nj}(s,t) = \sum_{l=\max\{0,[nt]+1-j\}}^{[ns]-j} c_l.$$

Then, under Assumption **4.1**,

$$\begin{aligned}
&\frac{1}{n^{d_x+d_y+2}}\mathrm{E}\bigg(\sum_{i=1}^{n}S_i\sum_{j=1}^{n}y_j\bigg)\\
&\qquad = \frac{1}{n^{d_x+d_y+2}}\mathrm{E}\bigg(\sum_{i=1}^{n}\sum_{k=-\infty}^{i}a_{nk}(i/n,0)u_k\sum_{j=-\infty}^{n}g_{nj}(1,0)w_j\bigg)\\
&\qquad = \frac{\sigma_{uw}}{n^{d_x+d_y+2}}\sum_{i=1}^{n}\sum_{k=-\infty}^{i}a_{nk}(i/n,0)g_{nk}(1,0). \qquad (7.11)
\end{aligned}$$

By similar reasoning to Theorem **4.4**, (7.11) converges to $\sigma_{uw}\kappa_{xy}$ where

$$\begin{aligned}
\kappa_{xy} = \int_0^1\bigg(&\int_0^t (t-\xi)^{d_x}(1-\xi)^{d_y}\mathrm{d}\xi\\
&+\int_{-\infty}^{0}\big((t-\xi)^{d_x}-(-\xi)^{d_x}\big)\big((1-\xi)^{d_y}-(-\xi)^{d_y}\big)\mathrm{d}\xi\bigg)\mathrm{d}t.
\end{aligned}$$

Putting these limits together allows the error-of-estimate distributions in the four cases of (7.10) to be calculated. In cases (7.10a) and (7.10b), writing $\int_0^1 X\mathrm{d}Y = \Xi_{xy} + \sigma_{uw}\lambda_{xy}$, the limiting distribution has the form

$$n^{1+d_x-d_y}(\hat{\beta}-\beta) \xrightarrow{\mathrm{d}} \frac{\int_0^1 X\mathrm{d}Y - \int_0^1 X\mathrm{d}\xi Y(1)}{\int_0^1 X^2\mathrm{d}\xi - \big(\int_0^1 X\mathrm{d}\xi\big)^2}. \qquad (7.12)$$

The estimator is consistent provided $d_y < 1 + d_x$ and this restriction holds for all cases under Assumption **4.1**. The phenomenon of cointegration requires only that y is stationary.

If the regressor is endogenous with $\sigma_{uw} \neq 0$ there is asymptotic bias, which in the unit root case is well known (see e.g. [3]) to play a decisive role in finite samples. In case (7.10a) the mean of the numerator in (7.12) is $\sigma_{uw}(\lambda_{xy} - \kappa_{xy})$.

Notwithstanding that the shocks are assumed serially independent under Assumption **4.1**, this asymptotic bias is not attributable to the regressor and regressand being dated contemporaneously. Writing $S_i = x_i + S_{i-1}$, the contribution of x_i to the form of λ_{xy} is negligible, as pointed out on page 64.

In case (7.10b) there is no bias, by construction. However, in cases (7.10c) and (7.10d), the mean of the numerator of (7.8) dominates the mean deviation and is of order $n \log n$ and n, respectively. The second numerator term is $O(n^{d_x+d_y+1})$ by (7.9a) and (7.9c) and so also of small order, under these normalizations. In case (7.10c) the limit distribution has the form

$$\frac{n^{1+2d_x}}{\log n}(\hat{\beta}-\beta) \xrightarrow{\mathrm{d}} \frac{\gamma_{xy}}{\int_0^1 X^2 \mathrm{d}\xi - \left(\int_0^1 X \mathrm{d}\xi\right)^2}. \tag{7.13}$$

In case (7.10d) it is

$$n^{1+2d_x}(\hat{\beta}-\beta) \xrightarrow{\mathrm{d}} \frac{\gamma_{xy}+\sigma_{xy}}{\int_0^1 X^2 \mathrm{d}\xi - \left(\int_0^1 X \mathrm{d}\xi\right)^2}. \tag{7.14}$$

Since these limits do not depend on Ξ_{xy} (the stochastic part of $\int_0^1 X \mathrm{d}Y$) the weak convergence of this component does not contribute and (7.13) and (7.14) hold even under the baseline condition of Assumption **4.1**(a), without the restriction on $d_x + d_y$ appearing in (7.10b).

Similarly to (7.5) the error of estimate for the intercept, unadorned in this case, is

$$\hat{\alpha}-\alpha = \frac{\sum_i S_i^2 \sum_i y_i - \sum_i S_i \sum_i S_i y_i}{n\sum_i S_i^2 - \left(\sum_i S_i\right)^2}. \tag{7.15}$$

In cases (7.10a) and (7.10b), according to (7.9) the terms of the numerator of (7.15) are both $O(n^{2d_x+d_y+5/2})$ and

$$n^{1/2-d_y}(\hat{\alpha}-\alpha) \xrightarrow{\mathrm{d}} \frac{\int_0^1 X^2 \mathrm{d}\xi Y(1) - \int_0^1 X \mathrm{d}\xi \int_0^1 X \mathrm{d}Y}{\int_0^1 X^2 \mathrm{d}\xi - \left(\int_0^1 X \mathrm{d}\xi\right)^2}.$$

The intercept is consistently estimated in these cases. In cases (7.10c) and (7.10d), where $d_y \leq -d_x$, the first term in the numerator is $O(n^{2d_x+d_y+5/2})$ but the second term in the numerator is either $O(n^{d_x+5/2})$ in case (7.10d) or $O(n^{d_x+5/2}\log n)$ in case (7.10c), which dominate the first term. However, the error of estimate is at worst $O(n^{-d_x-1/2}\log n)$, which is of small order with $d_x > -\frac{1}{2}$. The intercept is therefore consistently estimated in all cases.

7.3 Implications for Modelling

It is a commonplace in econometrics that in a stationary world an endogenous regressor results in inconsistent regression, whereas in a cointegrating world, endogeneity may result in bias but the regression is still consistent. The interesting question is what the transition between these two worlds looks like, a question that may be answered precisely by considering fractionally integrated series.

Referring to the discussion on page 10, consider an overdifferenced fractional process x with $d_x < -\frac{1}{2}$. If x is defined by (1.1)+(1.2) then

$$S_i = \sum_{j=1}^{i} x_j = \sum_{j=-\infty}^{i} b_j^* u_j$$

where $b_j^* = \sum_{k=\max\{0,1-j\}}^{i-j} b_k$, just as in (2.25). If $b_k \sim d_x k^{d_x - 1}$ (omitting the slowly varying component) then eventually $b_j^* \sim j^{d_S - 1}$, as in (2.29), where $d_S = 1 + d_x < \frac{1}{2}$, so S_i is a stationary long memory process and the limits in (7.9a) and (7.9b) no longer apply. As explained in §7.1 they must be replaced by, respectively, $n^{-1}\sum_{i=1}^{n} S_i \to_{L_2} 0$ and $n^{-1}\sum_{i=1}^{n} S_i^2 \to_{L_2} \sigma_S^2$ where $\sigma_S^2 = \sigma_u^2 \sum_{j=0}^{\infty} b_j^{*2} < \infty$, as in (1.5). If $\sigma_{uw} \neq 0$ then instead of (7.14), the regression error-of-estimate converges in mean square to the limit σ_{Sy}/σ_S^2, where $\sigma_{Sy} = \sigma_{uw}\sum_{j=0}^{\infty} b_j^* c_j < \infty$ as in (4.3), matching (7.4) except with S replacing x. Thus, $\hat{\beta}$ is inconsistent and the solution for the asymptotic bias matches that for the model (7.1). Thinking of the 'model space' as the set of the possible values of d_x that might generate the data, $-\frac{1}{2}$ is the point in model space representing the boundary between regions of stationarity and nonstationarity of the process S. The transition from a consistent albeit biased regression to inconsistency due to an endogenous regressor occurs, as may be expected, at the boundary of the stationarity region.

In the case of a strictly exogenous regressor in (7.7), with $\sigma_{uw} = 0$ so that limit (7.10b) applies, the restriction $d_x + d_y > -\frac{1}{2}$ imposed by Assumption **6.9** is necessary for convergence to an a.s. continuous limit process, as shown in Theorem **6.13**. A natural question to ask is, what actually happens if $d_x + d_y < -\frac{1}{2}$? This has a ready answer from Theorem **4.3**, with S replacing x, noting that $d_x + d_y < -\frac{1}{2}$ is equivalent to $d_S + d_y < \frac{1}{2}$ since $d_S = 1 + d_x$. In (7.10b) the normalizing divisor has exponent smaller than $\frac{1}{2}$ and the sum consequently diverges, albeit having mean of zero. Changing the normalization of the numerator to $n^{-1/2}$ gives the limit for (7.8) as

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathrm{d}} \mathrm{N}\left(0, \frac{V_{Sy}}{\sigma_S^4}\right) \tag{7.16}$$

having the same form as (7.3), except with S replacing x. It is a curious fact that the complementary conditions set by Assumption **6.9** in Theorem **6.13** on the one hand and Theorem **4.3** on the other appear to be motivated in quite different ways in the logic of the respective proofs, when in reality they are two sides of the same coin.

A further question the curious reader might pose is, what happens in the state of the world where these inequalities are changed to the matching equality? Setting $d_x + d_y = -\frac{1}{2}$ (equivalently, $d_S + d_y = \frac{1}{2}$), the sum in (7.10b) is actually being normalized by $n^{-1/2}$ in this case. However, the distribution of the estimator still does not match (7.16) since, according to (4.8) (with d_x standing in for d_S), $V_{Sy} = \infty$. Neither of the two indicated limits applies in this case, but a logarithmic normalization to give a finite Gaussian limit might evidently be constructed.

Another important way in which the fractional cointegration scenario differs from the unit-root case concerns the possibility of mixed-normal inference. Phillips and Hansen ([56]) and Saikkonen ([62]) among other authors have proposed modified least squares estimators for unit root cointegration in which, by devices such as the addition of certain stationary variables to the equation, the residual term y is rendered orthogonal to x. The latter can then be treated as conditionally fixed, giving rise to asymptotically normal t ratios. This strategy cannot work in the case of a fractional residual. Unless $d_y = 0$, the distribution of $\int_0^1 X\mathrm{d}Y$ depends on components Ξ_1 and Ξ_3 and while both are conditionally Gaussian, the conditioning processes are different. Another possible pitfall is the use of prewhitening in the construction of test statistics, noting that an autoregressive filter cannot reduce a fractional process to white noise.

7.4 Cointegration with Drift

The development in §7.2 specified that the regressor be free of drift, as the partial sum of a zero mean process. If this assumption is relaxed, a different limit distribution is obtained with the trend dominating. In conventional unit root analysis, the standard procedure is to partial out any drift by inclusion of a trend dummy in the regression. This still results in a different limit distribution, but one not dependent on the drift parameter and which might be tabulated.

The same approach can be followed in fractional cointegration. If the regressor increment process has the form $x_i + \mu$ where x_i is given as usual by (1.1)+(1.2), the partial sum has the form $S_i + \mu i$ for $i = 1, \ldots, n$. Removing dependence on μ is conveniently performed by demeaning and detrending the variable and running a second stage regression on the resulting residual, which is equivalent to the multiple regression according to the well known Frisch-Waugh theorem. If S_i^* denotes the first-stage residual, the second stage error-of-estimate is just

$$\hat{\beta} - \beta = \frac{\sum_i S_i^* y_i}{\sum_i S_i^{*2}}. \tag{7.17}$$

The preliminary regression of $S_i + \mu i$ onto intercept and trend can be simplified by the approximation

$$\begin{bmatrix} n & \frac{1}{2}n(n+1) \\ \frac{1}{2}n(n+1) & \frac{1}{6}n(n+1)(2n+1) \end{bmatrix}^{-1} \sim \begin{bmatrix} 4n^{-1} & -6n^{-2} \\ -6n^{-2} & 12n^{-3} \end{bmatrix}.$$

An asymptotically equivalent form of the first-stage residual is then

$$\begin{aligned} S_i^* \sim S_i + \mu i &- \left(\frac{4}{n} \sum_k (S_k + \mu k) - \frac{6}{n^2} \sum_k k(S_k + \mu k) \right) \\ &- \left(\frac{12}{n^3} \sum_k k(S_k + \mu k) - \frac{6}{n^2} \sum_k (S_k + \mu k) \right) i \end{aligned}$$

which, it can be verified, does not depend on μ in the limit as $n \to \infty$. Further elementary calculations show that

$$\sum_i S_i^{*2} \sim \sum_i S_i^2 - \frac{1}{n}\Big(\sum_i S_i\Big)^2 - \frac{12}{n^3}\Big(\sum_i\Big(i-\frac{n}{2}\Big)S_i\Big)^2$$

so that Theorem **3.2** and the continuous mapping theorem give

$$\frac{1}{n^{2+2d_x}}\sum_i S_i^{*2} \xrightarrow{\mathrm{d}} \int_0^1 X^2\mathrm{d}\xi - \Big(\int_0^1 X\mathrm{d}\xi\Big)^2 - 12\Big(\int_0^1(\xi-\tfrac{1}{2})X\mathrm{d}\xi\Big)^2. \tag{7.18}$$

Also, subject to the conditions of either (7.10a) or (7.10b),

$$\begin{aligned}\frac{1}{n^{1+d_x+d_y}}\sum_i S_i^* y_i &\sim \frac{1}{n^{1+d_x+d_y}}\sum_i S_i y_i - \frac{1}{n^{3/2+d_x}}\sum_i S_i \frac{1}{n^{1/2+d_y}}\sum_i y_i \\ &\quad - \frac{12}{n^{5/2+d_x}}\sum_i\Big(i-\frac{n}{2}\Big)S_i\frac{1}{n^{3/2+d_y}}\sum_i\Big(i-\frac{n}{2}\Big)y_i \\ &\xrightarrow{\mathrm{d}} \int_0^1 X\mathrm{d}Y - \int_0^1 X\mathrm{d}\xi Y(1) - 12\int_0^1(\xi-\tfrac{1}{2})X\mathrm{d}\xi\Big(\int_0^1\xi\mathrm{d}Y - \tfrac{1}{2}Y(1)\Big).\end{aligned} \tag{7.19}$$

So if either $d_x + d_y > 0$, or $\sigma_{uw} = 0$ and $d_x + d_y > -\frac{1}{2}$, $n^{1+d_x-d_y}(\hat{\beta}-\beta)$ in (7.17) converges in distribution to the ratio of the limits in (7.19) and (7.18).

Given the existence of the fBM X, this development is identical to the usual unit root analysis, as treated in [13] among many other such references, in all respects except one. The exception is in the final term of (7.19), since the limit of the random sequence $n^{-3/2-d_y}\sum_i iy_i$ has not yet been examined. For the limit distribution in (7.19) to be well defined, this must be shown to have finite variance in the limit.

This development can conveniently follow the approach of §2.3 and §2.4. By analogy with (2.23), define an array $\{g_{ni}\}$ by

$$\sum_{i=1}^n iy_i = \sum_{i=1}^n i\sum_{l=0}^\infty c_l w_{i-l} = \sum_{i=-\infty}^n g_{ni}w_i \tag{7.20}$$

where if $c_l \sim d_y l^{d-1}$ (ignoring any scale factors) it can be verified that

$$g_{ni} = \sum_{l=\max\{0,1-i\}}^{n-i}(l+i)c_l \sim d_y\sum_{l=\max\{0,1-i\}}^{n-i}(l^d + il^{d-1}).$$

To calculate the mean square of (7.20), split the sum into the positive and non-positive indices. For $0 < x \le 1$,

$$g_{n[nx]} \sim d_y\int_0^{n-[nx]}\tau^{d_y}\mathrm{d}\tau + [nx]d_y\int_0^{n-[nx]}\tau^{d_y-1}\mathrm{d}\tau$$

$$\sim n^{d_y+1}\frac{d_y+x}{d_y+1}(1-x)^{d_y} \tag{7.21}$$

whereas for $-\infty < x \leq 0$,

$$\begin{aligned} g_{n[nx]} &\sim d_y\int_{-[nx]}^{n-[nx]}\tau^{d_y}\mathrm{d}\tau + [nx]d_y\int_{-[nx]}^{n-[nx]}\tau^{d_y-1}\mathrm{d}\tau \\ &\sim n^{d_y+1}\biggl(\frac{d_y}{d_y+1}\bigl((1-x)^{d_y+1}-(-x)^{d_y+1}\bigr) \\ &\qquad\qquad -(-x)\bigl((1-x)^{d_y}-(-x)^{d_y}\bigr)\biggr). \end{aligned} \tag{7.22}$$

Essentially, the requirement is now to show that the functions of x in (7.21) and (7.22) are both square-integrable. First, a straightforward calculation gives

$$\int_0^1\Bigl(\frac{d_y+x}{d_y+1}\Bigr)^2(1-x)^{2d_y}\mathrm{d}x = \frac{2d_y^2+d_y+1}{(d_y+1)(2d_y+1)(2d_y+3)}. \tag{7.23}$$

By contrast, integrating the square of (7.22) over $(-\infty,0]$ in closed form is not a trivial exercise. However, substituting the large-z approximation

$$(1+z)^a - z^a = az^{a-1} + O(z^{a-2})$$

for each of the terms shows that as $x \to -\infty$,

$$\begin{aligned} \frac{g_{n[nx]}}{n^{d_y+1}} &\sim d_y(-x)^{d_y} - (-x)d_y(-x)^{d_y-1} + O((-x)^{d_y-1}) \\ &= O((-x)^{d_y-1}). \end{aligned} \tag{7.24}$$

When n is large enough, the normalized $\{g_{ni}\}$ sequence is thus shown to be square-summable and similarly to Corollary **2.7**,

$$\lim_{n\to\infty}\mathrm{E}\biggl(\frac{1}{n^{3/2+dy}}\sum_{i=1}^n iy_i\biggr)^2 = \lim_{n\to\infty}\frac{\sigma_w^2}{n^{3+2dy}}\sum_{i=-\infty}^n g_{ni}^2 < \infty.$$

Under Assumption **4.1**, the distribution of $\int_0^1 \xi \mathrm{d}Y$ could be shown to be Gaussian by application of the methods of Chapters 2 and 3 to the sequence defined by (7.20), although this detail will not be pursued here.

Chapter 8

Autocorrelated Shocks

The asymptotic analysis developed in Chapters 4 through 7 has been based on Assumption **4.1**, specifying serial independence of the shock variables u_i and w_i. Because of minimal restrictions on the form of the linear coefficients at low orders of lag, local dependence that can be removed by linear filtering is accounted for (see Theorem **1.4**), so that the assumption of zero autocorrelations for the shock process is not unreasonably restrictive. Nonetheless, Assumption **4.1** remains a special case and the introductory discussion in §3.5 (see page 44) suggests one reason at least why it might be desirable to relax it.

All but one of the limit results appearing in Chapters 4–6 under Assumption **4.1** also hold under Assumption **8.1** below. (Theorem **6.16** is the exception.) The only change arising otherwise is the redefinition of the variance parameter, with the long-run variance ω_u^2 replacing σ_u^2. Indeed, there exists the possibility of nonlinear local dependence such as conditional heteroscedasticity, permitted under Assumption **8.1**, where there is no autocorrelation and $\omega_u^2 = \sigma_u^2$.

That said, it must be born in mind that rates of convergence to the limit are generally slower than when the shocks are independent. Examination of the proofs of results such as Theorem **2.10** and Theorem **3.10** in SLT gives an idea of how nominal and effective sample sizes may differ. Another major difference, following the precedent set by Theorem **2.10**, is that to reach these results, tedious algebraic manipulation is sometimes unavoidable.

8.1 Correlation Analysis

Extending Assumption **1.2** into the multivariate context must begin by defining some new symbols. Thus,

$$\gamma_{uw}(j) = \mathrm{E}(u_0 w_j) \text{ with } \boldsymbol{\omega}_{uw} = \textstyle\sum_{j=-\infty}^{\infty} \gamma_{uw}(j)$$

and

$$\mu_{uw}^4(j,k,l) = \mathrm{E}(u_0 w_k u_j w_{j+l}) \text{ with } \varpi_{uw}^4 = \textstyle\sum_{j,k,l=-\infty}^{\infty} \mu_{uw}^4(j,k,l).$$

Under stationarity these moments are invariant to the time index and in particular, $\gamma_{uw}(0) = \sigma_{uw}$ and $\mu_{uw}^4(0,0,0) = \mu_{uw}^4$, where the latter symbols are defined in

Assumption **4.1**(b).

8.1 Assumption

(a) Assumption **4.1**(a) holds and Assumption **1.2** holds for u_i and also for w_i with γ_w and ω_w^2 defined as for γ_u and ω_u^2 in Assumption **1.2**(a) and (1.4).

(b) $\gamma_{uw}(j) = O(|j|^{-1-\delta})$ for $\delta > 0$, and $|\omega_{uw}| < \infty$.

(c) for $\delta > 0$,

(i) $\mu_{uw}^4(j,k,l) = O(|k|^{-1-\delta}|l|^{-1-\delta})$

(ii) $\mu_{uw}^4(j,k,l) - \gamma_{uw}(k)\gamma_{uw}(l) = O(|j|^{-1-\delta})$

(iii) $\mu_{uw}^4(j,k,k+l) - \gamma_u(j)\gamma_w(j+l) = O(|k|^{-1-\delta})$

and $|\varpi_{uw}^4| < \infty$. □

In the context of this assumption and throughout this chapter, $0^{-1-\delta}$ represents 1 in calculations and similarly for zero raised to any negative power, where the counting index refers to a lag coefficient or order of autocovariance.

Case (i) of Assumption **8.1**(c) can be best appreciated by observing that under stationarity the expectations $\mathrm{E}(u_0 w_k u_j w_{j+l})$ are assumed to converge at the indicated rate to $\mathrm{E}(u_0 u_j w_{j+l})\mathrm{E}(w_0)$ as $|k| \to \infty$ and to $\mathrm{E}(u_0 u_j w_k)\mathrm{E}(w_0)$ as $|l| \to \infty$, the limits being zero in each case. According to (ii), the limit of $\mathrm{E}(u_0 w_k u_j w_{j+l})$ as $|j| \to \infty$ has the form $\mathrm{E}(u_0 w_k)\mathrm{E}(u_0 w_l)$ where the latter factors converge with k and l according to Assumption **8.1**(b). In case (iii), the limit of $\mathrm{E}(u_0 w_k u_j w_{j+k+l})$ as $|k| \to \infty$ is $\mathrm{E}(u_0 u_j)\mathrm{E}(w_0 w_{j+l})$. It is assumed implicitly in (i) that the divergences of the indices are not coordinated, by setting $k = l$ for example. This case is covered by (iii) with a given value of l.

The divergence rates in Assumption **8.1** are generally required to hold for any $\delta > 0$, in which case they are cited without comment. If δ must exceed some positive bound, to be specified, it is possible for δ to differ for u_i and w_i in Assumption **1.2**. In this case note that if either $\gamma_u(j) = 0$ or $\gamma_w(j) = 0$ for $j \neq 0$ then $\gamma_{uw}(j) = 0$ for $j \neq 0$ also, with a similar consideration for the fourth moments.

Under these assumptions the first modification called for is to (4.3), which now becomes

$$\mathrm{E}(x_i y_i) = \sigma_{xy} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} b_j c_k \gamma_{uw}(j-k). \tag{8.1}$$

To show this sum finite is no longer a trivial exercise, but summability of the autocovariances is sufficient. The following pair of mechanical lemmas supply the result, and have further applications. The proof uses the same basic technique that was applied in Theorem **1.4**. When applying the first of these lemmas, keep in mind the useful identity

$$\sum_{k=1}^{j-1}(j-k)^{d-1}k^{-1-\delta} = \sum_{m=1}^{j-1} m^{d-1}(j-m)^{-1-\delta}$$

where $m = j - k$. Hence, it doesn't matter which exponent plays which role in the sum.

8.2 Lemma If $d < 1$ and $\delta > 0$ then

$$\sum_{k=1}^{j-1} (j-k)^{d-1} k^{-1-\delta} \backsimeq j^{d-1}.$$

Proof Choose η from the interval $(1/(1+\delta), 1)$ and so write

$$\sum_{k=1}^{j-1} (j-k)^{d-1} k^{-1-\delta} = j^{d-1} \Big(\frac{j - j^\eta}{j}\Big)^{d-1} (A(j) + B(j)) \tag{8.2}$$

where $(j - j^\eta)/j \to 1$ as $j \to \infty$ and

$$A(j) + B(j) = \Bigg(\sum_{k=1}^{[j^\eta]-1} + \sum_{k=[j^\eta]}^{j-1} \Bigg) \Big(\frac{j-k}{j - j^\eta} \Big)^{d-1} k^{-1-\delta}.$$

Since $(j-k)/(j-j^\eta) > 1$ in $A(j)$ and $d - 1 < 0$ and also $k^{-1-\delta}$ is summable, $A(\infty) < \infty$. Setting $m = j - k$ gives

$$\begin{aligned}
B(j) &= \sum_{m=1}^{j-[j^\eta]} \Big(\frac{m}{j - j^\eta} \Big)^{d-1} (j-m)^{-1-\delta} \\
&= (j - j^\eta)^{1-d} j^{-\eta(1+\delta)} \sum_{m=1}^{j-[j^\eta]} m^{d-1} \Big(\frac{j-m}{j^\eta} \Big)^{-1-\delta} \\
&\leq (j - j^\eta)^{1-d} j^{-\eta(1+\delta)} \sum_{m=1}^{j-[j^\eta]} m^{d-1} \\
&\ll (j - j^\eta) j^{-\eta(1+\delta)} \leq j^{1-\eta(1+\delta)} \to 0
\end{aligned} \tag{8.3}$$

as $j \to \infty$. The first inequality in (8.3) holds since $j - m \geq j^\eta$ so that the second factor in the sum over m is smaller than 1, while the sum itself is of $O((j - j^\eta)^d)$ by integral approximation. The final bound is of small order by choice of η. ∎

8.3 Lemma If $|d_x| < \frac{1}{2}$, $|d_y| < \frac{1}{2}$ and $\delta > 0$ then

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} j^{d_x - 1} k^{d_y - 1} |j-k|^{-1-\delta} < \infty. \tag{8.4}$$

Proof Write the expression in (8.4) as $A+B$ where A contains those terms where $k < j$ and B the terms where $k \geq j$. Both A and B are shown to be finite. First, by Lemma **8.2** with $d = d_y$,

$$A = \sum_{j=1}^{\infty} j^{d_x - 1} \sum_{k=1}^{j-1} k^{d_y - 1} (j-k)^{-1-\delta} \backsimeq \sum_{j=1}^{\infty} j^{d_x + d_y - 2} < \infty.$$

Second, put $m = k - j$ to obtain

$$B = \sum_{j=1}^{\infty} j^{d_x - 1} \sum_{k=j}^{\infty} k^{d_y - 1}(k-j)^{-1-\delta} = \sum_{j=1}^{\infty} j^{d_x + d_y - 2} \sum_{m=0}^{\infty} \Big(\frac{j+m}{j}\Big)^{d_y - 1} m^{-1-\delta}$$
$$\ll \sum_{j=1}^{\infty} j^{d_x + d_y - 2} < \infty$$

where the inequality holds since $(j+m)/j \geq 1$ and $d_y - 1 < 0$, while $m^{-1-\delta}$ is summable. ∎

8.4 Theorem Under Assumption **8.1**, $\sigma_{xy} < \infty$.

Proof Direct from Lemma **8.3** given (8.1) and Assumption **8.1**, since the summability criteria are invariant to slowly varying components of b_j and c_k. ∎

These preliminaries lead to the next task which is to prove the generalized version of Theorem **4.2**.

8.5 Theorem With x_i and y_i defined by (4.1) where Assumption **8.1** holds,

$$\frac{1}{n}\sum_{i=1}^{n} x_i y_i \xrightarrow{L_2} \sigma_{xy}.$$

Proof The expression that must be shown to vanish is (4.4). For brevity define $V_n = \mathrm{E}\big(n^{-1}\sum_{i=1}^{n} x_i y_i - \sigma_{xy}\big)^2$. Under stationarity of u_i and w_i,

$$\sum_{i=1}^{n}\sum_{k=1}^{n} \mathrm{E}(u_{i-j}w_{i-l}u_{k-m}w_{k-p}) = n\sum_{k=1}^{n} \mathrm{E}(u_0 w_{j-l} u_{k-m} w_{k-p})$$

and (4.4) can be written as

$$V_n = \frac{1}{n^2}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}\sum_{p=0}^{\infty} b_j b_m c_l c_p$$
$$\times\, n\sum_{k=1}^{n}\big(\mu^4_{uw}(k-m, j-l, m-p) - \gamma_{uw}(j-l)\gamma_{uw}(m-p)\big).$$

Therefore, ignoring slowly varying components and applying Assumption **8.1**(c),

$$V_n \ll \frac{1}{n}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{l=0}^{\infty}\sum_{p=0}^{\infty} j^{d_x-1} m^{d_x-1} l^{d_y-1} p^{d_y-1}$$
$$\times \sum_{k=1}^{n} |k-m|^{-1-\delta}|j-l|^{-1-\delta}|m-p|^{-1-\delta}. \qquad (8.5)$$

Since $\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} j^{d_x-1}l^{d_y-1}|j-l|^{-1-\delta} < \infty$ by Lemma **8.3**, these terms do not affect the order of magnitude and a statement equivalent to (8.5) is

$$V_n \ll \frac{1}{n}\sum_{k=1}^{n}\sum_{m=0}^{\infty} m^{d_x-1}|k-m|^{-1-\delta}\sum_{p=0}^{\infty} p^{d_y-1}|m-p|^{-1-\delta}. \tag{8.6}$$

Consider the sum over p in (8.6). Decompose this sum into the terms with $p < m$ and $p \geq m$ and in the second case set $q = p - m$ and rearrange, to get

$$\begin{aligned}\sum_{p=0}^{\infty} p^{d_y-1}|m-p|^{-1-\delta} &= \sum_{p=0}^{m-1} p^{d_y-1}(m-p)^{-1-\delta} + m^{d_y-1}\sum_{q=0}^{\infty}\Big(\frac{q+m}{m}\Big)^{d_y-1} q^{-1-\delta} \\ &\ll m^{d_y-1}. \end{aligned} \tag{8.7}$$

This bound holds because the first term in (8.7) is bounded by Lemma **8.2** with $d = d_y$, while the sum in the second term is finite by summability of $q^{-1-\delta}$ since the other factor never exceeds 1. Substitute the bound in (8.7) back into (8.6) and again split the sum, this time into the terms with $m < k$ and with $m \geq k$. Apply Lemma **8.2** to the first of these components and since $k^{d_x+d_y-2}$ is summable, the result is

$$\begin{aligned} V_n &\ll \frac{1}{n}\sum_{k=1}^{n}\Bigg(\sum_{m=0}^{k-1} m^{d_x+d_y-2}(k-m)^{-1-\delta} \\ &\qquad + k^{d_x+d_y-2}\sum_{m=k}^{\infty}\Big(\frac{m}{k}\Big)^{d_x+d_y-2}(m-k)^{-1-\delta}\Bigg) \\ &\ll \frac{1}{n}\sum_{k=1}^{n} k^{d_x+d_y-2} = O(n^{-1}). \quad \blacksquare \end{aligned}$$

There is an instructive comparison here with the proof of Theorem **4.2**, which explicitly counted the nonzero terms of the sum V_n. The present argument works by showing that the non-negligible terms of the sum are sparse enough under Assumption **8.1** to permit summability, without having to itemize them.

The next result to be generalized is Theorem **4.3**. The conditions here do not impose mutual independence of the shocks u and w, only that they be uncorrelated with each other at all orders of lag.

8.6 Theorem With x_i and y_i defined by (4.1), under Assumption **8.1** and if $\gamma_{uw}(j) = 0$ for all j and $d_x + d_y < \frac{1}{2}$,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i y_i \xrightarrow{\mathrm{d}} \mathrm{N}(0, V_{xy})$$

where $V_{xy} < \infty$.

Proof As in Theorem **4.3**, the object is to show that $V_{xy} = E(\zeta^2) < \infty$ where $\zeta = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} b_k c_j Z(k,j)$ and $Z(k,j)$ is defined in (4.7). The proof is modified as follows. Since by assumption $u_{i-k}w_{i-j}$ has mean zero for all k and j, the random variables $Z(k,j)$ are Gaussian under Assumption **8.1** by a CLT for dependent data such as Theorem **3.10**. Thus, it can be verified that under Assumption **8.1**(c), Assumption **1.2** holds for random sequences of the form $\{u_{i-k}w_{i-j}\}_{i=-\infty}^{\infty}$. Whereas under Assumption **4.1** these random variables have the property

$$\mathrm{E}\big(Z(k,j)Z(k+p,j+p)\big) = \sigma_u^2\sigma_w^2$$

for $p \in \mathbb{N}$ and are independent otherwise, under Assumption **8.1**

$$\mathrm{E}\big(Z(k,j)Z(k+p+t,j+p+s)\big) = \gamma_u(t)\gamma_w(s)$$

for $p,t,s \in \mathbb{N}$, noting that $\sigma_u^2 = \gamma_u(0)$ and $\sigma_w^2 = \gamma_w(0)$. In place of (4.8), define for each t and s,

$$V_{xy}(t,s) = \gamma_u(t)\gamma_w(s)\Bigg(\sum_{k=0}^{\infty} b_{k+t}^2 \sum_{j=0}^{\infty} c_{j+s}^2 + 2\sum_{p=1}^{\infty}\Bigg(\sum_{k=0}^{\infty} b_k b_{k+p+t}\sum_{j=0}^{\infty} c_j c_{j+p+s}\Bigg)\Bigg).$$

Similarly to the proof of Theorem **4.3**, each of these sums is finite by assumption. Also

$$\sum\nolimits_{t,s}|\gamma_u(t)\gamma_w(s)| = \sum\nolimits_t |\gamma_u(t)|\sum\nolimits_s|\gamma_w(s)| < \infty$$

by Assumption **8.1**(a) and hence $V_{xy} = \mathrm{E}(\zeta^2) = \sum_{t,s} V_{xy}(t,s) < \infty$. ∎

8.2 The Covariance Decomposition

The next step requiring modification is the decomposition of G_n in (4.12)–(4.14). Under Assumption **8.1** it is no longer the case that $\mathrm{E}(G_{1n}) = 0$ and $\mathrm{E}(G_{3n}) = 0$. Instead, let a sequence $\{L_n\}$ be chosen such that $L_n \to \infty$ but $L_n/n \to 0$ as $n \to \infty$. Then in place of (4.12)–(4.14) substitute the definitions

$$G_{1n} = \frac{1}{nK(n)}\sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{j=0}^{\infty} b_j u_{k-j}\sum_{l=0}^{i+j-k-L_n} c_l w_{i+1-l} \tag{8.8}$$

$$G_{2n} = \frac{1}{nK(n)}\sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{j=0}^{\infty} b_j u_{k-j}\sum_{v=-L_n}^{L_n} c_{i+j-k+1+v} w_{k-j+v} \tag{8.9}$$

$$G_{3n} = \frac{1}{nK(n)}\sum_{i=1}^{n-1}\sum_{k=1}^{i}\sum_{j=0}^{\infty} b_j u_{k-j}\sum_{l=i+j-k+2+L_n}^{\infty} c_l w_{i+1-l}. \tag{8.10}$$

$G_n = G_{1n} + G_{2n} + G_{3n}$ as before, but $2L_n$ terms have been moved from G_{1n} and G_{3n} into G_{2n}. (An empty sum equals zero, note.) In view of Assumption **8.1**(c), under these definitions $\mathrm{E}(G_{1n}) = O(L_n^{-\delta})$ and $\mathrm{E}(G_{3n}) = O(L_n^{-\delta})$. With $L_n \sim n^{\alpha}$

for $\alpha < 1$, the convergence rate of $n^{-\alpha\delta}$ shows how the best choice of L_n might relate to the degree of dependence.

Expression (4.16) now has to be replaced by

$$\mathrm{E}(G_{2n}) = \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=0}^{\infty} a_{n,i-k}(i/n, 0) c_{k+1} F_n(k) + o(1) \tag{8.11}$$

where

$$F_n(k) = \sum_{l=-L_n}^{L_n} \frac{c_{k+1+l}}{c_{k+1}} \gamma_{uw}(l) \backsimeq \sum_{l=-L_n}^{L_n} \left(1 + \frac{l}{k}\right)^{d_y - 1} \gamma_{uw}(l).$$

For finite values of l, $(1 + l/k)^{d_y-1} = 1 + O(k^{-1})$. Assumption **8.1**(b) implies that only finite values of l contribute significantly to $F_n(k)$ and hence that $F_n(k) \to \omega_{uw}$ as $k \to \infty$ and $n \to \infty$. The counterpart expression to (4.17) (the terms of (8.11) with $k < i$) therefore takes the form

$$\begin{aligned} \frac{1}{nK(n)} \sum_{i=1}^{n-1} \sum_{k=0}^{i-1} a_{n,i-k}(i/n, 0) c_{k+1} F_n(k) &\sim \frac{d_y}{n^2} \sum_{i=1}^{n-1} \sum_{k=1}^{i} \left(\frac{k}{n}\right)^{d_x + d_y - 1} F_n(k) \\ &\to \omega_{uw} d_y \int_0^1 \int_0^\tau \zeta^{d_x + d_y - 1} \mathrm{d}\zeta \mathrm{d}\tau \\ &= \frac{\omega_{uw} d_y}{(d_x + d_y)(1 + d_x + d_y)} \end{aligned} \tag{8.12}$$

as $n \to \infty$. The convergence of F_n can be asserted here, since the contribution of finite values of k to the limiting sum is of small order in n. This matches the limit in (4.17) except that ω_{uw} replaces σ_{uw}. Exactly the same modification applies to (4.18).

These considerations suffice to verify the following generalization of Theorem **4.4**.

8.7 Theorem Under Assumption **8.1** and with G_{2n} defined by (8.9), $\mathrm{E}(G_{2n}) \to \omega_{uw}\lambda_{xy}$ as $n \to \infty$ where λ_{xy} is defined in (4.15). □

The next result, explicitly invoking the autocorrelation structure of Assumption **8.1**(c), is the generalized form of Theorem **4.8**.

8.8 Theorem Under Assumption **8.1** and with G_{2n} defined by (8.9), $G_{2n} - \mathrm{E}(G_{2n}) \to_{L_2} 0$.

Proof The argument in Theorem **4.8** is modified as follows. The term (4.41) needs to be replaced by

$$P_{ik} \equiv \sum_{j=0}^{\infty} b_j \sum_{v=-L_n}^{L_n} c_{i+1-k+j+v}(u_{k-j} w_{k-j+v} - \gamma_{uw}(v)).$$

The inequality (4.43) whose minorant appears in inequality (4.42) is therefore replaced by

$$|\mathrm{E}(P_{ik}P_{i-m,k-p})| \leq \sum_{j=0}^{\infty}\sum_{l=0}^{\infty}\Bigg|b_jb_l \sum_{v=-L_n}^{L_n}\sum_{v'=-L_n}^{L_n} c_{i+1-k+j+v}c_{i-m+1-k+p+l+v'} \times \big(\mathrm{E}(u_{k-j}w_{k-j+v}u_{k-p-l}w_{k-p-l+v'}) - \gamma_{uw}(v)\gamma_{uw}(v')\big)\Bigg|. \quad (8.13)$$

In place of (4.44), the expression in question becomes

$$|\mathrm{E}(P_{ik}P_{i-m,k-p})| \leq \sum_{j=0}^{\infty}|b_jb_{j-p}c_{i+1-k+j}c_{i-m+1-k+j}H_n(j,i,k,m,p)| \quad (8.14)$$

where

$$H_n(j,i,k,m,p) = \sum_{q=p-j}^{p+j}\frac{b_{j-p+q}}{b_{j-p}}\sum_{v=-L_n}^{L_n}\sum_{v'=-L_n}^{L_n}\frac{c_{i+1-k+j+q+v}}{c_{i+1-k+j}}\frac{c_{i-m+1-k+j+q+v'}}{c_{i-m+1-k+j}} \times \big(\mu^4_{uw}(q,v,v') - \gamma_{uw}(v)\gamma_{uw}(v')\big). \quad (8.15)$$

In (8.15), observe that $q = p + l - j$ so that the double sum over j and l in (8.13) is constructed by diagonals.

The approach of Theorem **4.8** carries through unchanged provided the function $H_n(j,i,k,m,p)$ is bounded in the limit for all values of the arguments, noting that these are liable to diverge as $n \to \infty$. According to Assumptions **8.1**(b) and (c),

$$\sum_{q=-\infty}^{\infty}\sum_{v=-L_n}^{L_n}\sum_{v'=-L_n}^{L_n}\big(\mu^4_{uw}(q,v,v') - \gamma_{uw}(v)\gamma_{uw}(v')\big) \to \varpi^4_{uw} - \omega^2_{uw} < \infty$$

as $n \to \infty$. This summability implies that only finite values of v, v', and q appear in non-negligible terms of the sum in (8.15). Attention therefore focuses on the various ratios of coefficients. For finite values of q, note that

$$\frac{b_{j-p+q}}{b_{j-p}} \backsimeq \left(1 + \frac{q}{j-p}\right)^{d_x-1} = 1 + O((j-p)^{-1}).$$

Similarly, for arbitrary argument r and finite values of v and q,

$$\frac{c_{r+q+v}}{c_r} \backsimeq \left(1 + \frac{v+q}{r}\right)^{d_y-1} = 1 + O(r^{-1}).$$

In all but a finite number of the terms of the majorant of (4.42), the indices $r = i+1-k+j$ in (8.15) are diverging as $n \to \infty$, and since the sum is divergent, either $r = O(n)$ or r relates to a collection of terms that are of small order in the normalized sum. The same is true of $r' = i - m + 1 - k + j$. These facts establish that $H_n(j,i,k,m,p) = O(1)$ as each of its arguments diverges. If the expression in (8.14) is decomposed into either B_{11} and B_{12} as in (4.46) or into B_{21} and B_{22} as in (4.48), with $H_n(j,i,k,m,p)$ replacing $\mu^4_{uw} - \sigma^2_{uw}$ in each case, the bounds established in (4.51), (4.53), (4.55), and (4.57) all continue to apply. ∎

8.3 Stochastic Integrals

Having redefined G_{2n} in (8.9), it is next necessary to consider the effects of the corresponding changes in G_{1n} and G_{3n}, following the analysis of Chapter 5. In the first case, considering the change from (4.12) to (8.8), refer to the reorganization of the sum detailed in pages 74–75. Since k has been replaced by $k + L_n$ in (8.8), it can be verified that the upper limit on p in the second member of (5.3) has to change from m to $m - L_n$. Hence, q_{nm} in (5.4) becomes $q_{n,m-L_n}$ and so a_{nmp} becomes $a_{n,m-L_n,p}$. The lag on u_p in moving average (5.3) is increased from 1 to $L_n + 1$ so that any correlation between the variables is rendered asymptotically negligible, according to Assumption **8.1**(b).

Consider how this effects the result of Lemma **5.2**. Since t is defined by $m = [nt]$ in (5.13) and $([nt]+L_n)/n = t+o(1)$, applying Theorem **5.1** shows that the result of Lemma **5.2** is unchanged. The manipulations involving $A(t, s)$, and also its absolute bound $\bar{A}(t, s)$ in Lemma **6.1**, all go through unchanged by the redefinition of G_{1n}. The one notable effect of the change is to resolve the dilemma raised in the remark on page 79, of the singularity in the integral $A(t, s)$ at the point $s = t$. Now, this point has been moved from G_{1n} to G_{2n} and the integral can be unambiguously defined subject to the strict inequality $s < t$.

Exactly the same considerations apply to the redefined G_{3n}. In (5.6), h_{np} becomes $h_{n,p-L_n}$, and hence e_{npm} becomes $e_{n,p-L_n,m}$ so that the lag on w_m relative to u_p in the moving average is increased from 1 to $1+L_n$. The asymptotic formulae $E(s, t)$ and $\bar{E}(s, t)$ nonetheless remain similarly unchanged, apart from imposition of the strict inequality $t < s$.

8.4 Weak Convergence

In Chapter 6, there are changes to be made to Theorem **6.3** and **6.4**, but here the amendments are straightforward. Thus,

8.9 Theorem Under Assumption **8.1**,

(i) $\lim_{n\to\infty} n^{-1}\mathrm{E}\left(\sum_{m=-nN}^{n-1}(q_{n,m-L_n} - q^N_{n,m-L_n})w_{m+1}\right)^2 = O(N^{2d_x-1})$

(ii) $\lim_{n\to\infty} n^{-1}\mathrm{E}\left(\sum_{m=-\infty}^{-nN-1} q_{n,m-L_n}w_{m+1}\right)^2 = O(N^{2(d_x+d_y-1)})$. $\square$

The changes to the proof of **6.3** do not need to be set out in great detail. In (6.11), it is a matter of changing $q_{nm} - q^N_{nm}$ to $q_{n,m-L_n} - q^N_{n,m-L_n}$ and accordingly, in the second member the upper bound of the sum over p is changed to $-nN - L_n - 1$. In the third member, $\sigma^2_u\sigma^2_w$ changes to $\omega^2_u\omega^2_w$ with the addition of a residual term $O(L_n^{-2\delta})$, with the corresponding change to the final asymptotic bound. The modifications to part (ii) of the proof are similar. In the same way, Theorem **6.4** is replaced by

8.10 Theorem Under Assumption **8.1**,

(i) $\lim_{n\to\infty} n^{-1}K(n)^{-2}\mathrm{E}\big(\sum_{p=-nN}^{n-1}(h_{n,p-L_n} - h^N_{n,p-L_n})u_p\big)^2 = O(N^{2d_y-1})$

(ii) $\lim_{n\to\infty} n^{-1}K(n)^{-2}\mathrm{E}\big(\sum_{p=-\infty}^{-nN-1} h_{n,p-L_n}u_p\big)^2 = O(N^{2(d_x+d_y-1)})$. □

Next, consider the weak convergence results. Under Assumption **1.2**, the functional central limit theorem of Chapter 3 was proved with the approach of Theorem **3.11** replacing that of Theorem **3.2**. To reproduce the results of Chapter 6 under Assumption **8.1** a parallel approach is adopted. The essential change is to apply the method of Lemma **3.14** to establish the Gaussianity, in place of Lemma **3.4**.

8.11 Lemma The conclusion of Lemma **6.12** holds under Assumptions **8.1** and **3.9**.

Proof Similarly to Theorem **3.14**, Gaussianity is to be shown for the components $q_{1n}(t)$ and $q^N_{2n}(t)$ in (6.29) by testing the conditions of Theorem **3.10**. In this case, the role of the scale constants c_{nm} is played by the constants c_{np}, defined in (6.32) for the case of q_{1n} and in (6.40) for the case of q^N_{2n}. The main task is to verify the conditions of Assumption **3.9**. The same considerations arise as in the discussion of page 48, except that the dependence of the conditions on the sign of d in the univariate analysis now relate to d_x. Overlooking the non-trending factors attached to (6.32) and (6.40), the arguments in the two cases align closely. Thus, if $d_x > 0$, since $\max_{1\le p\le[nt]} c_{np} = o(n^{-1/2})$ in (6.32) it is easily verified that conditions (3.44) and (3.45) are satisfied for any $\alpha > 0$. If $d_x < 0$, $\max_{1\le p\le[nt]} c_{np} = o(n^{-1/2-d_x})$ in (6.32), but setting $\alpha > -2d_x$ in Assumption **3.9** meets the requirement. The form of (6.40) is only an order-of-magnitude approximation to that of (3.53), but with M^k_{nj} defined similarly to (3.54) but with c_{np} in (6.40) replacing c_{ni} in that formula,

$$M^k_{nj} \backsimeq \frac{((j-1)B_n + nk)^{d_x-1}}{n^{1/2+d_x}}.$$

With this understanding, in the cases $k > 0$,

$$\max_{1\le j\le r_n} M^k_{nj} = M^k_{n0} = O(n^{-3/2}) = o(B_n^{-1/2})$$

where the last equality holds for any $B_n = o(n)$. Also, in view of the square-summability which holds for any $d_x < \frac{1}{2}$,

$$\sum_{j=1}^{r_n}(M^k_{nj})^2 \ll \frac{B_n^{2d_x-2}}{n^{2d_x+1}} = O(B_n^{-1}).$$

For the case $k = 0$, $\max_{1\le j\le r_n} M^0_{nj} = O(n^{-d_x-1/2})$ with $d_x < 0$, so here too setting $\alpha > -2d_x$ is required. With these considerations, Gaussianity is established as before for $q_{1n}(t)$ and the N components of $q^N_{2n}(t)$. To extend this result to $q^N_n(t)$ requires a proof of independent increments, but this goes through by the argument of Theorem **3.13**, without any change except the substitution of $z_{ni} = a_{nmi}u_i/(\sqrt{n}K(n))$ for $z_{ni} = a_{ni}(s,t)u_i/\kappa(n)$. With these changes the order of magnitude in (3.48) continues to apply. ∎

8.12 Lemma The conclusion of Lemma **6.14** also holds under Assumptions **8.1**, **6.9**(a), and **3.15**.

Proof The only changes needing to be made to the proof of Lemma **6.14** are the replacement of Theorem **A.5** by Theorem **3.17** to show uniform integrability of maximal partial sums and the implicit adjustment to the proof of Theorem **6.13** via the citation of Assumption **3.15**. ∎

8.13 Theorem The conclusions of Theorems **6.11** and **6.15** hold under Assumptions **8.1** and **6.9**.

Proof The proofs of Theorem **6.11** and the parallel Theorem **6.15** can be reiterated with the changes that Lemmas **8.11** and **8.12** are invoked in place of Lemmas **6.12** and **6.14**. ∎

The final consideration is Theorems **6.16** and **6.17**, establishing convergence to stochastic integrals. Only one step in the proof of this result depended directly on the independence of the shocks. This was the L_2-approximation of G_{1n}^N by G_{1n}^{N*} in (6.66), for which the sequence k_n was required only to diverge more slowly than n. Now k_n might require further restriction, with Assumption **8.1** needing to be supplemented by a restriction on δ beyond positivity.

8.14 Theorem Under Assumptions **6.9** and **8.1**, $G_{1n}^N \to_{\mathrm{d}} \Xi_1^N$ if in addition, setting $L_n \sim n^\alpha$ and $k_n \sim n^\mu$, there exist $\mu < 1$ and $\alpha < 1$ such that

$$\max\left\{\frac{1-\mu(d_x+\frac{1}{2})}{\alpha} - 1, \frac{1-2\mu(d_x+\frac{1}{2})}{1-\mu}\right\} < \delta \tag{8.16}$$

where δ is defined by Assumption **8.1**(b) and (c).

Proof Under functional form (8.8) the generic form of the covariance segments, replacing (6.61), is

$$G_{1n}^N = \frac{1}{nK(n)} \sum_{m=1-nN+L_n}^{n-1} \sum_{p=1-nN}^{m-L_n} a_{nmp} u_p w_{m+1}. \tag{8.17}$$

The decomposition in (6.63), after modification to ensure that the increments are independent in the limit, takes the form

$$G_{1n}^{N*} = \frac{1}{nK(n)} \sum_{j=1}^{k_n} \left(\sum_{p=n_0}^{n_{j-1}} a_{nn_{j-1}p} u_p \sum_{m=n_{j-1}+1+L_n}^{\min\{n,n_j+L_n\}} w_m \right).$$

The modified form of (6.65) is then

$$G_{1n}^N - G_{1n}^{N*} = \frac{1}{nK(n)} \sum_{j=1}^{k_n} \sum_{m=n_{j-1}+L_n+1}^{\min\{n,n_j+L_n\}-1} \left(\sum_{p=n_{j-1}+1}^{m-L_n} a_{nmp} u_p \right.$$

$$+ \sum_{p=n_0}^{n_{j-1}} (a_{nmp} - a_{nn_{j-1}p}) u_p \Bigg) w_{m+1}. \tag{8.18}$$

For convenience of notation, let it be assumed that if $n_{k_n} = n$, the length of sample available is actually $n + L_n$. Also note that m may be written m_j without ambiguity.

In this setup, (6.66) may be replaced by $\mathrm{E}(G_n^N - G_n^{N*})^2 \le 2(A_n + B_n)$ where

$$\begin{aligned} A_n = &\frac{1}{n^2 K(n)^2} \sum_{j=1}^{k_n} \sum_{j'=1}^{k_n} \sum_{m_j = n_{j-1}+1+L_n}^{n_j - 1 + L_n} \sum_{p_j = n_{j-1}+1}^{m_j - L_n} a_{nm_j p_j} \\ &\times \sum_{m_{j'} = n_{j'-1}+1+L_n}^{n_{j'} - 1 + L_n} \sum_{p_{j'} = n_{j'-1}}^{m_{j'} - L_n} a_{nm_{j'} p_{j'}} \mathrm{E}(u_{p_j} w_{m_j+1} u_{p_{j'}} w_{m_{j'}+1}). \end{aligned} \tag{8.19}$$

and

$$\begin{aligned} B_n = &\frac{1}{n^2 K(n)^2} \sum_{j=1}^{k_n} \sum_{j'=1}^{k_n} \sum_{m_j = n_{j-1}+1+L_n}^{n_j - 1 + L_n} \sum_{p_j = n_0}^{n_{j-1}} (a_{nm_j p_j} - a_{nn_{j-1} p_j}) \\ &\times \sum_{m_{j'} = n_{j'-1}+1+L_n}^{n_{j'} - 1 + L_n} \sum_{p_{j'} = n_0}^{n_{j'-1}} (a_{nm_{j'} p_{j'}} - a_{nn_{j'-1} p_{j'}}) \mathrm{E}(u_{p_j} w_{m_j+1} u_{p_{j'}} w_{m_{j'}+1}). \end{aligned} \tag{8.20}$$

According to Assumption **8.1**(c),

$$\begin{aligned} \mathrm{E}(u_{p_j} w_{m_j+1} u_{p_{j'}} w_{m_{j'}+1}) &\le \mu_{uw}^4(|p_j - p_{j'}|, m_j - p_j, m_{j'} - p_{j'}) \\ &= \gamma_{uw}(m_j - p_j)\gamma_{uw}(m_{j'} - p_{j'}) + \Big((\mu_{uw}^4(|p_j - p_{j'}|, m_j - p_j, m_{j'} - p_{j'}) \\ &\qquad - \gamma_{uw}(m_j - p_j)\gamma_{uw}(m_{j'} - p_{j'}) \Big) \\ &\ll L_n^{-2(1+\delta)} + |p_j - p_{j'}|^{-1-\delta}. \end{aligned} \tag{8.21}$$

The number of instances of expectation (8.21) appearing in (8.19) is of order $(n/k_n)^4$. Since $a_{nm_j p_j}/K(n) = O(1)$ and $|p_j - p_{j'}| \ll (n/k_n)|j - j'|$, it is possible to bound A_n as

$$A_n \ll \frac{k_n^2}{n^2} \Big(\frac{n}{k_n}\Big)^4 L_n^{-2(1+\delta)} + \frac{1}{n^2} \Big(\frac{n}{k_n}\Big)^{3-\delta} \sum_{j=1}^{k_n} \sum_{j'=1}^{k_n} |j - j'|^{-1-\delta}.$$

In the second of these terms, writing m for $|j - j'|$, in view of the summability of $m^{-1-\delta}$ it is found that

$$\sum_{j=1}^{k_n} \sum_{j'=1}^{k_n} |j - j'|^{-1-\delta} = \sum_{j=1}^{k_n} \Bigg(\sum_{m=1}^{j-1} m^{-1-\delta} + \sum_{m=1}^{k_n - j} m^{-1-\delta} \Bigg) \ll k_n. \tag{8.22}$$

Hence, after simplification, the bound on A_n has the form

$$A_n \ll \frac{n^2}{k_n^2} L_n^{-2(1+\delta)} + \frac{n^{1-\delta}}{k_n^{2-\delta}} = O(n^{\max\{C_1,C_2\}}) \tag{8.23}$$

where $C_1 = 2(1-\mu-\alpha(1+\delta))$ and $C_2 = 1-\delta-\mu(2-\delta)$.

For B_n consider the double sum in (8.20) which for each pair (j,j') contains the product of two terms. Taking the j^{th} term, applying the c_r inequality[1] with $r = 2$ and then Lemma **6.5**(iv) gives, since $L_n = o(n)$,

$$\begin{aligned}
&\frac{1}{nK(n)} \sum_{m_j=n_{j-1}+1+L_n}^{n_j-1+L_n} \left| \sum_{p_j=n_0}^{n_{j-1}} (a_{nm_jp_j} - a_{nn_{j-1}p_j}) \right| \\
&\ll \sum_{m_j=n_{j-1}+1+L_n}^{n_j-1+L_n} \left(\frac{1}{nK(n)^2} \sum_{p_j=n_0}^{n_{j-1}} (a_{nm_jp_j} - a_{nn_{j-1}p_j})^2 \right)^{1/2} \\
&\ll \sum_{m_j=n_{j-1}+1+L_n}^{n_j-1+L_n} \left(\frac{m_j - n_{j-1}}{n} \right)^{d_x+1/2} \\
&\ll \frac{(n_j - n_{j-1})^{d_x+3/2}}{n^{d_x+1/2}} \ll \frac{n}{k_n^{d_x+3/2}}.
\end{aligned}$$

The same bound holds for j', so for each pair (j,j') the products are bounded by a scale factor of order $n^2/k_n^{2d_x+3}$. Applying (8.22), in a similar way to (8.23) the bound on B_n takes the form

$$\begin{aligned}
B_n &\ll k_n^2 \frac{n^2}{k_n^{2d_x+3}} L_n^{-2(1+\delta)} + \frac{n^2}{k_n^{2d_x+3}} \left(\frac{n}{k_n}\right)^{-1-\delta} \sum_{j=1}^{k_n} \sum_{j'=1}^{k_n} |j-j'|^{-1-\delta} \\
&= O(n^{\max\{D_1,D_2\}})
\end{aligned} \tag{8.24}$$

where $D_1 = 2(1-\mu(d_x+\frac{1}{2})-\alpha(1+\delta))$ and $D_2 = 1-\delta-\mu(2d_x+1-\delta)$.

It is evident that the bound in (8.24) dominates that in (8.23) for every $d_x < \frac{1}{2}$, so the proof is concluded by noting that (8.16) is the bound on δ as a function of μ and α that ensures both $D_1 < 0$ and $D_2 < 0$. ∎

In the usual way, the counterpart result for G_{3n}^N is stated for the record.

8.15 Theorem $G_{3n}^N \to_{\mathrm{d}} \Xi_3^N$ under Assumptions **6.9** and **8.1** if, similarly to Theorem **8.14**, there exist $\mu < 1$ and $\alpha < 1$ such that

$$\max\left\{ \frac{1-\mu(d_y+\frac{1}{2})}{\alpha} - 1, \frac{1-2\mu(d_y+\frac{1}{2})}{1-\mu} \right\} < \delta \quad \square \tag{8.25}$$

[1] Proved as SLT Theorem 2.21.

As with Theorem **6.16**, the argument is identical following substitution of the complementary formulae, noting only that Lemma **6.6** has to be cited in place of Lemma **6.5**.

Theorems **8.14** and **8.15** are the only results to be encountered in this chapter that may impose a positive lower bound on δ. However, conditions (8.16) and (8.25) should be viewed as no more than sufficient bounds on the amount of dependence permitted and are very likely to be stronger than necessary. First, note how the c_r inequality was needed to bound B_n in (8.20) to deal with the fact that the squared sums include $O(n^2)$ terms, when all but $O(n)$ of these terms are constrained to zero in Theorems **6.16** and **6.17**. Another consideration is that the various values of δ specified by Assumption **8.1** (six in total) need not be identical, whereas conditions (8.16) and (8.25) necessarily cite the largest of these, should they differ. In spite of these qualifications, if the processes are long memory with positive fractional parameter the conditions can be met for any $\delta > 0$, by setting α and μ close enough to 1. The necessity to constrain δ arises only in the antipersistent cases where d_x or d_y are negative.

Taken together, all the foregoing considerations make it possible to state the generalization of Theorem **6.10**, which is most simply given as follows.

8.16 Theorem Theorem **6.10** continues to hold without Assumption **4.1** if Assumption **8.1** holds for δ such that conditions (8.16) and (8.25) are satisfied. □

8.5 Variance Formulae

The one task remaining is to generalize variance formulae (6.48) and (6.60), by a fairly straightforward application of the type of argument leading to (8.11).

8.17 Theorem If Assumption **8.1** holds, then (6.48) is replaced by

$$\mathrm{E}(Q(t)^2) = \begin{cases} \omega_u^2 \int_{-\infty}^t A^2(t,s)\mathrm{d}s, & d_y > 0 \\ \omega_u^2 \int_{-\infty}^t A^{*2}(t,s)\mathrm{d}s, & d_y < 0. \end{cases} \tag{8.26}$$

Proof With q_{nm} defined in (5.4), Assumption **8.1**(a) implies

$$\begin{aligned} \mathrm{E}(q_{nm}^2) &= \frac{1}{nK(n)^2}\mathrm{E}\bigg(\sum_{p=-\infty}^{m} a_{nmp}u_p\bigg)^2 \\ &= \frac{1}{nK(n)^2}\sum_{p=-\infty}^{m} a_{nmp}^2\bigg(\sigma_u^2 + 2\sum_{r=1}^{\infty}\Big(1 + \frac{a_{nm(p-r)} - a_{nmp}}{a_{nmp}}\Big)\gamma_u(r)\bigg) \\ &= \frac{1}{nK(n)^2}\sum_{p=-\infty}^{m} a_{nmp}^2\bigg(\omega_u^2 + 2\sum_{r=1}^{\infty}\frac{a_{nm(p-r)} - a_{nmp}}{a_{nmp}}\gamma_u(r)\bigg). \end{aligned} \tag{8.27}$$

The proof is therefore completed by showing that the second term in the parentheses in (8.27) is of small order. Given the summability of the autocovariances, it suffices to show that the terms of this sum are of sufficiently small order as $n \to \infty$, for each finite r.

Consider the case $d_y > 0$. Define t and s by the substitutions $[nt]$ for m and $[ns]$ for p. When r is finite a straightforward modification of the arguments leading to Lemma **5.2**, also noting $a_{n[nt][ns]} = O(K(n))$, gives

$$
\begin{aligned}
&\frac{a_{n[nt]([ns]-r)} - a_{n[nt][ns]}}{a_{n[nt][ns]}} \\
&\backsimeq \frac{1}{K(n)} \sum_{l=\max\{1-[nt],0\}}^{n-1-[nt]} l^{d_y-1}\Big(\big((l+n(t-s)+r)^{d_x} - (l+n(t-s))^{d_x}\big) \\
&\qquad\qquad - 1_{\{s\le 0\}}\big((1-ns+r)^{d_x} - (1-ns)^{d_x}\big)\Big) \\
&= \frac{1}{n} \sum_{l=\max\{1-[nt],0\}}^{n-1-[nt]} \left(\frac{l}{n}\right)^{d_y-1}\left(\left(\frac{l}{n}+t-s\right)^{d_x}\left(\left(1+\frac{r}{l+n(t-s)}\right)^{d_x} - 1\right)\right. \\
&\qquad\qquad \left. - 1_{\{s\le 0\}}\left(\left(\frac{1}{n}-s\right)^{d_x}\left(\left(1+\frac{r}{1-ns}\right)^{d_x}-1\right)\right)\right) \\
&= O(n^{-1})
\end{aligned}
\tag{8.28}
$$

noting that for fixed r, the contents of the large parentheses in the penultimate member of (8.28) are $O(1/n)$, for each l of the sum.

Hence, from (8.27) and **8.1**(a), noting $q_n(t) = q_{n[nt]}$ and that the $\gamma_u(r)$ are summable,

$$
\mathrm{E}(q_n(t)^2) = \frac{\omega_u^2}{nK(n)^2} \sum_{s=-\infty}^{t} a_{n[nt][ns]}^2 + O(n^{-1}) \to \omega_u^2 \int_{-\infty}^{t} A(t,s)^2 \mathrm{d}s
$$

as $n \to \infty$. If $d_y < 0$ then the alternative representation of (5.5) using (5.26) must be adopted, but a similar argument based on Lemma **5.3** yields the same conclusion. ∎

The companion result for G_{3n} is stated as follows for completeness, the proof being left for the reader to supply.

8.18 Theorem If Assumption **8.1** holds, then (6.60) is replaced by

$$
\mathrm{E}(H(s)^2) = \begin{cases} \omega_w^2 \int_{-\infty}^{s} E^2(s,t)\mathrm{d}t, & d_y > 0 \\ \omega_w^2 \int_{-\infty}^{s} E^{*2}(s,t)\mathrm{d}t, & d_y < 0. \end{cases} \quad \square
$$

Chapter 9

Frequency Domain Analysis

This chapter sets out the basics of an alternative approach to long memory analysis. The frequency domain is the favoured setting for nonparametric investigations of long memory since the parameter d can be estimated from the periodogram without any further assumptions about functional form. These techniques are well covered in the literature and will not be discussed here since the focus, as before, is on modelling long memory and the convergence of partial sums to fractional Brownian motion. Although the techniques of analysis are very different, the findings match those already obtained in the time domain, highlighting the fact that these are two complementary ways to study the same models. The modelling framework is somewhat limited in scope compared to the time domain, but some asymptotic results can be derived more easily and elegantly in the frequency context.

The chapter draws material from [17], joint work with Nigar Hashimzade and inspired by [11] among other sources. Other useful references include [10], [9], [73], [61], [23], [57], and [51]. A minor issue of notation arises in this chapter, because the frequent appearance of complex-valued terms threatens confusion with the use of the symbol i as the observation index in discrete time. The symbols t and s are therefore used in this context, with the symbol r representing a location in the time continuum.

9.1 Harmonizable Representation

The so-called harmonizable representation (equivalently, spectral representation) of a stationary stochastic process in the time domain assigns a distribution to random variations at different frequencies, instead of at different points in time. The source of the variations is taken to be a continuous, complex-valued process $U : [-\pi, \pi] \mapsto \mathbb{C}$, where $U(\lambda)$ for $\lambda \in [-\pi, \pi]$ denotes a process coordinate and $\overline{U}(\lambda)$ is its complex conjugate. The key feature of U is that it has orthogonal increments. Letting $\mathrm{d}\lambda$ denote an increment of the line and $\mathrm{d}U(\lambda)$ the variation of the process over this increment, the following properties are assumed, for $0 < \sigma_u^2 < \infty$:

$$d\overline{U}(\lambda) = dU(-\lambda) \tag{9.1a}$$

$$E(dU(\lambda)) = 0 \tag{9.1b}$$

$$E(dU(\lambda)d\overline{U}(\mu)) = \begin{cases} \sigma_u^2 d\lambda, & \lambda = \mu \\ 0, & \text{otherwise.} \end{cases} \tag{9.1c}$$

The process is symmetric about zero apart from the switch to the complex conjugate according to (9.1a). The two-sided domain $[-\pi, \pi]$ is in truth a mathematical convenience rather than a necessity, since one or the other half of it contains all the information about the process. Property (9.1c) implies, for the case $\lambda_2 > \lambda_1$,

$$\begin{aligned} E|U(\lambda_2) - U(\lambda_1)|^2 &= E\left(\int_{\lambda_1}^{\lambda_2} dU(\lambda) \int_{\lambda_1}^{\lambda_2} d\overline{U}(\mu)\right) \\ &= \sigma_u^2 \int_{\lambda_1}^{\lambda_2} d\lambda = \sigma_u^2(\lambda_2 - \lambda_1). \end{aligned} \tag{9.2}$$

A case fulfilling the conditions of (9.1) has the form $U(\lambda) = \sigma_u(A(\lambda) + iB(\lambda))/\sqrt{2}$ and $U(-\lambda) = \sigma_u(A(\lambda) - iB(\lambda))/\sqrt{2}$ for $0 \le \lambda \le \pi$, where A and B are segments of standard Brownian motions with $A(0) = B(0) = 0$. There is no explicit requirement in spectral theory that U be Gaussian, but a process with finite variance that is both continuous and has independent increments must also be Gaussian.[1]

The role played by U in defining the distribution of a time-domain process is as the integrator function in a stochastic Stieltjes integral. The leading example is the harmonizable representation of a white noise time domain sequence $\{u_1, \dots, u_n\}$, with mean zero and variance σ_u^2. This is connected to U via a Fourier transform, according to

$$u_t = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{it\lambda} dU(\lambda), \ t = 1, \dots, n. \tag{9.3}$$

At given time t, the oscillating function $e^{it\lambda}$ selects for each λ variously large and small contributions from U, to be added to u_t. In this way the function in (9.3) contributes random variations to u_t at different frequencies. The identity

$$\int_{-\pi}^{\pi} e^{it\lambda} dU(\lambda) = \int_{-\pi}^{\pi} e^{-it\lambda} d\overline{U}(\lambda)$$

means that under the assumptions of (9.1), u_t in (9.3) and its moments are real-valued.

The domain of λ is bounded by $\pm\pi$ since π corresponds to the highest frequency over which variations can be observed in a sample of length n. The function $e^{it\pi}$ oscillates $n/2$ times as t ranges from 1 to n, and to see higher frequencies requires a longer sample. Similarly the lowest observable frequency, with a half-cycle over the sample period, is $\lambda = 1/n$.

---

[1] See SLT Theorem 28.21.

The fact that $\{u_t\}$ in (9.3) forms an uncorrelated sequence is verified by

$$\begin{aligned}\mathrm{E}(u_t u_s) &= \frac{1}{2\pi}\mathrm{E}\left(\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}t\lambda}\mathrm{d}U(\lambda)\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}s\mu}\mathrm{d}\overline{U}(\mu)\right)\\ &= \frac{\sigma_u^2}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}(t-s)\lambda}\mathrm{d}\lambda\\ &= \begin{cases}\sigma_u^2, & t=s\\ 0, & \text{otherwise}\end{cases} \end{aligned}\tag{9.4}$$

(see (B.17)). The uncorrelatedness property maps directly via the Fourier transform from the uniform distribution of U over frequencies, as demonstrated by (9.2), giving rise to form of the expectation in (9.4).

Given (9.3) as a starting point, an autocorrelated process is constructed by inserting a (in general) complex-valued function $h(\lambda)$, known as a transfer function or frequency response function, into the integral. The transfer function defines in effect a stochastic process $h(\lambda)U(\lambda)$ for $\lambda \in [-\pi, \pi]$, having heteroscedastic increments and so assigning greater or lesser variations, on average, to different frequencies. These map into different modes of autocorrelation under the Fourier transform.

To construct the harmonizable representation of a general moving average process $x_t = \varphi(B)u_t$ where $\varphi(B)$ is a lag polynomial of infinite order, define the transfer function

$$h(\lambda) = \varphi(\mathrm{e}^{-\mathrm{i}\lambda}) = \sum_{j=0}^{\infty}\varphi_j \mathrm{e}^{-\mathrm{i}j\lambda}. \tag{9.5}$$

Substitute from (9.3) so as to write

$$\begin{aligned}x_t = \sum_{j=0}^{\infty}\varphi_j u_{t-j} &= \frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\varphi_j\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}(t-j)\lambda}\mathrm{d}U(\lambda)\\ &= \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}t\lambda}h(\lambda)\mathrm{d}U(\lambda)\end{aligned}\tag{9.6}$$

for $t = 1, \ldots, n$. Thus, in the linear framework h has the functional form of the lag polynomial with $\mathrm{e}^{-\mathrm{i}\lambda}$ taking the place of the lag operator B. The transfer function of a stationary ARMA(p, q) process $\phi(B)x_t = \theta(B)u_t$ takes the form

$$h(\lambda) = \frac{\theta(\mathrm{e}^{-\mathrm{i}\lambda})}{\phi(\mathrm{e}^{-\mathrm{i}\lambda})}. \tag{9.7}$$

More generally, any continuous function on $[0, \pi]$ can serve as a transfer function and a popular procedure (e.g. [24]) is to specify it semiparametrically.

To quantify the mapping on the real line, the squared modulus of h with suitable scaling factors defines the spectral density of the process,

$$f(\lambda) = \frac{\sigma_u^2}{2\pi}|h(\lambda)|^2. \tag{9.8}$$

Adapting the manipulation of (9.4), application of identity (B.2) of Appendix B gives the autocovariances as

$$\begin{aligned}\gamma_k = \mathrm{E}(x_t x_{t-k}) &= \frac{\sigma_u^2}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}k\lambda}|h(\lambda)|^2 \mathrm{d}\lambda \\ &= 2\int_0^{\pi} f(\lambda)\cos(k\lambda)\mathrm{d}\lambda,\ \ k = 0, 1, 2, \ldots \end{aligned} \tag{9.9}$$

The better-known spectral density formula

$$f(\lambda) = \frac{\gamma_0}{2\pi} + \frac{1}{\pi}\sum_{j=1}^{\infty}\gamma_j \cos j\lambda \tag{9.10}$$

can be deduced by substituting it into (9.9) and using solution (B.18) to produce

$$\begin{aligned}&2\int_0^{\pi}\left(\frac{\gamma_0}{2\pi}+\frac{1}{\pi}\sum_{j=1}^{\infty}\gamma_j\cos(j\lambda)\right)\cos(k\lambda)\mathrm{d}\lambda \\ &\quad = \frac{\gamma_0}{\pi}\int_0^{\pi}\cos(k\lambda)\mathrm{d}\lambda + \frac{2}{\pi}\sum_{j=1}^{\infty}\gamma_j\int_0^{\pi}\cos(j\lambda)\cos(k\lambda)\mathrm{d}\lambda \\ &\quad = \gamma_k.\end{aligned}$$

For the linear moving average case in particular, (9.5) and (B.2) produce the expansion

$$|h(\lambda)|^2 = \varphi(\mathrm{e}^{-\mathrm{i}\lambda})\varphi(\mathrm{e}^{\mathrm{i}\lambda}) = \sum_{j=0}^{\infty}\varphi_j^2 + 2\sum_{k=1}^{\infty}\left(\sum_{j=0}^{\infty}\varphi_j\varphi_{j+k}\right)\cos k\lambda. \tag{9.11}$$

Formulae (9.8) and (9.10) then give

$$\gamma_k = \sigma_u^2\sum_{j=0}^{\infty}\varphi_j\varphi_{j+k}$$

providing an alternative derivation of (1.6).

9.2 The Fractional Model

The simplest case of long memory in this framework is the fractionally integrated moving average, having transfer function

$$h(\lambda) = \sum_{j=0}^{\infty} b_j \mathrm{e}^{-\mathrm{i}\lambda j} = (1-\mathrm{e}^{-\mathrm{i}\lambda})^{-d} \tag{9.12}$$

where $b_j = \Gamma(d+j)/(\Gamma(d)\Gamma(j+1))$ as in (1.12) and the second equality is obtained by applying the generalized binomial expansion in (1.8). For this example the spectral density has the alternative forms (with the benefit of trigonometric identities (B.2) and (B.9))

$$f(\lambda) = \frac{\sigma_u^2}{2\pi}|1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2d} = \frac{\sigma_u^2}{2^{d+1}\pi}(1 - \cos\lambda)^{-d} = \frac{\sigma_u^2}{2^{2d+1}\pi}(\sin\lambda/2)^{-2d}. \quad (9.13)$$

In view of the fact that $(2\sin\lambda/2)^2 = |\lambda|^2 + O(|\lambda|^4)$, $|h(\lambda)|^2$ is approximated by $|\lambda|^{-2d}$ at low frequencies with $|\lambda|$ close to zero. If $d > 0$, f diverges at the origin, a phenomenon characterizing all long memory processes having the properties described in Chapter 1. When $d < 0$ on the other hand, $f(0) = 0$, which is the spectral property characterizing antipersistence. By contrast, weakly dependent processes have spectral densities that are bounded at the origin and also, except in the antipersistent case, bounded away from zero at the origin.

The estimation of the spectral density is a central topic in time series analysis, but it is important not to lose sight of the fact that it measures only one aspect of a dynamic process. The model in (9.6) is said to be causal, because x_t reflects the arrow of time in being driven by present and past shocks while it is independent of future shocks. Contrast (9.12) with a non-causal, forward-looking moving average in which $b_j = \Gamma(d+j)/(\Gamma(d)\Gamma(1+j))$ is the coefficient of u_{t+j} for each $j > 0$, while the coefficient of u_{t-j} is zero. The transfer function for this model is $(1 - \mathrm{e}^{\mathrm{i}\lambda})^{-d}$ and the spectral density is also (9.13).

More dramatically, consider a symmetric two-sided moving average model (also non-causal)

$$x_t = \sum_{k=-\infty}^{\infty} b_k^* u_{t-k}$$

where $b_k^* = b_{-k}^* = \sum_{j=0}^{\infty} b_j b_{k+j}$. Calling the transfer function of this model $h^*(\lambda)$, it can be verified (compare (9.11)) that

$$h^*(\lambda) = \sum_{k=-\infty}^{\infty} b_k^* \mathrm{e}^{\mathrm{i}k\lambda} = b(\mathrm{e}^{-\mathrm{i}\lambda})b(\mathrm{e}^{\mathrm{i}\lambda}) = |h(\lambda)|^2.$$

In particular, in the fractionally integrated case with parameter $d/2$ so that $b_j = \Gamma(d/2 + j)/(\Gamma(d/2)\Gamma(j + 1))$, (9.12) gives

$$h^*(\lambda) = (1 - \mathrm{e}^{-\mathrm{i}\lambda})^{-d/2}(1 - \mathrm{e}^{\mathrm{i}\lambda})^{-d/2} = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-d}.$$

The spectral density of this model is also identical with (9.13). In this case the transfer function is real, not complex, which is always the case with time-symmetric models, but this information is lost in taking the modulus.

The implication of these examples is that spectral densities contain no information about time ordering and directions of causation. However, one very useful feature of $f(\lambda)$ is the connection via (9.9) to the autocovariance sequence, which is likewise invariant to the direction of causality. Here, for the case of (9.12), is the harmonic counterpart of the time domain derivation of γ_k in Theorem **1.3**.

9.1 Theorem If $b_j = \Gamma(d+j)/(\Gamma(d)\Gamma(j+1))$ in (9.6) with $|d| < \frac{1}{2}$,

$$\gamma_k = \sigma_u^2 \frac{\Gamma(1-2d)\Gamma(d+k)}{\Gamma(1-d+k)} \frac{\sin \pi d}{\pi}. \tag{9.14}$$

Proof Inserting (9.13) into formula (9.9) gives

$$\gamma_k = \frac{\sigma_u^2}{2^{2d}\pi} \int_0^{\pi} (\sin \lambda/2)^{-2d} \cos(k\lambda)\mathrm{d}\lambda. \tag{9.15}$$

The integral solution in (B.19) with $x = \lambda/2$, $\nu = 1-2d$, and $a = 2k$, using (B.14) and also noting $\cos(k\pi) = (-1)^k$ for integer k by (B.4), produces

$$\begin{aligned}\int_0^{\pi} (\sin \lambda/2)^{-2d} \cos(k\lambda)\mathrm{d}\lambda &= \frac{\pi\Gamma(2-2d)(-1)^k}{2^{-2d}(1-2d)\Gamma(1-d+k)\Gamma(1-d-k)} \\ &= \frac{\Gamma(1-2d)\Gamma(d+k)}{2^{-2d}\Gamma(1-d+k)} \sin \pi d. \end{aligned} \tag{9.16}$$

Here, the second equality uses (B.13), (B.15), and then (B.5). The proof is completed by substituting (9.16) into (9.15). ∎

9.3 The Partial Sum Process

The normalized partial sum of the fractional process defined by (9.12) has harmonizable representation as follows, after substituting from (9.6) and resolving the sum of the terms $\mathrm{e}^{\mathrm{i}t\lambda}$ as a geometric series.

$$\begin{aligned} X_n(r) &= \frac{1}{n^{d+1/2}} \sum_{t=1}^{[nr]} x_t \\ &= \frac{1}{n^{d+1}\sqrt{2\pi/n}} \int_{-\pi}^{\pi} \sum_{t=1}^{[nr]} \mathrm{e}^{\mathrm{i}t\lambda}(1-\mathrm{e}^{-\mathrm{i}\lambda})^{-d}\mathrm{d}U(\lambda). \\ &= \frac{1}{n^{d+1}\sqrt{2\pi/n}} \int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}([nr]+1)\lambda} - \mathrm{e}^{\mathrm{i}\lambda}}{\mathrm{e}^{\mathrm{i}\lambda}-1}(1-\mathrm{e}^{-\mathrm{i}\lambda})^{-d}\mathrm{d}U(\lambda). \end{aligned} \tag{9.17}$$

The question of interest is, how should this formulation behave as the sample size increases, to complement the time domain device of mapping integer dates into a continuum? The trick is to make a change of variable in the integral from λ to λ/n. The function $\mathrm{e}^{\mathrm{i}t\pi/n}$ oscillates just half a cycle as t ranges from 1 to n, and changing the range of the integral from $[-\pi, \pi]$ to $[-n\pi, n\pi]$ extends the domain of U to accommodate the higher frequencies observable as n increases. Letting $X(r)$ denote the limiting case of $X_n(r)$, this is found heuristically as follows, making the change of variable in (9.17), rearranging, and letting $n \to \infty$.

$$X_n(r) = \frac{1}{\sqrt{2\pi}} \int_{-n\pi}^{n\pi} \frac{\mathrm{e}^{\mathrm{i}([nr]+1)\lambda/n} - \mathrm{e}^{\mathrm{i}\lambda/n}}{n(\mathrm{e}^{\mathrm{i}\lambda/n}-1)} \frac{(1-\mathrm{e}^{-\mathrm{i}\lambda/n})^{-d}}{n^d}\mathrm{d}U(\lambda)$$

$$\to X(r) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{i\lambda r} - 1}{i\lambda} (i\lambda)^{-d} dU(\lambda). \tag{9.18}$$

This argument does not amount to a proof of weak convergence, but the limit formula shows the harmonizable representation of fractional Brownian motion. If U is Gaussian, such proofs do not involve a central limit theorem as such and are a matter of showing uniform tightness of the sequence of measures and L_2-convergence of the increments.

The spectral density of this continuous-time process is $\sigma_u^2|\lambda|^{-2d}/2\pi$. Just as the generalized binomial form (1.12) is not the only model that converges to (2.1), so the limit in (9.18) features a transfer function that is common to all long memory models in the neighbourhood of zero, not only (9.12). However, to show convincingly that (9.18) is indeed the harmonizable representation of (2.1), it needs at least to be shown that the increment variances match as functions of d, apart from possible scale factors. This is done as follows.

9.2 Theorem If X is the limit process having the harmonizable representation in (9.18) with $|d| < \frac{1}{2}$,

$$E(X(r+\delta) - X(r))^2 = \frac{\sigma_u^2 \Upsilon_d}{\Gamma(d+1)^2} \delta^{2d+1} \tag{9.19}$$

where Υ_d is defined in (2.6).

Proof Using (9.18), (9.4), (B.2), and (B.9),

$$\begin{aligned} E(X(r+\delta) - X(r))^2 &= \frac{1}{2\pi} E\bigg(\int_{-\infty}^{\infty} e^{i\lambda r} \frac{e^{i\lambda\delta} - 1}{i\lambda} (i\lambda)^{-d} dU(\lambda) \\ &\qquad \times \int_{-\infty}^{\infty} e^{-i\mu r} \frac{e^{-i\mu\delta} - 1}{-i\mu} (-i\mu)^{-d} dU(\mu) \bigg) \\ &= \frac{\sigma_u^2}{\pi} \int_0^{\infty} |e^{i\lambda\delta} - 1|^2 |\lambda|^{-2d-2} d\lambda \\ &= \frac{4\sigma_u^2}{\pi} \int_0^{\infty} \sin^2(\lambda\delta/2) \lambda^{-2d-2} d\lambda. \end{aligned} \tag{9.20}$$

Setting $\mu = -2d-1$ and $a = \delta/2$ in (B.21), also noting that $\cos(-\pi d - \pi/2) = -\sin(\pi d)$ by (B.6), yields the result

$$\begin{aligned} \int_0^{\infty} \sin^2(\lambda\delta/2) \lambda^{-2d-2} d\lambda &= \frac{\Gamma(-2d-1)\sin(\pi d)\delta^{2d+1}}{2} \\ &= \frac{\pi\delta^{2d+1}}{4\Gamma(2d+2)\cos(\pi d)} \end{aligned} \tag{9.21}$$

where the second equality applies successively (B.15) with $x = -2d-1$, then (B.8), (B.6), and (B.7). Substituting into (9.20) gives

$$E(X(r+\delta) - X(r))^2 = \frac{\sigma_u^2 \delta^{2d+1}}{\Gamma(2d+2)\cos(\pi d)}. \quad \blacksquare \tag{9.22}$$

Applying (2.13) to (9.22) shows that this expression matches (9.19) when Υ_d is given by (2.6). Alternatively, a direct match is found using formula (2.15). Remarkably, the time domain formula in (2.4) appears at no stage in this derivation.

Comparing (9.19) with Theorem **2.2**, the formulae differ by the factor $1/\Gamma(d+1)^2$ whose role is to cancel the numerator of (2.15). The customary division by $\Gamma(d+1)$ of the formula in (2.1) can be understood as done to have the moments of the time-domain and harmonizable representations of the model agree, but as remarked on page 23, its inclusion is at root a matter of how to define the moving average coefficients. The factor appears in the denominator of (1.12) which matches the model whose transfer function appears in (9.12). If $1/\Gamma(d)$ were to be replaced by d in the latter formula, the denominator in (9.19) would disappear. It is difficult to pinpoint a reason to prefer one or other normalization. There is economy of notation on the one hand, but on the other hand there is the aesthetic appeal of the classic fractional model (1.12), based on the generalized binomial expansion.

A more significant difference between the representations is to be found in the comparison of Corollary **2.8** and Theorem **2.10**. The harmonic framework cannot accommodate a nonparametric dependence setup of the type captured by Assumption **1.2**(b). Short-run dynamics can enter only via the transfer function. For example, the process defined in (1.25) has $h(\lambda) = (1 - \mathrm{e}^{-\mathrm{i}\lambda})^{-d}\theta(\mathrm{e}^{-\mathrm{i}\lambda})$. Paralleling the development in (9.17) and (9.18), since $\theta(\mathrm{e}^{-\mathrm{i}\lambda/n}) \to \theta(1)$ as $n \to \infty$ when the lag coefficients are summable, when $x_t = (1-B)^{-d}\theta(B)u_t$ with u_t from (9.3) the limiting case of the partial sum process as $n \to \infty$ is found as

$$X_n(r) = \frac{1}{n^{d+1/2}}\sum_{t=1}^{[nr]} x_t \to X(r) = \frac{\theta(1)}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{\mathrm{i}\lambda r}-1}{\mathrm{i}\lambda}(\mathrm{i}\lambda)^{-d}\mathrm{d}U(\lambda).$$

Similarly to what was shown in Theorem **1.4**, the only effect of the weakly dependent shocks on the limit distribution is the possible change of scale.

9.4 Covariance Analysis

The next step is to study the relationships between different fractional processes. Consider a pair of frequency-domain processes U and W satisfying (9.1), with respective variances σ_u^2 and σ_w^2 and the additional condition

$$\mathrm{E}(\mathrm{d}U(\lambda)\mathrm{d}\overline{W}(\mu)) = \begin{cases} \sigma_{uw}\mathrm{d}\lambda, & \lambda = \mu \\ 0, & \text{otherwise} \end{cases} \tag{9.23}$$

where σ_{uw} is the covariance linking these variables. (Needless to say, there is no connection between this usage of the symbols U and W with that in §5.3.) Further suppose that X and Y are fractional Brownian motions with harmonic representations

$$X(r) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{\mathrm{i}\lambda r}-1}{\mathrm{i}\lambda}(\mathrm{i}\lambda)^{-d_x}\mathrm{d}U(\lambda) \tag{9.24}$$

and

$$Y(r) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{i\mu r} - 1}{i\mu} (i\mu)^{-d_y} dW(\mu) \tag{9.25}$$

for $r \in [0,1]$, as in (9.18).

The covariance of contemporaneous increments is found by a variation of Theorem **9.2** with an additional trick based on two useful identities. Since $e^{i\pi/2} = i$ and $e^{-i\pi/2} = -i$, it follows that

$$(i\lambda)^{-d} = |\lambda|^{-d} e^{-i\pi d \operatorname{sgn}(\lambda)/2} \tag{9.26}$$

and

$$(-i\lambda)^{-d} = |\lambda|^{-d} e^{i\pi d \operatorname{sgn}(\lambda)/2} \tag{9.27}$$

where $\operatorname{sgn}(\lambda)$ denotes the sign of λ, either $+1$ or -1.

9.3 Theorem If $|d_x| < \frac{1}{2}$ and $|d_y| < \frac{1}{2}$,

$$E(X(r+\delta) - X(r))Y((r+\delta) - Y(r))) = \frac{\sigma_{uw} \cos(\pi(d_x - d_y)/2)\delta^{d_x+d_y+1}}{\Gamma(d_x + d_y + 2)\cos(\pi(d_x + d_y)/2)}. \tag{9.28}$$

Proof Using (9.4) and then (B.2) and (B.9) in the final equality,

$$\begin{aligned}
&E(X(r+\delta) - X(r))Y((r+\delta) - Y(r))) \\
&\quad = \frac{1}{2\pi} E\Bigg(\int_{-\infty}^{\infty} e^{i\lambda r} \frac{e^{i\lambda\delta} - 1}{i\lambda} (i\lambda)^{-d_x} dU(\lambda) \\
&\qquad\qquad \times \int_{-\infty}^{\infty} e^{-i\mu r} \frac{e^{-i\mu\delta} - 1}{-i\mu} (-i\mu)^{-d_y} d\overline{W}(\mu) \Bigg) \\
&\quad = \frac{\sigma_{uw}}{2\pi} \int_{-\infty}^{\infty} |e^{i\lambda\delta} - 1|^2 |\lambda|^{-d_x - d_y - 2} e^{-i\pi \operatorname{sgn}(\lambda)(d_x - d_y)/2} d\lambda \\
&\quad = \frac{\sigma_{uw}}{\pi} (e^{-i\pi(d_x - d_y)/2} + e^{i\pi(d_x - d_y)/2}) \int_0^{\infty} |e^{i\lambda\delta} - 1|^2 \lambda^{-d_x - d_y - 2} d\lambda \\
&\quad = \frac{4\sigma_{uw}}{\pi} \cos(\pi(d_x - d_y)/2) \int_0^{\infty} \sin^2(\lambda\delta/2) \lambda^{-d_x - d_y - 2} d\lambda.
\end{aligned} \tag{9.29}$$

The integral in (9.29) is identical with that of the univariate case in (9.21), except for the replacement of $2d$ by $d_x + d_y$. ∎

It will not escape notice that formula (9.28) can also be written as

$$\frac{\sigma_{uw} \Upsilon_{xy} \delta^{d_x+d_y+1}}{\Gamma(d_x + 1)\Gamma(d_y + 1)}$$

where Υ_{xy} is defined in (4.34) and also in (2.53), and so matches the formula in (2.52) apart from the choice of normalization discussed in the previous section.

9.5 Stochastic Integral

The next result, for comparison with the time domain calculation leading to (4.22), is the expected value of the stochastic integral of X with respect to Y. The differential increment of Y is found from the harmonic representation (9.25) as

$$\begin{aligned}\mathrm{d}Y(r) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\partial}{\partial r}\frac{\mathrm{e}^{\mathrm{i}\mu r}-1}{\mathrm{i}\mu}(\mathrm{i}\mu)^{-d_y}\mathrm{d}W(\mu)\mathrm{d}r\\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{\mathrm{i}\mu r}(\mathrm{i}\mu)^{-d_y}\mathrm{d}W(\mu)\mathrm{d}r.\end{aligned}$$

Teaming this formula with (9.24) and forming the integral with respect to time gives

$$\begin{aligned}\int_0^1 X\mathrm{d}Y = \frac{1}{2\pi}\int_0^1\Bigg(\int_{-\infty}^{\infty}&\frac{\mathrm{e}^{\mathrm{i}\lambda r}-1}{\mathrm{i}\lambda}(\mathrm{i}\lambda)^{-d_x}\mathrm{d}U(\lambda)\\ &\times\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}\mu r}(-\mathrm{i}\mu)^{-d_y}\mathrm{d}\overline{W}(\mu)\Bigg)\mathrm{d}r.\end{aligned}$$

Substituting from (9.26) and (9.27) and applying (9.23), the expected value of this random variable has the form

$$\mathrm{E}\left(\int_0^1 X\mathrm{d}Y\right) = \frac{\sigma_{uw}}{2\pi}\int_0^1\int_{-\infty}^{\infty}\frac{1-\mathrm{e}^{-\mathrm{i}\lambda r}}{\mathrm{i}\lambda}|\lambda|^{-d_x-d_y}(\mathrm{e}^{\mathrm{i}\pi\mathrm{sgn}(\lambda)(d_y-d_x)/2})\mathrm{d}\lambda\mathrm{d}r. \tag{9.31}$$

9.4 Theorem If $|d_x| < \frac{1}{2}$, $|d_y| < \frac{1}{2}$, and $d_x + d_y > 0$,

$$\mathrm{E}\left(\int_0^1 X\mathrm{d}Y\right) = \sigma_{uw}\frac{\Gamma(1-d_x-d_y)\sin\pi d_y}{\pi(1+d_x+d_y)(d_x+d_y)}. \tag{9.32}$$

Proof To evaluate the double integral in (9.31), first make a change of variable $v = \lambda r$. Since $0 \le r \le 1$, $\mathrm{sgn}(\lambda) = \mathrm{sgn}(v)$ and (9.31) factorizes into two integrals, where the second one can be split into positive and negative regions over which $\mathrm{sgn}(v)$ is constant. Thus,

$$\begin{aligned}\mathrm{E}\left(\int_0^1 X\mathrm{d}Y\right) &= \frac{\sigma_{uw}}{2\pi}\int_0^1 r^{d_x+d_y}\mathrm{d}r\int_{-\infty}^{\infty}\frac{1-\mathrm{e}^{-\mathrm{i}v}}{\mathrm{i}v}|v|^{-d_x-d_y}\mathrm{e}^{-\mathrm{i}\pi\mathrm{sgn}(v)(d_x-d_y)/2}\mathrm{d}v\\ &= \frac{\sigma_{uw}}{2\pi(1+d_x+d_y)}\left(\mathrm{e}^{-\mathrm{i}\pi(d_x-d_y)/2}J + \mathrm{e}^{\mathrm{i}\pi(d_x-d_y)/2}\bar{J}\right)\end{aligned} \tag{9.33}$$

where, since $1-\mathrm{e}^{-\mathrm{i}v} = 1-\cos v + \mathrm{i}\sin v$ from (B.1),

$$\begin{aligned}J &= \int_0^{\infty}\frac{1-\mathrm{e}^{-\mathrm{i}v}}{\mathrm{i}v}|v|^{-d_x-d_y}\mathrm{d}v\\ &= -\mathrm{i}\int_0^{\infty}\frac{1-\cos v}{v^{1+d_x+d_y}}\mathrm{d}v + \int_0^{\infty}\frac{\sin v}{v^{1+d_x+d_y}}\mathrm{d}v.\end{aligned}$$

To resolve the first term of J, first apply (B.9) and then (B.21) with $\mu = -d_x - d_y$ and $a = \frac{1}{2}$, followed by (B.15) and (B.8), to get successively

$$\begin{aligned}\int_0^\infty \frac{1-\cos v}{v^{1+d_x+d_y}}\mathrm{d}v &= 2\int_0^\infty \frac{\sin^2(v/2)}{v^{1+d_x+d_y}}\mathrm{d}v \\ &= -\Gamma(-d_x-d_y)\cos(\pi(d_x+d_y)/2) \\ &= \frac{\pi}{2\Gamma(1+d_x+d_y)\sin(\pi(d_x+d_y)/2)}. \end{aligned} \tag{9.34}$$

For the second term of J, (B.20) with $\mu = -d_x - d_y$ and $a = 1$, then (B.15) and (B.8), give

$$\begin{aligned}\int_0^\infty \frac{\sin v}{v^{1+d_x+d_y}}\mathrm{d}v &= \Gamma(-d_x-d_y)\sin(\pi(-d_x-d_y)/2) \\ &= \frac{\pi}{2\Gamma(1+d_x+d_y)\cos(\pi(d_x+d_y)/2)}. \end{aligned} \tag{9.35}$$

Gathering the terms (9.34) and (9.35) under a common denominator and applying (B.1) and (B.8) yields

$$J = \frac{-\mathrm{i}\mathrm{e}^{\mathrm{i}\pi(d_x+d_y)/2}\pi}{\Gamma(1+d_x+d_y)\sin(\pi(d_x+d_y))}.$$

Then, with the further assistance of (B.3), (B.15), and (B.13),

$$\begin{aligned}\mathrm{e}^{-\mathrm{i}\pi(d_x-d_y)/2}J + \mathrm{e}^{\mathrm{i}\pi(d_x-d_y)/2}\bar{J} &= \frac{-\mathrm{i}(\mathrm{e}^{\mathrm{i}\pi d_y} - \mathrm{e}^{-\mathrm{i}\pi d_y})\pi}{\Gamma(1+d_x+d_y)\sin(\pi(d_x+d_y))} \\ &= \frac{2\pi\sin(\pi d_y)}{\Gamma(1+d_x+d_y)\sin(\pi(d_x+d_y))} \\ &= \frac{2\Gamma(1-d_x-d_y)\Gamma(d_x+d_y)\sin\pi d_y}{\Gamma(1+d_x+d_y)} \\ &= \frac{2\Gamma(1-d_x-d_y)\sin\pi d_y}{d_x+d_y}. \end{aligned} \tag{9.36}$$

Finally, on substituting (9.36) into (9.33) the result is (9.32). ∎

As expected, the formula in (9.32) matches λ_{xy} in (4.22) apart from scale factors $\Gamma(d_x+1)$ and $\Gamma(d_y+1)$ whose role has been discussed. Theorems **9.1**, **9.2**, **9.3**, and **9.4** all reproduce the corresponding formulae obtained for the time-domain representation of the fractional process, showing that (2.1) and (9.4) really are alternative representations of the same model. Nonetheless, a comparison of the proofs of Theorems **9.4** and **4.5** is an intriguing exercise, to say the least.

Chapter 10

Autoregressive Roots near Unity

Thinking of the fractional process in its partial-sum manifestation as providing a way of embedding the unit root within a more general class of nonstationary processes, it merits comparison with another model class that gives rise to Ornstein-Uhlenbeck processes in the limit instead of fractional Brownian motions. What is different about the near-unit root approach is that an array framework is essential. The concept of 'close to unity' is linked to sample size, the limit results being obtained by considering an autoregressive coefficient whose proximity to unity depends on n. Unlike the fractional case, there exist no stationary processes whose normalized partial sums can give rise to the asymptotic limits obtained in this theory. The theory does not so much point to an alternative modelling methodology as to attempt to throw light on the transition between the unit root case and the stable root or mildly explosive root cases of autoregressive models. This chapter is inspired mainly by a seminal paper of Peter Phillips [53]. Related references include among many others [54], [12], [58], and most recently [55].

10.1 Generalizing Unit Roots

Let β be a fixed parameter, define $b_{nj} = \mathrm{e}^{-\beta j/n}$ and let the shock sequence $\{u_i\}_{i=1}^n$ satisfy Assumption **1.1**. For each $n \in \mathbb{N}$ consider the triangular moving average array

$$x_{ni} = \sum_{j=0}^{i-1} b_{nj} u_{i-j},\ i = 1, \dots, n. \tag{10.1}$$

With $x_{n0} = 0$ the process has the autoregressive representation

$$x_{ni} = \mathrm{e}^{-\beta/n} x_{n,i-1} + u_i,\ i = 1, \dots, n \tag{10.2}$$

where $\mathrm{e}^{-\beta/n} = 1 - \beta/n + O(n^{-2})$, so that when n is large this is the unit root model with a small-order adjustment.

In the case with $\beta = 0$, which is exactly the unit root process, $n^{-1/2}x_{n[nt]} \to_{\mathrm{d}} \sigma_u B(t)$ for $t \in [0,1]$, where B is regular Brownian motion. The case $\beta \neq 0$ models a situation in which the autoregressive root, while not unity, is close enough that the process still diverges like $n^{1/2}$. After normalization the limit process, unlike Brownian motion, has dependent increments. With $\beta > 0$, as $n \to \infty$ the distribution is eventually stationary and independent of initial conditions. By contrast, the 'local to unity' fractional process (partial sum of a stationary fractional with parameter $d < \frac{1}{2}$) diverges like $n^{1/2+d}$. After normalization the corresponding limit process (2.1) also has dependent increments but, like Brownian motion, is nonstationary. The three classes of model therefore offer a notable contrast of limit properties.

10.2 The Covariance Function

It will be useful in the sequel to generalize the sequences in (10.1) to the form $\{x_{ni}, i = 1, \ldots, [nt]\}$ for any $t > 0$. The case $t > 1$ is important since it allows the asymptotic analysis to extend to continuous-time processes defined on the positive half-line. Letting $j = [nt] - i$, the substitution $\mathrm{e}^{\beta j/n}u_{[nt]-j} = \mathrm{e}^{\beta([nt]-i)/n}u_i$ is convenient for defining the normalized continuous time càdlàg process

$$X_n(t) = \frac{x_{n[nt]}}{\sqrt{n}} = \frac{\mathrm{e}^{-\beta[nt]/n}}{\sqrt{n}} \sum_{i=1}^{[nt]} \mathrm{e}^{\beta i/n} u_i. \tag{10.3}$$

10.1 Theorem Under Assumption **1.1**,

$$\lim_{n\to\infty} \mathrm{E}(X_n(t)X_n(s)) = \sigma_u^2 \frac{\mathrm{e}^{-\beta|t-s|} - \mathrm{e}^{-\beta(t+s)}}{2\beta}. \tag{10.4}$$

Proof First consider the variance, setting $s = t$. By Assumption **1.1**, $X_n(t)$ is a sum of $[nt]$ independent increments and the variance is

$$\mathrm{E}(X_n(t)^2) = \frac{\sigma_u^2}{n} \mathrm{e}^{-2\beta[nt]/n} \sum_{i=1}^{[nt]} \mathrm{e}^{2\beta i/n}. \tag{10.5}$$

Expressing $\mathrm{e}^{2\beta i/n}$ in power series form as $\sum_{k=0}^{\infty}(2\beta i/n)^k/k!$ and substituting the integral approximation formula

$$\sum_{i=1}^{[nt]} (i/n)^k = \frac{t^{k+1}}{k+1} + O(1/n)$$

for each $k \geq 0$, (10.5) can be written as

$$\mathrm{E}(X_n(t)^2) = \frac{\sigma_u^2}{n} \mathrm{e}^{-2\beta[nt]/n} \sum_{i=1}^{[nt]} \left(\sum_{k=0}^{\infty} \frac{(2\beta i/n)^k}{k!} \right)$$

$$= \sigma_u^2 \frac{\mathrm{e}^{-2\beta t}}{2\beta} \sum_{k=0}^{\infty} \frac{(2\beta t)^{k+1}}{(k+1)!} + O(1/n)$$

$$= \sigma_u^2 \frac{1 - \mathrm{e}^{-2\beta t}}{2\beta} + O(1/n). \tag{10.6}$$

In the same way, under Assumption **1.1**,

$$\begin{aligned} \mathrm{E}(X_n(t)X_n(s)) &= \frac{1}{n}\mathrm{e}^{-\beta([nt]+[ns])/n} \sum_{i=1}^{[nt]} \sum_{j=1}^{[ns]} \mathrm{e}^{\beta(i+j)/n} \mathrm{E}(u_i u_j) \\ &= \frac{\sigma_u^2}{n}\mathrm{e}^{-\beta([nt]+[ns])/n} \sum_{i=1}^{[n\min\{t,s\}]} \mathrm{e}^{2\beta i/n} \\ &= \sigma_u^2 \frac{\mathrm{e}^{-\beta(t+s)}}{2\beta} \sum_{k=0}^{\infty} \frac{(2\beta \min\{t,s\})^{k+1}}{(k+1)!} + O(1/n) \\ &= \sigma_u^2 \frac{\mathrm{e}^{-\beta(t+s)}}{2\beta} \left(\mathrm{e}^{2\beta \min\{t,s\}} - 1\right) + O(1/n). \end{aligned}$$

The limit (10.4) follows noting that

$$2\min\{t,s\} - t - s = -|t-s|. \quad \blacksquare \tag{10.7}$$

10.3 Weak Convergence

The limit in (10.4) specifies the covariance function of the Gaussian stochastic process $\sigma_u J_\beta : [0,\infty) \mapsto \mathbb{R}$ where

$$J_\beta(t) = \int_0^t \mathrm{e}^{-\beta(t-\xi)} \mathrm{d}B(\xi) \tag{10.8}$$

and B is standard Brownian motion. Thus,

$$\mathrm{E}(J_\beta(t)J_\beta(s)) = \mathrm{e}^{-\beta(t+s)} \int_0^{\min\{t,s\}} \mathrm{e}^{2\beta\xi}\mathrm{d}\xi = \frac{\mathrm{e}^{-\beta(t+s)}}{2\beta}\left(\mathrm{e}^{2\beta\min\{t,s\}} - 1\right). \tag{10.9}$$

If $\beta > 0$, it is easily seen in view of (10.7) that as $\min\{t,s\} \to \infty$,

$$\mathrm{E}(J_\beta(t)J_\beta(s)) \to \frac{\mathrm{e}^{-\beta|t-s|}}{2\beta}. \tag{10.10}$$

The limiting case of the variance is therefore simply $1/2\beta$. When t and s are sufficiently large the covariance function depends on $|t-s|$ but not on t or s and the limit process is accordingly stationary.

It is similarly not difficult to verify that $\mathrm{e}^{-\beta|t-s|}/2\beta$ is the covariance function of the stationary stochastic process

$$J_\beta^*(t) = \frac{\mathrm{e}^{-\beta t}}{\sqrt{2\beta}} B(\mathrm{e}^{2\beta t}). \tag{10.11}$$

When $\beta > 0$, $J_\beta(t)$ and $J_\beta^*(t)$ are equivalent Gaussian processes when t is large enough, although distinct when t is small with covariance functions (10.9) and (10.10) respectively. J_β is the so-called Ornstein-Uhlenbeck process, formally defined as the solution of the stochastic differential equation

$$\mathrm{d}J_\beta(t) = -\beta J_\beta(t) + \mathrm{d}B(t). \tag{10.12}$$

The limit process is also well defined for the case $\beta < 0$ since none of these results depend on the sign of β except for the asymptotic stationarity just noted. The limit formulae in (10.6) and (10.4) as $n \to \infty$ continue to apply, but instead of the variance converging to the finite limit of $\sigma_u^2/2\beta$ as $t \to \infty$, it diverges. The finite-n autoregression in (10.2) has a mildly explosive root in this case.

The next question to be resolved is the link between (10.8) and the empirical process (10.2). Consider the stochastic process

$$\mathrm{e}^{\beta s} B(s) : [0, t] \mapsto \mathbb{R}$$

where as before B denotes Brownian motion on the positive half-line. An increment of this product has the form

$$\mathrm{d}\big(\mathrm{e}^{\beta s} B(s)\big) = \beta \mathrm{e}^{\beta s} B(s)\mathrm{d}s + \mathrm{e}^{\beta s}\mathrm{d}B(s).$$

Integrating from 0 to t given $B(0) = 0$ yields the integration-by-parts formulation

$$\mathrm{e}^{\beta t} B(t) = \beta \int_0^t \mathrm{e}^{\beta s} B(s)\mathrm{d}s + \int_0^t \mathrm{e}^{\beta s}\mathrm{d}B(s).$$

Multiplying by $\mathrm{e}^{-\beta t}$ and rearranging gives for J_β in (10.8) the relation

$$J_\beta(t) = B(t) - \beta \int_0^t \mathrm{e}^{-\beta(t-s)} B(s)\mathrm{d}s. \tag{10.13}$$

Thus, J_β with $\beta > 0$ can be viewed as the sum of a Brownian motion and a negatively signed bias term based on an average of its recent past variations tending to push the process in the direction of mean reversion, the characteristic that is also evident from the form of (10.12).

To show that $\sigma_u J_\beta$ is indeed the weak limit of X_n in (10.3) consider the element W_n of the space of càdlàg processes on $[0, \infty)$, where

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} u_i. \tag{10.14}$$

Under Assumption **1.1**, Theorem **3.2** with $d = 0$ (for example) is an FCLT for a unit root process that generalizes straightforwardly from the unit interval to the case of $[0, t]$. This gives

$$W_n \xrightarrow{\mathrm{d}} \sigma_u B \tag{10.15}$$

where B is Brownian motion on the positive half-line. Noting $W_n(0) = 0$ and $W_n(i/n) - W_n((i-1)/n) = n^{-1/2}u_i$, write the telescoping sum representation

$$\begin{aligned}
\mathrm{e}^{\beta[nt]/n} W_n(t) &= \sum_{i=1}^{[nt]} \Big(\mathrm{e}^{\beta i/n} W_n(i/n) - \mathrm{e}^{\beta(i-1)/n} W_n((i-1)/n) \Big) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \mathrm{e}^{\beta i/n} u_i + (\mathrm{e}^{\beta/n} - 1) \sum_{i=1}^{[nt]} \mathrm{e}^{\beta(i-1)/n} W_n((i-1)/n). \qquad (10.16)
\end{aligned}$$

The first right-hand side sum of (10.16) is $\mathrm{e}^{\beta[nt]/n} X_n(t)$ according to (10.3). Considering the second sum, $\mathrm{e}^{\beta/n} - 1 = \beta/n + O(1/n^2)$. Since W_n is a step function with $W_n(s) = W_n(i/n)$ when $i/n \le s < (i+1)/n$,

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{[nt]} \mathrm{e}^{\beta(i-1)/n} W_n((i-1)/n) &= \sum_{i=0}^{[nt]-1} \mathrm{e}^{\beta i/n} W_n(i/n) \int_{i/n}^{(i+1)/n} \mathrm{d}s \\
&= \int_0^t \mathrm{e}^{\beta s} W_n(s) \mathrm{d}s + O(1/n). \qquad (10.17)
\end{aligned}$$

Substituting (10.17) into (10.16) and rearranging yields the relation

$$X_n(t) = W_n(t) - \beta \int_0^t \mathrm{e}^{-\beta(t-s)} W_n(s) \mathrm{d}s + O(1/n). \tag{10.18}$$

This is the finite-sample counterpart of (10.13). These arguments strongly indicate that the process $X_n(t)$ has a Gaussian weak limit. Since J_β in (10.13) is a continuous functional of B, the following theorem is a consequence of (10.18) and the continuous mapping theorem.[1]

10.2 Theorem If $W_n \to_{\mathrm{d}} \sigma_u B$ then $X_n \to_{\mathrm{d}} \sigma_u J_\beta$. $\square$

However, it is also of interest to pursue direct arguments for the FCLT, of the type of used in Theorem **3.2**. There is the same kind of problem here as was found with fractional processes, that X_n is not a simple cumulation of shocks since the moving average coefficients depend on t as well as the shock date i. However, for any fixed t, $X_n(t)$ is a linear function of shocks with nonsummable coefficients, so that a CLT can operate. The important steps in the following sketch proof have been worked out in detail in Theorem **3.2** and it suffices here to reproduce them briefly.

---

[1] See SLT Theorem 29.4.

10.3 Theorem Under Assumption **1.1**, $X_n \to_{\mathrm{d}} \sigma_u J_\beta$.

Proof (Outline) Consider the conditions of Lemma **3.4**. Setting

$$c_{ni} = n^{-1/2}\mathrm{e}^{-\beta([nt]-i)/n} = O(n^{-1/2}) \tag{10.19}$$

it is verified in the same way using Theorem **A.4** that

$$\sum_{i=1}^{n} c_{ni}^2 \mathrm{E}(u_i^2 1_{\{|c_{ni}u_i|>\varepsilon\}}) = o(n^{1-r/2})$$

for $\varepsilon > 0$, so the Lindeberg condition is satisfied. Under the assumptions this means in view of (10.6) that $X_n(t) \to_{\mathrm{d}} \mathrm{N}(0, \sigma_u^2(1-\mathrm{e}^{-2\beta t})/2\beta)$ for each $t > 0$. Subject to almost sure continuity, the process X_n/σ_u has a limit with characteristics matching (10.8).

To show uniform tightness, the main requirement is that the collection

$$\left\{ \sup_{t \le s \le t+\delta} \frac{(X_n(s) - X_n(t))^2}{v_n^2(t,\delta)}, n \in \mathbb{N} \right\} \tag{10.20}$$

is uniformly integrable for all $0 < \delta < 1$ and $t \in [0, 1-\delta]$, where $v_n^2(t,\delta) = \mathrm{E}(X_n(t+\delta) - X_n(t))^2$. Write

$$\begin{aligned} X_n(s) - X_n(t) &= \frac{1}{n^{1/2}} \sum_{i=[nt]+1}^{[ns]} \mathrm{e}^{-\beta([ns]-i)/n} u_i \\ &\quad + \frac{1}{n^{1/2}} \sum_{i=1}^{[nt]} (\mathrm{e}^{-\beta([ns]-i)/n} - \mathrm{e}^{-\beta([nt]-i)/n}) u_i \\ &= Y_{1n}(s,t) + Y_{2n}(s,t). \end{aligned}$$

The squares of these two sums may be considered separately, according to Theorem **A.7**.

In the case of Y_{1n}, direct application of Theorem **A.5** is not possible for the reasons encountered in Lemma **3.6**, that the moving average is not a simple partial sum. In this case the solution is to consider the partial sum process

$$Y_{1n}^*(s,t) = \frac{1}{n^{1/2}} \sum_{i=[nt]+1}^{[ns]} \mathrm{e}^{\beta([n(t+\delta)]-i)/n} u_i.$$

The moving average coefficients $\mathrm{e}^{\beta([n(t+\delta)]-i)/n}$ for $i = [nt]+1, \ldots, [n(t+\delta)]$ can be written (with order inverted) as the collection

$$\{1, \mathrm{e}^{\beta/n}, \mathrm{e}^{2\beta/n}, \ldots, \mathrm{e}^{([n(t+\delta)]-[nt]-1)\beta/n}\}. \tag{10.21}$$

This does not depend on s and all the elements appear in the terminal coordinate of the process $Y_{1n}^*(t+\delta, t)$. With

$$v_{1n}^2(t,\delta) = \frac{1}{n} \sum_{i=[nt]+1}^{[n(t+\delta)]} \mathrm{e}^{-2\beta([n(t+\delta)]-i)/n} \tag{10.22}$$

the collection $\{\sup_{t\leq s\leq t+\delta} Y_{1n}^*(s,t)^2/v_{1n}^2(t,\delta), n \in \mathbb{N}\}$ satisfies the conditions of Theorem **A.5**. This holds under any permutation p of the moving average weights in (10.21). Let the version of Y_{1n}^* with weights so permuted be denoted Y_{1np}^*. For any n there exists a permutation, say p_n^*, for which

$$\sup_{t\leq s\leq t+\delta} \frac{Y_{1n}(s,t)^2}{v_{1n}^2(t,\delta)} \leq \sup_{t\leq s\leq t+\delta} \frac{Y_{1np_n^*}^*(s,t)^2}{v_{1n}^2(t,\delta)}. \tag{10.23}$$

The majorants of (10.23) for $n \in \mathbb{N}$ form a uniformly integrable collection by Theorem **A.5**, and this property extends to the collection of the minorants of the inequality, by Theorem **A.6**.

The term $Y_{2n}(s,t)$ is not a partial sum over s and can be dealt with similarly to its counterpart in the proof of Lemma **3.6**. With

$$v_{2n}^2(t,\delta) = \frac{1}{n}\sum_{i=1}^{[nt]}(\mathrm{e}^{-\beta([n(t+\delta)]-i)/n} - \mathrm{e}^{-\beta([nt]-i)/n})^2 \tag{10.24}$$

notice that $v_{2n}^2(t,t-s) \leq v_{2n}^2(t,\delta)$ for every s. The collections $\{Y_{2n}(s,t)^2/v_{2n}^2(t,\delta), n \in \mathbb{N}\}$ are uniformly integrable by Theorem **A.5** and Theorem **A.6** for each $s \in [t, t+\delta]$ and hence in particular for the supremum with respect to s.

According to (10.22) $v_{1n}^2(t,\delta) = O(\delta)$, being the sum of $[n(t+\delta)]-[nt]$ positive terms, all below 1 and divided by n. Also, according to (10.24), $v_{2n}^2(t,\delta)$ is a sum of $[nt]$ terms divided by n and the terms have the form

$$(\mathrm{e}^{-\beta([n(t+\delta)]-i)/n} - \mathrm{e}^{-\beta([nt]-i)/n})^2 = O(\delta^2).$$

Since $v_n^2(t,\delta) = v_{1n}^2(t,\delta)+v_{2n}^2(t,\delta)$, the uniform integrability of (10.20) now follows by Theorem **A.7** and the uniform tightness proof is completed by an application of Lemma **3.8** with $L = 0$, $U = 1$ and $d = 0$. ∎

10.4 Stochastic Integral

Applying a telescoping sum argument to the squares of the process leads to another interesting asymptotic relation. Recalling that $x_{ni} = \mathrm{e}^{-\beta i/n}\sum_{k=1}^{i}\mathrm{e}^{\beta k/n}u_k$, let $z_{ni} = \mathrm{e}^{\beta i/n}x_{ni}$ and so note that

$$z_{ni}^2 = z_{n,i-1}^2 + \mathrm{e}^{2\beta i/n}u_i^2 + 2\mathrm{e}^{\beta i/n}z_{n,i-1}u_i. \tag{10.25}$$

Substituting from (10.25) for $z_{ni}^2 - z_{n,i-1}^2$ produces

$$\begin{aligned}
x_{ni}^2 - x_{n,i-1}^2 &= \mathrm{e}^{-2\beta i/n}z_{ni}^2 - \mathrm{e}^{-2\beta(i-1)/n}z_{n,i-1}^2 \\
&= \mathrm{e}^{-2\beta i/n}(1-\mathrm{e}^{2\beta/n})z_{ni}^2 + \mathrm{e}^{-2\beta(i-1)/n}(z_{ni}^2 - z_{n,i-1}^2) \\
&= \mathrm{e}^{-2\beta i/n}(1-\mathrm{e}^{2\beta/n})z_{ni}^2 + \mathrm{e}^{-2\beta(i-1)/n}(\mathrm{e}^{2\beta i/n}u_i^2 + 2\mathrm{e}^{\beta i/n}z_{n,i-1}u_i) \\
&= (1-\mathrm{e}^{2\beta/n})x_{ni}^2 + \mathrm{e}^{2\beta/n}u_i^2 + 2\mathrm{e}^{\beta/n}x_{n,i-1}u_i.
\end{aligned}$$

Given (10.3) and that $x_{n0} = 0$, the telescoping sum therefore has the form

$$X_n(1)^2 = \frac{x_{nn}^2}{n} = \frac{1}{n}\sum_{i=1}^{n}(x_{ni}^2 - x_{n,i-1}^2)$$
$$= (1 - \mathrm{e}^{2\beta/n})\frac{1}{n}\sum_{i=1}^{n} x_{ni}^2 + \mathrm{e}^{2\beta/n}\frac{1}{n}\sum_{i=1}^{n} u_i^2 + 2\mathrm{e}^{\beta/n}\frac{1}{n}\sum_{i=1}^{n} x_{n,i-1}u_i. \quad (10.26)$$

Consider the terms of this equality. $X_n(1)^2 \to_{\mathrm{d}} \sigma_u^2 J_\beta(1)^2$ by Theorem **10.3** and the continuous mapping theorem. Applying a standard manipulation from unit root theory,[2] Theorem **10.3** also gives

$$\frac{1}{n^2}\sum_{i=1}^{n} x_{ni}^2 = \frac{1}{n}\sum_{i=1}^{n} X_n^2(i/n) = \sum_{i=1}^{n}\int_{(i-1)/n}^{i/n} X_n^2(i/n)\mathrm{d}t \xrightarrow{\mathrm{d}} \sigma_u^2\int_0^1 J_\beta^2(t)\mathrm{d}t.$$

Next, with W_n defined in (10.14), write $\Delta W_n(i/n) = W_n(i/n+1/n) - W_n(i/n) = n^{-1/2}u_{i+1}$ so that

$$\frac{1}{n}\sum_{i=1}^{n} x_{n,i-1}u_i = \sum_{i=0}^{n-1} X_n(i/n)\Delta W_n(i/n) \xrightarrow{\mathrm{d}} \sigma_u\int_0^1 J_\beta(t)\mathrm{d}B(t)$$

where the indicated limit is an Itô integral with respect to Brownian motion, having zero mean.[3] Finally, applying the law of large numbers to the sequence u_i^2 and noting $\mathrm{e}^{\beta/n} \to 1$ and $n(1-\mathrm{e}^{2\beta/n}) \to -2\beta$, after dividing by σ_u^2 and rearrangement, as $n \to \infty$ the almost sure limiting form of the relation in (10.26) is

$$\int_0^1 J_\beta(t)\mathrm{d}B(t) = \tfrac{1}{2}\big(J_\beta(1)^2 - 1\big) + \beta\int_0^1 J_\beta(t)^2\mathrm{d}t. \quad (10.27)$$

In one sense, this expression might be viewed as generalizing to $\beta \neq 0$ the well-known property of Brownian motions, that $\int_0^1 B\mathrm{d}B = \frac{1}{2}(B(1)^2 - 1)$ with probability 1. However, the true generalization has the form

$$\int_0^1 J_\beta \mathrm{d}J_\beta = \tfrac{1}{2}(J_\beta(1)^2 - 1). \quad (10.28)$$

To verify this claim consider the increment of (10.3),

$$\Delta X_n(t) = \frac{\mathrm{e}^{-\beta([nt]+1)/n}}{n^{1/2}}\sum_{i=1}^{[nt]+1}\mathrm{e}^{\beta i/n}u_i - \frac{\mathrm{e}^{-\beta[nt]/n}}{n^{1/2}}\sum_{i=1}^{[nt]}\mathrm{e}^{\beta i/n}u_i$$
$$= \frac{u_{[nt]+1}}{n^{1/2}} + (\mathrm{e}^{-\beta/n} - 1)\frac{\mathrm{e}^{-\beta[nt]/n}}{n^{1/2}}\sum_{i=1}^{[nt]}\mathrm{e}^{\beta i/n}u_i$$

[2] Compare SLT Theorem 32.2.

[3] Section 32.2 of SLT gives technical details on this convergence.

$$= \Delta W_n(t) - \beta \frac{1}{n} X_n(t) + O(n^{-2}). \tag{10.29}$$

Multiply (10.29) by $X_n(t)$ to give

$$X_n(t)\Delta X_n(t) = X_n(t)\Delta W_n(t) - \beta \frac{1}{n} X_n^2(t) + O(n^{-2})$$

and then integrate with respect to t over $[0, 1]$. By arguments of the kind detailed above, the normalized limiting relation as $n \to \infty$ has the form

$$\int_0^1 J_\beta(t)\mathrm{d}J_\beta(t) = \int_0^1 J_\beta(t)\mathrm{d}B(t) - \beta \int_0^1 J_\beta^2(t)\mathrm{d}t \tag{10.30}$$

where the right-hand side of the equation can be taken as defining the left-hand side. This calculation shows that (10.30) is not an Itô integral and has a nonzero mean whose sign depends on β. That this is negative when $\beta > 0$ is a further indication of the mean-reversion tendency of the Ornstein-Uhlenbeck process.

10.5 Autocorrelated Shocks

This chapter has so far invoked Assumption **1.1** on the usual grounds of clarity and simplicity, but it would not be difficult to extend the weak convergence argument to Assumption **1.2**, using Lemma **3.14** as the model to modify Theorem **10.3**. The conditions of Assumption **3.9** are easily verified for the c_{ni} in (10.19).

However, extending Theorem **10.1** to autocorrelated shock processes is complicated in much the same way that going from Corollary **2.8** to Theorem **2.10** proved non-trivial. The next result extends Theorem **10.1** to weakly dependent shocks, although for the case of the variance only. The counterpart of (10.4) with the replacement of σ_u^2 by ω_u^2 can be obtained in the same way, with some further notational overhead that the reader can supply as desired.

10.4 Theorem If Assumption **1.2** holds for u_i in (10.3) then

$$\lim_{n\to\infty} \mathrm{E}(X_n(t)^2) = \omega_u^2 \frac{1 - \mathrm{e}^{-2\beta t}}{2\beta},\ t \geq 0.$$

Proof Choose a monotone integer sequence $\{B_n \in \mathbb{N}\}$ such that $B_n \to \infty$ but $B_n/n \to 0$ as $n \to \infty$, and put $r_n = [nt/B_n]$ for $t \in (0, 1]$ whenever n is large enough that $r_n \geq 1$. Assume for convenience that t takes a value such that $[nt] = r_n B_n$. This is harmless since by taking n large enough, every t can be made arbitrarily close to a value obeying the restriction.

Then, (10.3) can be written as

$$\mathrm{e}^{\beta[nt]/n} X_n(t) = \frac{[nt]^{1/2}}{n^{1/2}} \frac{1}{r_n^{1/2}} \sum_{j=1}^{r_n} \left(\frac{1}{B_n^{1/2}} \sum_{i=(j-1)B_n+1}^{jB_n} \mathrm{e}^{\beta i/n} u_i \right). \tag{10.31}$$

Define $a_n^* = \mathrm{e}^{-\beta B_n/n} - 1$ and

$$a_{nji} = \frac{\mathrm{e}^{\beta i/n} - \mathrm{e}^{\beta j B_n/n}}{\mathrm{e}^{\beta j B_n/n}} = \mathrm{e}^{\beta(i - jB_n)/n} - 1.$$

The r_n bracketed terms in (10.31) can then be decomposed as

$$\frac{1}{B_n^{1/2}} \sum_{i=(j-1)B_n+1}^{jB_n} \mathrm{e}^{\beta i/n} u_i = \mathrm{e}^{\beta j B_n/n} S_{nj} + \mathrm{e}^{\beta j B_n/n} a_n^* S_{nj}^* \tag{10.32}$$

where

$$S_{nj} = \frac{1}{B_n^{1/2}} \sum_{i=(j-1)B_n+1}^{jB_n} u_i$$

and

$$S_{nj}^* = \frac{1}{B_n^{1/2}} \sum_{i=(j-1)B_n+1}^{jB_n} \frac{a_{nji}}{a_n^*} u_i.$$

The signs of a_{nji} and a_n^* depend on that of β, but always match since $i \le jB_n$, so their ratio is positive in every case and $0 \le a_{nji}/a_n^* \le 1$.

With these definitions, consider $\mathrm{e}^{2\beta[nt]/n}\mathrm{E}(X_n^2(t))$. Multiplying out the square of (10.31) after substituting (10.32), there are three types of summand: those involving squares and products of the S_{nj} and those involving squares and products of the S_{nj}^*, in each case contributing r_n^2 terms; and those involving products of S_{nj}^* with S_{nj} contributing $2r_n^2$ terms).

The sum of the first type has the form

$$\mathrm{E}\left(\frac{[nt]^{1/2}}{n^{1/2}} \frac{1}{r_n} \sum_{j=1}^{r_n} \mathrm{e}^{\beta j B_n/n} S_{nj}\right)^2 = \frac{[nt]}{n}(T_{1n} + 2T_{2n}) \tag{10.33}$$

where

$$T_{1n} = \frac{1}{r_n} \sum_{j=1}^{r_n} \mathrm{e}^{2\beta j B_n/n} \mathrm{E}(S_{nj}^2)$$

and

$$T_{2n} = \frac{1}{r_n} \sum_{j=2}^{r_n} \mathrm{e}^{\beta j B_n/n} \sum_{m=1}^{j-1} \mathrm{e}^{\beta(j-m)B_n/n} \mathrm{E}(S_{nj} S_{n,j-m}).$$

Since $\mathrm{E}(S_{nj}^2) \to \omega_u^2$ as $n \to \infty$ for all j and $r_n B_n/n \to t$,

$$T_{1n} \to \omega_u^2 \int_0^1 \mathrm{e}^{2\beta t x} \mathrm{d}x = \omega_u^2 \frac{\mathrm{e}^{2\beta t} - 1}{2\beta t}. \tag{10.34}$$

Also, $T_{2n} = O(n^{-\delta})$ in view of the fact that $|\mathrm{E}(S_{nj} S_{n,j-m})| = O(m^{-1-\delta} B_n^{-\delta})$ for all j under Assumption **1.2**, according to the calculation in (2.60).

Considering the terms of the second type, $\mathrm{E}(S_{nj}^{*2}) = O(1)$, since this is true of $\mathrm{E}(S_{nj}^2)$. These are quadratic forms with respect to the same matrix of expected products, with weights that are unity in the latter case and drawn from the interval $[0, 1]$ in the former case. Since $|a_n^*| = |\mathrm{e}^{-\beta/r_n} - 1| = O(r_n^{-1})$, it follows that

$$\frac{1}{r_n}\sum_{j=1}^{r_n} \mathrm{e}^{2\beta j B_n/n} a_n^{*2}\mathrm{E}(S_{nj}^{*2}) = O(r_n^{-2}).$$

The contribution of the cross-product terms is also of small order by the same reasoning as for the S_{nj}. Finally, $|\mathrm{E}(S_{nj}S_{nj}^*)| = O(1)$ by similar reasoning so the contemporaneous products of the third type are bounded by

$$\frac{1}{r_n}\sum_{j=1}^{r_n} a_n^*|\mathrm{E}(S_{nj}S_{nj}^*)| = O(r_n^{-1})$$

with a similar conclusion for the cross-products. All the components involving the S_{ni}^* therefore vanish asymptotically. The theorem follows by (10.33), (10.34) and (10.31). ∎

A

Appendix: Useful Results

This appendix proves some results that do not relate explicitly to the fractional model but turn out to be useful, in particular for establishing various steps in the proof of the fractional FCLT. First, here are two well-known identities relating to a random variable X and constants $a > 0$ and $x > 0$.[1]

A.1 Theorem $\mathrm{E}(|X|^a 1_{\{|X| \le x\}}) = a\int_0^x \xi^{a-1} P(|X| > \xi)\mathrm{d}\xi - x^a P(|X| > x)$. $\square$

A.2 Corollary $\mathrm{E}(|X|^a) = a\int_0^\infty \xi^{a-1} P(|X| > \xi)\mathrm{d}\xi$. $\square$

The first follows easily, applying integration by parts after writing $P(|X| > x) = 1 - F(x)$ where F is the c.d.f. of X. The second one shows the case $x = \infty$ making use of $x^a = a\int_0^x \xi^{a-1}\mathrm{d}\xi$.

A.3 Lemma For a random variable X, L_r-boundedness implies

$$P(|X| > \eta) = O(\eta^{-r}\log(\eta)^{-1-\mu}) \tag{A.1}$$

for $\mu > 0$.

Proof Corollary **A.2** with $a = 1$ gives the equality $\mathrm{E}(|X|) = \int_0^\infty P(|X| > \xi)\mathrm{d}\xi$. Applied to the case $|X|^r$, this means that

$$\mathrm{E}(|X|^r) = \int_0^\infty P(|X|^r > \xi)\mathrm{d}\xi. \tag{A.2}$$

L_r-boundedness implies integrability on the right hand side so that

$$P(|X|^r > \xi) = O(\xi^{-1}\log(\xi)^{-1-\mu})$$

as $\xi \to \infty$, from which (A.1) follows on setting $\eta = \xi^{1/r}$. ∎

[1] These are proved in SLT as Theorem 9.21 and Corollary 9.22, respectively.

A.4 Theorem If a random variable X is L_r-bounded for $r \geq 2$ then as $\eta \to \infty$,

$$\mathrm{E}(X^2 1_{\{|X|>\eta\}}) = o(\eta^{2-r}). \quad \square \tag{A.3}$$

Proof Begin by writing

$$\begin{aligned}\mathrm{E}(X^2 1_{\{|X|>\eta\}}) &= \mathrm{E}(X^2) - \mathrm{E}(X^2 1_{\{|X|\leq\eta\}}) \\ &= 2\int_\eta^\infty \xi P(|X| > \xi)\mathrm{d}\xi + \eta^2 P(|X| > \eta)\end{aligned} \tag{A.4}$$

where the second equality is got by applying Corollary **A.2** and Theorem **A.1** with $a = 2$.

For clarity the cases $r > 2$ and $r = 2$ are treated separately. In the first case, Lemma **A.3** implies that $\eta^r P(|X| > \eta) = O(\log(\eta)^{-1-\mu})$ as $\eta \to \infty$. Hence, when η is large enough, $\xi > \eta$ implies

$$\xi^r P(|X| > \xi) \leq \eta^r P(|X| > \eta). \tag{A.5}$$

Multiplying through (A.5) by ξ^{1-r} and integrating both sides from η to ∞ produces the inequality

$$\begin{aligned}\int_\eta^\infty \xi P(|X| > \xi)\mathrm{d}\xi &\leq \eta^r P(|X| > \eta) \int_\eta^\infty \xi^{1-r}\mathrm{d}\xi \\ &= \frac{1}{r-2}\eta^2 P(|X| > \eta).\end{aligned} \tag{A.6}$$

Substitute (A.6) into (A.4) and then apply (A.1) to get (A.3).

For the case $r = 2$, first set $\nu = \mu/2$ so that Lemma **A.3** gives

$$\eta^2 \log(\eta)^{1+\nu} P(|X| > \eta) = O(\log(\eta)^{-\nu}). \tag{A.7}$$

When η is large enough, $\xi > \eta$ therefore implies

$$\xi^2 \log(\xi)^{1+\nu} P(|X| > \xi) \leq \eta^2 \log(\eta)^{1+\nu} P(|X| > \eta).$$

Multiply both sides of this inequality by $\xi^{-1}\log(\xi)^{-1-\nu}$ and integrate from η to ∞ to get

$$\int_\eta^\infty \xi P(|X| > \xi)\mathrm{d}\xi \leq \eta^2 \log(\eta)^{1+\nu} P(|X| > \eta) \int_\eta^\infty \xi^{-1}\log(\xi)^{-1-\nu}\mathrm{d}\xi. \tag{A.8}$$

To solve the right-hand-side integral, make the change of variable $u = \log(\xi)$ so that $\mathrm{d}u = \xi^{-1}d\xi$ and the integrand is $u^{-1-\nu}$. The solution therefore has the form

$$\int_\eta^\infty \xi^{-1}\log(\xi)^{-1-\nu}\mathrm{d}\xi = \frac{1}{-\nu}\left[\log(\xi)^{-\nu}\right]_\eta^\infty = \frac{1}{\nu\log(\eta)^\nu}. \tag{A.9}$$

Substituting from (A.8) into (A.4) gives, according to (A.7),

$$\mathrm{E}(X^2 1_{\{|X|>\eta\}}) \leq \eta^2\Big(\frac{2\log(\eta)}{\nu} + 1\Big)P(|X| > \eta)$$

$$= O(\log(\eta)^{-\mu}) = o(1)$$

as $\eta \to \infty$ which matches (A.3) for this case. ∎

The following is a leading implication of **A.4**.

A.5 Theorem Let $\{u_i,\ -\infty < i < \infty\}$ be independently distributed and uniformly L_r-bounded for $r \geq 2$, and let $S_k = \sum_{j=1}^{k} c_{nj} u_j$ where $\{c_{nj}\}$ is a triangular array of constants and $\sum_{j=1}^{n} c_{nj}^2 = O(1)$ as $n \to \infty$. The collection $\{\max_{1 \leq k \leq n} S_k^2, n \in \mathbb{N}\}$ is uniformly integrable with

$$\mathrm{E}\Big(\max_{1\leq k\leq n} S_k^2 1_{\{\max_{1\leq k\leq n}|S_k|>\eta\}}\Big) = o(\eta^{2-r}) \tag{A.10}$$

as $\eta \to \infty$. □

This result strengthens the simple demonstration of uniform integrability by specifying the rate of convergence, which is important for certain applications. As often as not it is applied in cases where the sum bounds $1, \ldots, n$ are replaced by integer functions of n, such as $[nt]+1, \ldots, [ns]$ for any pair of real numbers s and t with $s > t$. It is easily confirmed that the argument is just as valid for such choices as for the baseline case of $t = 0$ and $s = 1$.

Proof of A.5 For finite constants $K_1 > 0$, $K_2 > 0$, and $K_3 \geq K_2 \sum_{j=1}^{n} c_{nj}^2$ where $K_3 < \infty$ exists by assumption,

$$\begin{aligned}\mathrm{E}\Big(\max_{1\leq k\leq n}|S_k|^r\Big) &\leq K_1 \mathrm{E}|S_n|^r \leq K_2 \mathrm{E}\Bigg(\sum_{j=1}^{n} c_{nj}^2 u_j^2\Bigg)^{r/2} \\ &\leq K_2 \sum_{j=1}^{n} c_{nj}^2 \mathrm{E}|u_j|^r \leq K_3 \max_{1\leq j\leq n} \mathrm{E}|u_j|^r \end{aligned} \tag{A.11}$$

In (A.11) the first inequality is the Doob inequality,[2] the second the Burkholder inequality[3] and the third is by convexity. Noting that

$$\max_{1\leq k\leq n}|S_k|^r = \Big(\max_{1\leq k\leq n}|S_k|\Big)^r$$

and that the random variable $\max_{1\leq k\leq n}|S_k|$ is L_r-bounded by (A.11) and the assumption on $\{u_j\}$, it follows by Theorem **A.4** that (A.10) holds as $\eta \to \infty$. Given the assumptions, this conclusion does not depend on n and so holds for $n \in \mathbb{N}$. ∎

This result holds equally in the case where $\{u_i\}$ is a martingale difference, although that extension is not used in the present applications. Both the Doob and

[2] See SLT Theorem 16.21.

[3] See SLT Theorem 16.24.

Burkholder inequalities are martingale results, therefore holding under independence, although the somewhat simpler Marcinkiewicz-Zygmund inequality[4] can substitute for Burkholder's in that case.

Finally, some simple but useful facts are stated formally so that they may be cited at different points in the development.

A.6 Theorem If U and V are L_2-bounded random variables and $|U| \leq |V|$ a.s., then for a constant $M > 0$,

$$\mathrm{E}(U^2 1_{\{|U|>M\}}) \leq \mathrm{E}(V^2 1_{\{|V|>M\}}).$$

Proof The assumption implies $U^2 1_{\{|U|>M\}} \leq V^2 1_{\{|V|>M\}}$ a.s.. For any pair of integrable random variables X and Y, $X \leq Y$ a.s. implies that $\mathrm{E}(X) \leq \mathrm{E}(Y)$.[5] ∎

A.7 Theorem If X and Y are L_2-bounded random variables,

$$\mathrm{E}((X+Y)^2 1_{\{|X+Y|>M\}}) \leq 2\mathrm{E}(X^2 1_{\{|X|>M/2\}}) + 2\mathrm{E}(Y^2 1_{\{|Y|>M/2\}})$$

for $M > 0$.

Proof For nonnegative L_2-bounded random variables U and V,

$$\begin{aligned}
&\mathrm{E}\big((U+V)^2 1_{\{U+V>M\}}\big) \\
&\qquad = \mathrm{E}\big((U+V)^2 1_{\{U+V>M\}\cap\{U\geq V\}}\big) + \mathrm{E}\big((U+V)^2 1_{\{U+V>M\}\cap\{U<V\}}\big) \\
&\qquad \leq \mathrm{E}\big(2U^2 1_{\{U+V>M\}\cap\{U\geq V\}}\big) + \mathrm{E}\big(2V^2 1_{\{U+V>M\}\cap\{U<V\}}\big) \\
&\qquad \leq 2\mathrm{E}\big(U^2 1_{\{U>M/2\}}\big) + 2\mathrm{E}\big(V^2 1_{\{V>M/2\}}\big).
\end{aligned} \tag{A.12}$$

Setting $U = |X|$ and $V = |Y|$, $|X+Y| \leq |X| + |Y|$ by the triangle inequality and

$$\mathrm{E}((X+Y)^2 1_{\{|X+Y|>M\}}) \leq \mathrm{E}((|X|+|Y|)^2 1_{\{|X|+|Y|>M\}})$$

by Theorem **A.6**, which in conjunction with (A.12) completes the proof. ∎

A.8 Theorem Let the function $L : [0,\infty) \mapsto \mathbb{R}$ be slowly varying.

(i) If $\rho > -1$,

$$\int_0^v y^\rho L(y)\mathrm{d}y \sim \frac{v^{1+\rho}L(v)}{1+\rho} \text{ as } v \to \infty$$

(ii) If $\rho < -1$,

$$\int_v^\infty y^\rho L(y)\mathrm{d}y \sim -\frac{v^{1+\rho}L(v)}{1+\rho} \text{ as } v \to \infty.$$

---

[4] For example, see [30] Theorem 8.1.

[5] Shown by, for example, SLT Lemma 4.11.

Proof For (i), make the change of variable $y = sv$ where $0 \leq s \leq 1$ and write

$$\int_0^v y^\rho L(y)\mathrm{d}y = v \int_0^1 (sv)^\rho L(sv)\mathrm{d}s = v^{1+\rho} L(v) \int_0^1 s^\rho \frac{L(sv)}{L(v)}\mathrm{d}s.$$

For (ii), make the change of variable $y = sv$ where $1 \leq s < \infty$ and write

$$\int_v^\infty y^\rho L(y)\mathrm{d}y = v \int_1^\infty (sv)^\rho L(sv)\mathrm{d}s = v^{1+\rho} L(v) \int_1^\infty s^\rho \frac{L(sv)}{L(v)}\mathrm{d}s.$$

In each case, the theorem follows since $L(sv)/L(v) \to 1$ as $v \to \infty$. ∎

In part (ii), the case $\rho = -1$ can have a solution depending on the form of L. Specifically, in equation (A.9), $L(y) = \log(y)^{-1-a}$ for $a > 0$. The integral has exact solution $1/(a \log(v)^a)$.

B

Appendix: Identities and Integral Solutions

$$\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i}\sin x \tag{B.1}$$

$$\cos x = \frac{\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x}}{2} \tag{B.2}$$

$$\sin x = \frac{\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x}}{2\mathrm{i}} \tag{B.3}$$

$$\cos(x \pm \pi) = -\cos x \tag{B.4}$$

$$\sin(x \pm \pi) = -\sin x \tag{B.5}$$

$$\cos\left(\frac{\pi}{2} - x\right) = \sin x \tag{B.6}$$

$$\sin\left(\frac{\pi}{2} - x\right) = \cos x \tag{B.7}$$

$$\sin(2x) = 2\cos x \sin x \tag{B.8}$$

$$\sin^2 x = \frac{1 - \cos(2x)}{2} \tag{B.9}$$

$$\sin x + \sin y = 2\cos\left(\frac{x-y}{2}\right)\sin\left(\frac{x+y}{2}\right) \tag{B.10}$$

$$\cos(x-y) - \cos(x+y) = 2\sin x \sin y \tag{B.11}$$

$$\Gamma(x) = \int_0^\infty t^{x-1}\mathrm{e}^{-t}\mathrm{d}t \quad (x > 0) \tag{B.12}$$

$$\Gamma(x+1) = x\Gamma(x) \;\; (x \neq 0, -1, -2, \ldots) \tag{B.13}$$

$$B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}\quad (x>0,\ y>0) \tag{B.14}$$

$$\Gamma(x)\Gamma(1-x)\sin(\pi x)=\pi\quad (x\notin\mathbb{Z}) \tag{B.15}$$

$$\text{Stirling's approximation: } \log(n!)=n\log n-n+O(\log n) \tag{B.16}$$

$$\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}j\lambda}\mathrm{d}\lambda=\begin{cases}2\pi & j=0\\ 0, & j\neq 0,\ \text{integer}\end{cases} \tag{B.17}$$

$$\int_0^{\pi}\cos kx\cos jx\mathrm{d}x=\begin{cases}\pi/2 & j=k\\ 0 & j\neq k\end{cases},\ j,k\text{ integers} \tag{B.18}$$

$$\int_0^{\pi}\sin^{\nu-1}x\cos ax\mathrm{d}x=\frac{\pi}{2^{\nu-1}\nu B\Big(\frac{\nu+a+1}{2},\frac{\nu-a+1}{2}\Big)}\cos\Big(\frac{a\pi}{2}\Big) \tag{B.19}$$

$(\nu>0)$ [27], eqn. 3.631.8.

$$\int_0^{\infty}x^{\mu-1}\sin(ax)\mathrm{d}x=\frac{\Gamma(\mu)}{a^{\mu}}\sin\Big(\frac{\mu\pi}{2}\Big) \tag{B.20}$$

$(a>0,\ 0<|\mu|<1)$ [27], eqn. 3.761.4.

$$\int_0^{\infty}x^{\mu-1}\sin^2(ax)\mathrm{d}x=-\frac{\Gamma(\mu)}{2^{\mu+1}a^{\mu}}\cos\Big(\frac{\mu\pi}{2}\Big) \tag{B.21}$$

$(a>0,\ -2<\mu<0)$ [27], eqn. 3.823.

$$F(a,b;c;z)=\frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{j=0}^{\infty}\frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(c+j)\Gamma(j+1)}z^j \tag{B.22}$$

$(|z|\leq 1, c\neq 0,-1,-2,\ldots)$ [1], eqn. 15.1.1.

$$F(a,b;c;1)=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \tag{B.23}$$

$(c-a-b>0)$ [1], eqn. 15.1.20.

$$F(a,b;c;z)=\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1\tau^{b-1}(1-\tau)^{c-b-1}(1-\tau z)^{-a}d\tau \tag{B.24}$$

$(c>b>0,\ |z|\leq 1)$ [1], eqn 15.3.1.

Bibliography

[1] Abramowitz, M. and I. A. Stegun (1965) *Handbook of Mathematical Functions*, Dover, New York.

[2] Abadir, K., W. Distaso, L. Giraitis and H. L. Koul (2014) Asymptotic normality for weighted sums of linear processes, *Econometric Theory* 30, 252–284.

[3] Banerjee, A., J. J. Dolado, J. W. Galbraith, and D. F. Hendry (1993) *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data,* Advanced Texts in Econometrics. Oxford, Oxford University Press.

[4] Beran, J. (1994) *Statistics for Long Memory Processes*. New York: Chapman and Hall.

[5] Beran, J. Y. Feng, S. Ghosh and R. Kulik (2013) *Long-Memory Processes, Probabilistic Properties and Statistical Methods* Springer, Heidelberg.

[6] Billingsley, P. (1968) *Convergence of Probability Measures*, John Wiley and Sons, New York.

[7] Billingsley, P. (1999) *Convergence of Probability Measures* (2nd Edn.), John Wiley and Sons, New York.

[8] Bondon, P. and W. Palma (2007). A class of antipersistent processes. *Journal of Time Series Analysis* 28, 261–273.

[9] Box, G. E. P., G. M. Jenkins, G. C. Reinsel and G. M. Ljung (2016) *Time Series Analysis: Forecasting and Control* (5th Edn.) John Wiley & Sons, Hoboken NJ.

[10] Brockwell, P. J. and R. A. Davis (1991) *Time Series: Theory and Methods* 2nd Edn, Springer-Verlag, New York.

[11] Chan, N. H. and N. Terrin (1995) Inference for unstable long memory processes with applications to fractional unit root autoregressions. *Annals of Statistics* 23, 1662–1683.

[12] Chan, N.H. and C.-Z. Wei (1987) Asymptotic inference for nearly nonstationary AR(1) processes, *Annals of Statistics* 15, 1050–1063.

[13] Davidson, J. (2000) *Econometric Theory*, Blackwell Publishers, Oxford.

[14] Davidson, J. (2021) *Stochastic Limit Theory: An Introduction for Econometricians*, 2nd Edn. Oxford University Press, Oxford.

[15] Davidson, J. and R. M. de Jong (2000) The functional central limit theorem and convergence to stochastic integrals II: fractionally integrated processes. *Econometric Theory* 16 (5), 643–666.

[16] Davidson, J. (2004) Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business and Economics Statistics* 22 (1), 16–29.

[17] Davidson, J. and N. Hashimzade (2008) Alternative Frequency and Time Domain Versions of Fractional Brownian Motion, *Econometric Theory* 24(1), 256–293.

[18] Davidson, J. and N. Hashimzade (2009a) Type I and type II fractional Brownian motions: a reconsideration *Computational Statistics and Data Analysis* 53(6) (2009) 2089–2106.

[19] Davidson, J. and N. Hashimzade (2009b) Representation and Weak Convergence of Stochastic Integrals with Fractional Integrator Processes", *Econometric Theory* 25 (6), 1589–1624.

[20] Davydov, Yu. A. (1970) The invariance principle for stationary processes *Theory of Probability and its Applications* 15(3), 487–498.

[21] De Jong, R. M. and Davidson, J. (2000) The functional central limit theorem and convergence to stochastic integrals I: weakly dependent processes. *Econometric Theory* 16 (5), 621–642.

[22] Doukhan, P., G. Oppenheim, and M. Taqqu (eds.) (2003) *Theory and applications of long-range dependence.* Birkhauser, Boston.

[23] Fishman, G. S. (1969) *Spectral Methods in Econometrics* Harvard University Press, Cambridge Mass.

[24] Geweke, J. and S. Porter-Hudak, 1983, The estimation and application of long memory time series models, *Journal of Time Series Analysis* 4, 221–238.

[25] Giraitis, L., H. L. Koul, D. Surgailis (2012) *Large Sample Inference for Long Memory Processes*, Imperial College Press, London.

[26] Gorodetskii, V. V. (1977) On convergence to semi-stable Gaussian processes, *Theory of Probability and its Applications* 22(3), 498–508.

[27] Gradsteyn, I. S. and I. M. Ryzhik (2007) *Table of Integrals, Series and Products* (7th Edn.) Academic Press, Burlington MA.

[28] Granger, C. W. J. (1980) ‘Long memory relationships and the aggregation of dynamic models’, *Journal of Econometrics* 14, 227–238.

[29] Granger, C. W. J. and R. Joyeux (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis* 1(1) 15–29.

[30] Gut, A. (2013), *Probability: a Graduate Course* (2nd Edn.), Springer, New York.

[31] Hassler, U. (2019) *Time Series Analysis with Long Memory in View*, John Wiley & Sons, Hoboken NJ.

[32] Hurvich, C. M and K. I. Beltrao (1994) Automatic semiparametric estimation of the memory parameter of a long memory time series, *Journal of Time Series Analysis* 15, 285–302.

[33] Hosking, J. R. M. (1981) Fractional differencing. *Biometrika* 68,1, 165–76.

[34] Hosking, J. R. M. (1984a) Modelling persistence in hydrological time series using fractional differencing, *Water Resources Research* 20(12) 1898–1908

[35] Hosking, J. R. M. (1984b) Asymptotic distributions of the sample mean, autocovariances and autocorrelations of long memory time series. Technical summary report 2752, Mathematics Research Center, University of Wisconsin, Madison WI

[36] Hosking, J. R. M. (1996) Asymptotic distributions of the sample mean, autocovariances and autocorrelations of long memory time series, *Journal of Econometrics* 73, 261-284.

[37] Hosoya, Y. (2003) Fractional invariance principle, *Journal of Time Series Analysis* 26(3) 463–486.

[38] Hurst, H. E. (1951) Long term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers* 116, 770–779.

[39] Johansen S. and M. Ø. Nielsen (2011) A necessary moment condition for the fractional functional central limit theorem, *Econometric Theory* 28(3) 671–679.

[40] Lamperti, J. (1962) Semi-stable stochastic processes. *Transactions of the American Mathematical Society* 104(1), 62–78.

[41] Liu, M. (1998) Asymptotics of nonstationary fractional integrated series, *Econometric Theory* 14, 641–662.

[42] Lo, A. W. (1991) Long-term memory in stock market prices, *Econometrica* 59(5) 1279–1313.

[43] Mandelbrot, B. B. and J. R. Wallis (1968) Noah, Joseph and operational hydrology *Water Resources Research* 4(5) 909–918.

[44] Mandelbrot, B. B. and J. W. van Ness 1968. Fractional Brownian motions, fractional noises and applications. SIAM Review 10, 4, 422–437.

[45] Marinucci, D. and P. M. Robinson, (1999) Weak convergence to fractional Brownian motion. *Stochastic Processes and their Applications* 80, pp.103–120.

[46] Marinucci, D. and P. M. Robinson (1999) Alternative forms of fractional Brownian motion. *Journal of Statistical Inference and Planning* 80, 111–122.

[47] Mason. D. M. (2016) The Hurst phenomenon and the rescaled range statistic, *Stochastic Processes and their Applications* 126, 3790–3807.

[48] Moulines, E. and P. Soulier (1999) Broad band log-periodogram estimation of time series with long-range dependence. *Annals of Statistics* 27, 1415–1439.

[49] Odaki, M. (1993) On the invertibility of fractionally differenced ARIMA processes. *Biometrika* 80(3) 703–709.

[50] Palma, W. (2007) *Long Memory Time Series: Theory and Methods*, John Wiley & Sons, Hoboken NJ.

[51] Palma, W. (2016) *Time Series Analysis,* John Wiley & Sons, Hoboken NJ.

[52] Pipiras, V. and M. Taqqu (2017) *Long-Range Dependence and Self-Similarity,* Cambridge University Press, Cambridge.

[53] Phillips, P. C. B. (1987) Towards a unified asymptotic theory for autoregression, *Biometrika* 74, 535–547.

[54] Phillips, P. C. B. (1988) Regression theory for near-integrated time series, *Econometrica* 56, 1021–1043.

[55] Phillips, P. C. B. (2023) Estimation and inference with near unit roots, *Econometric Theory* 39, 221–263.

[56] Phillips, P. C. B. and B. E. Hansen (1990) Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

[57] Phillips, P. C. B. and C. S. Kim (2007) Long run covariance matrices for fractionally integrated processes, *Econometric Theory* 23, 1233–1247.

[58] Phillips, P. C. B. and T. Magdalinos (2007) Limit theory for moderate deviations from a unit root, *Journal of Econometrics* 136, 115–130.

[59] Robinson, P. M. (1994) Time series with strong dependence. In *Advances in Econometrics, Econometric Society 6th World Congress. Vol. 1* (Sims, ed.) Cambridge University Press.

[60] Robinson, P. M. (1994) Semiparametric analysis of long memory time series. *Annals of Statistics* 22(1), 515–539.

[61] Rozanov, Yu. A. (1967) *Stationary Random Processes*, Holden Day, San Francisco.

[62] Saikkonen, P. (1991) Asymptotically efficient estimation of cointegration regressions. *Econometric Theory* 7, 1–21.

[63] Samorodnitsky, G. (2016) *Stochastic Processes and Long Range Dependence*, Springer International, Switzerland.

[64] Skorokhod, A. V. (1956), Limit theorems for stochastic processes, *Theory of Probability and its Applications* 1, 261–90.

[65] Sowell, F. (1990) The fractional unit root distribution. *Econometrica* 58 (2) 495–505.

[66] Sowell, F. (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 53, 165–188.

[67] Sutcliffe, J., S. Hurst, A. G. Awadallah, E. Brown and K. Hamed (2016) Harold Edwin Hurst: the Nile and Egypt, past and future, *Hydrological Sciences Journal* 61(9) 1557–1570.

[68] Taqqu, M. S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichskeitstheorie Verw. Geb.* 31, 287–302.

[69] Taqqu, M. S. (2003) Fractional Brownian motion and long-range dependence, Chapter 1 of [22].

[70] Tousson, O. (1925) Memoire sur l'Histoire du Nil, *Memoires de l'Institut d'Egypte,* Vol. 8, Imprimeries de l'Institut Francais, Cairo.

[71] Wang, Q., Y-X. Lin and C. M. Gulati (2003) Asymptotics for general fractionally integrated processes with applications to unit root tests, *Econometric Theory* 19, 143–164.

[72] Wu, W. B. and Shao, X. (2006). Invariance principles for fractionally integrated non-linear processes. *IMS Lecture Notes–Monograph Series: Recent Developments in Non-parametric Inference and Probability* 50, 20–30.

[73] Yaglom, A. M. (1962) *Introduction to the Theory of Stationary Random Functions* Prentice-Hall, Englewood Cliffs, NJ.

Index