NONLINEARITY AND TEMPORAL DEPENDENCE

By

Xiaohong Chen, Lars P. Hansen and Marine Carrasco

May 2008
Revised October 2009

COWLES FOUNDATION DISCUSSION PAPER NO. 1652R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Nonlinearity and Temporal Dependence *

Xiaohong Chen[†] Yale University
Lars Peter Hansen University of Chicago
Marine Carrasco University of Montreal

October 4, 2009

Abstract

Nonlinearities in the drift and diffusion coefficients influence temporal dependence in diffusion models. We study this link using three measures of temporal dependence: $\rho-mixing$, $\beta-mixing$ and $\alpha-mixing$. Stationary diffusions that are $\rho-mixing$ have mixing coefficients that decay exponentially to zero. When they fail to be $\rho-mixing$, they are still $\beta-mixing$ and $\alpha-mixing$; but coefficient decay is slower than exponential. For such processes we find transformations of the Markov states that have finite variances but infinite spectral densities at frequency zero. The resulting spectral densities behave like those of stochastic processes with long memory. Finally we show how state-dependent, Poisson sampling alters the temporal dependence.

JEL classification: C12; C13; C22; C50.

Keywords: diffusion; strong dependence; long memory; poisson sampling; quadratic forms.

*Forthcoming: *Journal of Econometrics*. First version: May 1999. We thank a co-editor, Peter Robinson, an associate editor and anonymous referees for valuable suggestions which greatly improved the paper. We also thank Rabi Bhattacharya, Richard Bradley, Javier Hidalgo, Jose Scheinkman, Harald Uhlig, Noah Williams and James Wolter for helpful conversations, Noah Williams for research assistance, participants of 2000 World Congress of the Econometric Society and 2002 Montreal CIRANO Conference on Multivariate Financial Econometrics for comments. Chen and Hansen gratefully acknowledge support from the National Science Foundation.

[†]Corresponding author: Department of Economics, Yale University, Box 208281, New Haven, CT 06520-8281, USA. Phone (203) 432-5852, Fax (203) 432-6167, xiaohong.chen@yale.edu

1 Introduction

This paper studies how nonlinearity induces temporal dependence in continuous-time Markov models. Our interest in temporal dependence stems from a variety of empirical evidence. Bond prices are known to be highly persistent and the conditional volatilities of financial time series are often temporally clustered, *e.g.*, see Ding et al. (1993), Bollerslev and Mikkelsen (1996), and Andersen et al. (2001). Using linear time series methods, many researchers have documented the presence of *long memory* in empirical descriptions of data. Here we investigate when high degrees of temporal dependence are present in first-order (and hence *finite memory*) Markov models. In these models the dependence emerges because of the nonlinearities in the evolution equation for the Markov state.

Since hedging and pricing theory in financial economics often assume securities follow continuous-time nonlinear diffusions, we use these models to capture nonlinearities in time series. For pedagogical and analytical convenience, we primarily treat the case of scalar diffusions. That is, we study the solution to a stochastic differential equation:

$$dx_t = \mu(x_t)dt + \sigma(x_t)dW_t$$

where μ is the drift coefficient, σ^2 is the diffusion coefficient and $\{W_t\}$ is a one-dimensional standard Brownian motion. While scalar diffusion processes are special, they provide a very nice laboratory for our analysis. They allow for nonlinearities in the time series to be captured fully by two functions: the drift and the diffusion coefficients. By focusing on diffusion models, we are able to show clearly the mechanism whereby nonlinearities in a time series get transmitted into temporal dependence. After studying scalar diffusions in some detail, we explore the implications of subordination whereby the scalar diffusion operates according to a stochastic time scale. We conclude with some multivariate extensions.

We classify the temporal dependence of scalar diffusions using three alternative notions of mixing: $\rho-mixing$, $\beta-mixing$ and $\alpha-mixing$. The $\rho-mixing$ coefficients measure the temporal decay of maximally autocorrelated (nonlinear) functions of the Markov state. When $\rho - mixing$ coefficients decay to zero, the spectral density of any process with finite second moments formed by taking a nonlinear function of the state has a continuous (and hence finite) spectral density at all frequencies, including frequency zero. Thus $\rho - mixing$ gives an operational way to classify the dependence of a nonlinear Markov process. The process is weakly dependent if the $\rho - mixing$ coefficients decay to zero and strongly dependent if the $\rho - mixing$ coefficients are identically one. In our study of strongly dependent processes, we also use $\beta - mixing$ and $\alpha - mixing$, which for our purposes are more refined concepts. The $\beta - mixing$ or $\alpha - mixing$ coefficients for the strongly dependent processes have a decay rate that is slower than exponential and the implied spectral density functions for some of these processes diverge at frequency zero. Since the $\beta - mixing$ coefficients dominate the corresponding $\alpha - mixing$ coefficients, implications for $\beta - mixing$ have direct consequences for $\alpha - mixing$.

For both $\rho-$ and $\beta-$ mixing we are lead to use:

$$\frac{\mu}{\sigma} - \frac{\sigma'}{2}$$

to measure the *pull* from the boundaries. When this pull measure is negative at the right boundary and positive at the left boundary, Hansen and Scheinkman (1995) have shown how the non-zero boundary pull restrictions imply $\rho{-}mixing$ with exponential decay. We extend their result by showing that in the case of scalar diffusions, $\beta - mixing$ and $\rho - mixing$ with exponential decay are almost equivalent concepts, and both are implied by the non-zero boundary pull restrictions.

When the pull measure is zero at one of the boundaries, the Markov process is strongly dependent: the $\rho{-}mixing$ coefficients are identically one and the $\beta{-}mixing$ coefficients decay at a rate slower than exponential. We provide sufficient conditions for $\beta - mixing$ coefficients to decay at a polynomial rate. These conditions are expressed as restrictions on how slow the pull measure goes to zero at one of the boundaries. For some of these strongly dependent processes we may find functions of the Markov state with spectral densities that are infinite at frequency zero. Thus in spite of the first-order Markov property, best linear predictors are compelled to weight heavily past observations. We display scalar diffusions whose spectral densities near frequency zero diverge in the same manner as stationary linear models that are fractionally integrated.[1]

We also provide some characterizations of $\rho{-}mixing$ and $\alpha{-}mixing$ coefficients as functions of the stationary density and diffusion coefficient. These characterizations allow us to relate the behavior of the mixing coefficients to the thickness of the tails of the stationary density and to the growth of the diffusion coefficient as a function of the Markov state. This analysis includes multivariate diffusion processes.

Our findings complement the work of Granger and Teräsvirta (1999), who produce a (discrete-time) nonlinear Markov model with sample statistics that suggest evidence for long memory. We construct Markov models for which population spectral densities diverge because of the forms of the nonlinearities in the conditional means and the conditional volatilities. Our construction differs from the nonlinear example in Granger and Teräsvirta (1999), since theirs is still weakly dependent as measured by the population attributes. It only looks like a long memory process from the vantage point of sample statistics.[2] In contrast, the first-order strictly stationary Markov examples we construct are strongly dependent in the population.

[1]While we provide simple sufficient conditions for $\beta - mixing$ decay rates in terms of the pull measures, there are alternative sufficient conditions for a diffusion process to be $\rho - mixing$ or $\beta - mixing$; see, e.g., Genon-Catalot et al. (2000) for $\rho - mixing$ and Veretennikov (1987) and Veretennikov (1997) for $\beta - mixing$. The sufficient conditions and proof strategies in these papers are different from ours. More importantly, they do not discuss the possibility of generating long memory type behavior from strongly dependent strictly stationary Markov diffusions.

[2]Similarly, Diebold and Inoue (2001) present several nonlinear time series models, including

We study how the temporal dependence of a diffusion is altered when it is sampled in a state dependent manner. For convenience, we use the Poisson sampling scheme advocated by Duffie and Glynn (2004) with a state dependent intensity parameter. The state dependence and hence endogeneity in the sampling alters the temporal dependence and the stationary distribution. We show how to adjust the measure of pull to take account of this endogeneity.

The rest of the paper is organized as follows. Section 2 briefly reviews alternative mixing concepts for a continuous-time stationary Markov process. Section 3 establishes that $\rho-mixing$ is equivalent to exponential ergodicity for a scalar diffusion, which in turn implies $\beta-mixing$ with exponential decay. Section 4 provides sufficient conditions for $\rho-,\beta-mixing$ with exponential decay rates expressed as boundary pull restrictions. Sections 5 and 6 study strongly dependent diffusion processes in the sense that $\rho_t \equiv 1$ for all $t \geq 0$, and $\beta-mixing$ with sub-exponential decay rates. While Section 5 provides sufficient conditions for $\beta-mixing$ with sub-exponential decay rates. Section 6 presents examples where some nonlinear transformations of the strongly dependent Markov diffusions behave as long memory processes. Section 7 considers the diffusions subject to a Poisson sampling. Section 8 explores some multivariate extensions obtained by featuring the behavior of the stationary density and the diffusion matrix. Section 9 gives some concluding remarks.

2 Review: mixing conditions

Consider a stationary Markov process $\{x_t\}$ on an open connected set $\Omega \subseteq \mathbb{R}^n$. For convenience, use a stationary distribution Q to initialize this process. Let L^p denote the space of Borel measurable functions that have finite $p-$th moments in accordance with the distribution Q:

$$\begin{aligned} L^p &= \left\{\phi : \Omega \to \mathbb{R} : ||\phi||_p = \left(\int |\phi|^p dQ\right)^{1/p} < \infty\right\},\ 1 \leq p < \infty, \\ L^\infty &= \left\{\phi : \Omega \to \mathbb{R} : ||\phi||_\infty = \operatorname{ess\,sup}_x |\phi(x)| = \inf_{c>0} \left(Q\left\{x : |\phi(x)| > c\right\} = 0\right) < \infty\right\} \end{aligned}$$

In L^2 we will use the familiar inner product: $\langle \phi, \psi \rangle = \int \phi\psi dQ$. Associated with the Markov process $\{x_t : t \geq 0\}$ is a semigroup of conditional expectation operators $\{\mathcal{T}_t : t \geq 0\}$ defined on L^2:

$$\mathcal{T}_t\phi(x) = E[\phi(x_t)|x_0 = x].$$

models with regime switching, structure changes, and permanent stochastic breaks, that look like long memory models from the vantage point of the sample variances of partial sums. Hidalgo and Robinson (1996) discuss the difficulty of distinguishing a structural break model from a long memory model.

For notational simplicity, we also let $\mathcal{T}$ be shorthand notation for $\mathcal{T}_1$. As in Hansen and Scheinkman (1995), we suppose that the semigroup is (right) continuous at $t = 0$. This allows us to construct a generator $\mathcal{A}$ on a domain D that is dense in L^2:

$$\mathcal{A}\phi(x) = \lim_{t \searrow 0} \frac{E[\phi(x_t)|x_0 = x] - \phi(x)}{t}$$

where the limit is defined using the mean-square norm on L^2. Thus the generator is the time zero derivative of the semigroup of conditional expectation operators.

2.1 Alternative Notions of Mixing

We consider three alternative notions of mixing. While these notions are defined more generally, we consider their specialization for Markov processes and use operator formulations as in Rosenblatt (1971). [3]

Let $Z = \{\phi \in L^2 : \int \phi dQ = 0\}$ denote the space of square-integrable functions with mean zero. The first measure of temporal dependence is the ρ-mixing (or maximal correlation) coefficients:[4]

Definition 2.1. *The $\rho - mixing$ coefficients are given by:*

$$\rho_t = \sup_{\phi \in Z, ||\phi||_2 = 1} ||\mathcal{T}_t \phi||_2$$

The process $\{x_t\}$ is $\rho - mixing$ if $\lim_{t\to\infty} \rho_t = 0$; and is $\rho - mixing$ with exponential decay rate if $\rho_t \leq \exp(-\delta t)$ for some $\delta > 0$.

Banon (1977) (Lemma 3.11) and Bradley (1986) (Theorem 4.2) established that for a stationary Markov process either the $\rho - mixing$ coefficients decay exponentially or they are identically equal to one.

The next measure of temporal dependence we consider is the α-mixing (or strong mixing):[5]

Definition 2.2. *The $\alpha - mixing$ coefficients are given by:*

$$\alpha_t = \sup_{\phi \in Z, ||\phi||_\infty = 1} ||\mathcal{T}_t \phi||_1.$$

The process $\{x_t\}$ is $\alpha - mixing$ if $\lim_{t\to\infty} \alpha_t = 0$; and is $\alpha - mixing$ with exponential decay rate if $\alpha_t \leq \gamma \exp(-\delta t)$ for some $\delta > 0$ and $\gamma > 0$. The process is $\alpha - mixing$ with a sub-exponential decay rate if $\alpha_t \leq \xi(t)$ for some positive non-increasing rate function ξ satisfying $\frac{1}{t} \log \xi(t) \to 0$, as $t \to \infty$.

[3]This subsection is largely based on Rosenblatt (1971) and Bradley (1986). They stated their results for discrete-time stationary Markov processes on general state spaces. However, it is easy to see that their results and proofs remain valid for continuous-time Markov processes.

[4]See the proof of Lemma VII.4.1 in Rosenblatt (1971).

[5]See the proof of Lemma VII.3.1 in Rosenblatt (1971)

Since the L^2 norm of a function ϕ is less than the L^∞ norm but exceeds the L^1 norm,

$$\frac{||\mathcal{T}_t\psi||_2}{||\psi||_2} \geq \frac{||\mathcal{T}_t\psi||_1}{||\psi||_\infty}.$$

Therefore,

$$\rho_t \geq \alpha_t.$$

In contrast to the $\rho-$*mixing* coefficients, the $\alpha-$*mixing* coefficients need *not* converge to zero at an exponential rate.

A third way way of measuring temporal dependence is given by the $\beta -$ *mixing* coefficients.[6]

Definition 2.3. *(Davydov (1973)) The $\beta -$ mixing coefficients are given by:*

$$\beta_t = \int \sup_{0\leq\phi\leq 1} \left| \mathcal{T}_t\phi(x) - \int \phi dQ \right| dQ.$$

The process $\{x_t\}$ is $\beta -$ mixing if $\lim_{t\to\infty} \beta_t = 0$; is $\beta -$ mixing with exponential decay rate if $\beta_t \leq \gamma \exp(-\delta t)$ for some $\delta > 0$ and $\gamma > 0$. The process is $\beta -$ mixing with sub-exponential decay rate if $\lim_{t\to\infty} \xi_t\beta_t = 0$ for some positive non-decreasing rate function ξ satisfying $\xi_t \to \infty$, $t^{-1}\ln \xi_t \to 0$ as $t \to \infty$.

A strictly stationary Markov process is $\beta -$ *mixing* if and only if it is (Harris) recurrent and aperiodic; see *e.g.*, Bradley (1986) (Theorem 4.3). Since

$$\sup_{|\phi|\leq 1} \int \left| \mathcal{T}_t\phi(x) - \int \phi dQ \right| dQ \leq \int \sup_{|\phi|\leq 1} \left| \mathcal{T}_t\phi(x) - \int \phi dQ \right| dQ,$$

it follows that

$$\beta_t \geq \alpha_t.$$

In contrast to the $\rho -$ *mixing* coefficients, the $\beta -$ *mixing* coefficients, like the $\alpha -$ *mixing* coefficients, need *not* converge to zero at an exponential rate. For general stationary Markov processes, the two dependence measures are not comparable: $\rho -$ *mixing* does not imply $\beta -$ *mixing* and $\beta -$ *mixing* does not imply $\rho -$ *mixing*; see *e.g.*, Bradley (1986). Nevertheless, all of the diffusion models we consider in this paper are $\beta -$ *mixing*, but some have $\rho_t \equiv 1$ for all t.

[6] $\beta -$ *mixing* is also called absolutely regular. It was studied by Volkonskii and Rozanov (1959), but they attribute the concept to Kolmogorov. The definition presented here is an alternative but equivalent one for a stationary Markov process, see *e.g.* Davydov (1973).

2.2 f- ergodicity

The notion $\beta - mixing$ for a Markov process is closely related to the concept called $f - ergodicity$ (in particular $1 - ergodicity$), see *e.g.*, Meyn and Tweedie (1993).

Definition 2.4. *Given a Borel measurable function $f \geq 1$, the Markov process $\{x_t\}$ is $f - ergodic$ if*

$$\lim_{t\to\infty} \sup_{0\leq\phi\leq f} \left| \mathcal{T}_t\phi(x) - \int \phi dQ \right| = 0 \text{ , for all } x \text{ .}$$

The Markov process $\{x_t\}$ is $f - uniformly\ ergodic$ if for all $t \geq 0$,

$$\sup_{0\leq\phi\leq f} \left| \mathcal{T}_t\phi(x) - \int \phi dQ \right| \leq cf(x)\exp(-\delta t)$$

for positive constants c and δ.[7]

A stationary process that is *f-uniformly ergodic* will be $\beta - mixing$ with exponential decay rate provided that $Ef(x_t) < \infty$. This connection is valuable because Meyn and Tweedie (1993) and Down et al. (1995) provide convenient drift conditions for f-uniform ergodicity. There exist other methods to establish $1 - ergodicity$ (i.e. when $f \equiv 1$) with sub-exponential decay rates, for example see Lindvall (1983).

3 Temporal dependence of a scalar diffusion

A scalar diffusion is typically represented as the solution to a stochastic differential equation:

$$dx_t = \mu(x_t)dt + \sigma(x_t)dW_t \tag{1}$$

with left boundary ℓ and right boundary r, either of which can be infinite. The function μ is the drift, σ^2 is the diffusion coefficient and $\{W_t\}$ is a standard Brownian motion.

Assumption 3.1. *μ and σ are continuous on (ℓ, r) with σ strictly positive on this interval.*

The generator of this scalar diffusion is known to be the differential operator:

$$\mathcal{A}\phi = \mu\phi' + \frac{1}{2}\sigma^2\phi''.$$

We will give a precise statement of the domain D of this generator subsequently.

[7]Our use of the term uniform ergodicity follows Down et al. (1995), but it differs from the use in Meyn and Tweedie (1993). Meyn and Tweedie (1993) define f *-uniform ergodicity* by requiring the left-hand side to converge to zero uniformly in x as t gets large.

The boundary behavior of a diffusion is characterized by the behavior of its scale function $S(.)$:

$$
\begin{aligned}
S(x) &= \int_a^x s(y)dy \text{ , for some fixed } a \in (\ell, r) \text{ ,} \\
s(y) &= \exp\left[-\int_a^y \frac{2\mu(u)}{\sigma^2(u)}du\right] .
\end{aligned}
$$

Since the scale function is increasing, it is well defined at both boundaries. A boundary is attracting when the scale function is finite at that boundary. We focus exclusively on the case in which neither boundary is attracting:

Assumption 3.2. *$S(\ell) = -\infty$ and $S(r) = +\infty$.*

We provide sufficient conditions that, among other things, guarantee that there exists a stationary distribution Q for the scalar diffusion. Since both boundaries are not attracting, this distribution, when it exists, is known to be unique and have a density q that is proportional to $\frac{1}{s\sigma^2}$. Thus stationarity is satisfied when:

Assumption 3.3. $\int_\ell^r \frac{1}{s(x)\sigma^2(x)}dx < \infty$.

From Has'minskii (1980) (Example 2, page 105), the process $\{x_t\}$ is Harris recurrent if Assumptions 3.1 and 3.2 hold. Hence $\{x_t\}$ is positive (Harris) recurrent if Assumptions 3.1, 3.2 and 3.3 hold. The diffusion is also aperiodic under these assumptions. Therefore,

Remark 3.4. *Under Assumptions 3.1, 3.2 and 3.3, the process $\{x_t\}$ is $\beta - mixing$.*

Hansen and Scheinkman (1995) provide sufficient conditions for a scalar diffusion to be $\rho - mixing$ with exponential decay, and Hansen et al. (1998) provide alternative conditions for a diffusion to be strongly dependent in the sense that $\rho_t \equiv 1$ for all t.

3.1 Diffusion in natural scale

For convenience, we transform the diffusion $\{x_t\}$ monotonically to its natural scale process $\{z_t\}$:

$$z_t = S(x_t).$$

Clearly the state space for $\{z_t\}$ is the entire real line $(-\infty, +\infty)$ under Assumption 3.2. Moreover, $\{z_t\}$ in the natural scale is known to be a local martingale (have zero drift) with diffusion coefficient:

$$\theta^2(z) = s^2[S^{-1}(z)]\sigma^2[S^{-1}(z)].$$

An equivalent statement of Assumption 3.3 is:

$$\int_{-\infty}^{+\infty} \frac{1}{\theta^2(z)}dz < \infty.$$

For the natural scale diffusion, the generator is the second-order differential operator

$$\mathcal{A}\phi = \frac{1}{2}\theta^2\phi''$$

defined on the domain:

$$D = \left\{ \phi : \phi' \text{ is absolutely continuous and } \begin{array}{l} \int \frac{1}{\theta^2(z)}\phi^2(z)dz < \infty \\ \int (\phi'(z))^2 dz < \infty \\ \int (\theta(z)\phi''(z))^2 dz < \infty \end{array} \right\}.$$

It follows from the definitions that the $\rho-$,$\beta-$ and $\alpha-$ mixing coefficients for $\{z_t\}$ are the same as those for $\{x_t\}$.

3.2 Equivalence of f-uniform ergodicity to ρ -mixing

We already pointed out that $f-$uniform ergodicity implies β - mixing with exponential decay. In this section, we explore further the near equivalence of $f-$uniform ergodicity and ρ - mixing for a stationary scalar diffusion. To establish such a link, we use a characterization of $f - uniform\ ergodicity$ by Down et al. (1995) and a characterization of $\rho - mixing$ by Hansen et al. (1998).

Let C^2 denote the space of functions mapping (ℓ, r) into $\mathbb{R}$ with continuous first and second derivatives. We construct the local operator:

$$\mathcal{B}V = \mu V' + \frac{1}{2}\sigma^2 V'',$$

which coincides with the generator $\mathcal{A}$ on the intersection of the domain D of the generator and C^2. A non-negative function V in C^2 is *norm-like* if $\{x : V(x) \leq v\}$ is compact in (ℓ, r) for any $v > 0$.

The following result is Theorem 5.2 of Down et al. (1995) (page 1681) specialized to our scalar diffusion $\{x_t\}$:

Theorem 3.5. *Suppose Assumptions 3.1 and 3.2 hold, and that there exists a non-negative function $V \in C^2$ (not necessarily norm-like) such that:*

$$\mathcal{B}V \leq -c(V + 1) + d\mathbf{1}_K \tag{2}$$

for some positive constants c and d and some compact set K. Then: (i) $\{x_t\}$ has a unique invariant probability measure Q; (ii) $\int V(x)dQ(x) < \infty$; (iii) $\{x_t\}$ is $(V+1)$-uniformly ergodic.

For a diffusion in natural scale, we consider solutions to the second-order differential equation:

$$\mathcal{B}\phi = -c\phi$$

for $\phi \in C^2$, and some $c > 0$, which is in the form of an eigenvalue problem. Weidmann (1987) (page 225) shows that there exists a nonnegative a such that for $c \geq a$ solutions ϕ cross the zero axis only a finite number of times and for $c < a$ they cross the axis an infinite number of times. Hansen et al. (1998) show that the corresponding diffusion is $\rho - mixing$ if and only if $a > 0$. We exploit this characterization of $\rho - mixing$ to establish the following two theorems.

Theorem 3.6. *Suppose that Assumptions 3.1, 3.2 and 3.3 hold, and $\{x_t\}$ is $\rho -$ mixing. Then there exists a nonnegative function $V \in C^2$ such that*

$$\frac{1}{2}\theta^2 V'' \leq -c(V+1) + d\mathbf{1}_K.$$

for some compact interval K and positive constants c and d. As a consequence, the diffusion is $(V+1)$ uniformly ergodic and β - mixing with exponential decay.

Proof. See Appendix. □

We next present a (partial) converse to this result.

Theorem 3.7. *Suppose that Assumptions 3.1 and 3.2 hold, and there is a nonnegative function $V \in C^2$, a compact interval K and positive constants c and d such that*

$$\frac{1}{2}\theta^2 V'' \leq -c(V+1) + d\mathbf{1}_K. \tag{3}$$

Then: $\{x_t\}$ satisfies Assumption 3.3 and is ρ -mixing.

Proof. See Appendix. □

4 Mixing with exponential decay rates

In this section we show that

$$\frac{\mu}{\sigma} - \frac{\sigma'}{2}$$

provides a measure of the *pull* of the diffusion by establishing formally its implications for *weak* dependence (*i.e.*, $\rho-$, $\beta-$mixing with exponential decay rates). We also propose extensions of this measure that avoid differentiability of the diffusion coefficient.

4.1 Natural Scale Case

We first study the temporal dependence of diffusions in the natural scale. As we will see, for these diffusions there is a direct link between the thickness of the tails of stationary density and the temporal dependence of the diffusion. We use this link to provide convenient sufficient conditions for decay rates of the mixing coefficients.

Theorem 4.1. *Suppose Assumptions 3.1 and 3.2 are satisfied. If*

$$\liminf_{|z|\to\infty} \frac{\theta(z)}{|z|} > 0, \tag{4}$$

then: (i) Assumptions 3.3 is satisfied; and (ii) $\{z_t\}$ is ρ - and β - mixing with exponential decay.

Proof. We establish this result by applying Theorems 3.6 and 3.7. We consider the following Lyapunov function:

$$V(z) = \frac{|z|^\alpha}{\alpha}, |z| \geq 1$$

for some $0 < \alpha < 1$, and we fill in the function V on $(-1, 1)$ so that it is C^2 and nonnegative. The constructed function is norm-like, and to guarantee inequality (3), we restrict θ^2 to satisfy:

$$\frac{\theta^2}{2} V'' \leq -cV + d \ , \ c > 0 \ , \tag{5}$$

for positive constants c and d. Inequality (5) will be satisfied if we can find positive constants c and d such that

$$\frac{1}{2}(\alpha - 1)\theta^2(z)|z|^{\alpha-2} \leq -c\frac{|z|^\alpha}{\alpha} + d \text{ for } |z| \geq 1.$$

If necessary, we adjust the constant d so that (5) is satisfied over the entire real line. Such constants exist given inequality (4). Clearly this condition is also sufficient for

$$\int \frac{1}{\theta^2(z)} dz < \infty$$

which guarantees the existence of a stationary distribution (Assumption 3.3). □

4.2 Diffusion with Nonzero Drift

If a diffusion $\{z_t\}$ in natural scale is stationary and $\beta - mixing$ with exponential decay rates, clearly so is the original process $\{x_t\}$ where $x_t = S^{-1}(z_t)$. Transforming the limit in (4) of Theorem 4.1 back to the original scale, we obtain:

Corollary 4.2. *Suppose that Assumptions 3.1 and 3.2 are satisfied. If*

$$\begin{aligned} \liminf_{x \nearrow r} \frac{s\sigma}{S} &> 0 \\ \limsup_{x \searrow \ell} \frac{s\sigma}{S} &< 0, \end{aligned}$$

then: $\{x_t\}$ satisfies Assumption 3.3, and is ρ - and β - mixing with exponential decay.

Remark 4.3. *Under Assumptions 3.1 and 3.2, if σ is differentiable, then the inequalities in Corollary 4.2 are implied by:*

$$\begin{aligned} \limsup_{x \nearrow r} \left(\frac{\mu}{\sigma} - \frac{\sigma'}{2} \right)(x) &< 0, \\ \liminf_{x \searrow \ell} \left(\frac{\mu}{\sigma} - \frac{\sigma'}{2} \right)(x) &> 0 \end{aligned} \tag{6}$$

We may think of

$$\left(\frac{\mu}{\sigma} - \frac{\sigma'}{2} \right)(x)$$

as providing a measure of the pull of the diffusion process $\{x_t\}$ (or $-\frac{1}{2}\theta'(z)$ as the pull measure for the natural scale diffusion $\{z_t\}$). For the scalar diffusion to be $\beta-$mixing with exponential decay we require that the pull be negative at the right boundary and positive at the left boundary. These restrictions are identical to the ones proposed by Hansen and Scheinkman (1995) for $\rho-$mixing with exponential decay, albeit their derivation is different than ours.

Remark 4.4. *Instead of assuming the differentiability of σ, we can require the existence of a positive function g that is differentiable and is dominated by σ:*

$$\frac{\sigma}{g} \geq 1.$$

The inequalities in Corollary 4.2 are now implied by:

$$\begin{aligned} \limsup_{x \nearrow r} \left(\frac{\mu}{\sigma^2} g - \frac{g'}{2} \right) &< 0 \\ \liminf_{x \searrow \ell} \left(\frac{\mu}{\sigma^2} g - \frac{g'}{2} \right) &> 0. \end{aligned}$$

Notice that these inequalities provide a trade-off between the drift and the diffusion behavior and cover examples in which the drift is dominated by the square root of the diffusion coefficient. In these cases, the exponential decay of the β-*mixing* coefficients is induced by the rapid increase in the volatility as a function of the Markov state. The drift may even be positive for states in the vicinity of the right boundary and negative in the vicinity of the left boundary (suggesting a pull to the left) while the resulting diffusion may still be stationary and uniformly ergodic. (See Conley et al. (1997) for a further discussion of volatility-induced stationarity.) Also, these conditions permit the diffusion coefficient to go to zero at either boundary.

We conclude this section by relating our work to that of Veretennikov (1987). Notice that for the special case when (i) σ is differentiable, we can transform our original process (1) using the twice differentiable function:

$$R(x) = \int_d^x \frac{1}{\sigma(u)} du,$$

where d is an interior point in the state space (ℓ, r). This scale transformation results in a new diffusion process $\{y_t = R(x_t) : t \geq 0\}$ with state space $(R(\ell), R(r))$, a unit diffusion coefficient $a^2 = 1$ and a drift given by our pull measure:

$$b = \frac{\mu}{\sigma} - \frac{\sigma'}{2}.$$

Under an additional condition (ii) $(R(\ell), R(r)) = (-\infty, \infty)$,[8] one can now apply Veretennikov (1987)'s theorem to establish $\beta - mixing$ with exponential decay based on the behavior of the drift coefficient b of the transformed diffusion with a constant diffusion coefficient. While this provides an alternative way to justify our (6) as a sufficient condition for $\beta - mixing$ with exponential decay, our derivation does not need the extra conditions (i) and (ii).[9]

5 $\beta-$Mixing with polynomial decay rates

Previously, we deduced sufficient condition for a diffusion process to be weakly dependent. In this section, we study diffusions which are strongly dependent in the sense that the $\rho - mixing$ coefficients $\rho_t \equiv 1$ for all $t \geq 0$, or $\beta - mixing$ coefficients decay at rates that are slower than exponential.

To study strong dependence, we focus on cases in which the pull measure

$$\frac{\mu}{\sigma} - \frac{\sigma'}{2}$$

converges to zero at one of the two boundaries. To characterize the strong dependence we study how fast this measure converges to zero. This leads us to investigate the limits:

$$\begin{aligned}
\nu^+ &\equiv \limsup_{x \nearrow r} \left[\left(\frac{\sigma^2}{\sigma\sigma' - 2\mu} \right)' - \frac{2\mu}{\sigma\sigma' - 2\mu} \right] \\
\nu^- &\equiv \liminf_{x \searrow \ell} \left[\left(\frac{\sigma^2}{\sigma\sigma' - 2\mu} \right)' - \frac{2\mu}{\sigma\sigma' - 2\mu} \right]
\end{aligned}$$

to obtain a more refined measure of strong dependence. These limits bound the behavior of the β-mixing coefficients.

5.1 Hitting Times

For a general discrete time Markov process, the $\beta-$mixing polynomial decay rates are implied by restrictions on the moments of the random time it takes to hit a compact

[8] Notice that for volatility-induced stationarity type of diffusion models, it may well be $R(\ell) > -\infty$ and/or $R(r) < \infty$. So $(R(\ell), R(r)) = (-\infty, \infty)$ is an additional restriction.

[9] In particular, our Corollary 4.2 *cannot* be derived from Veretennikov (1987)'s theorem.

set; see e.g., Tuominen and Tweedie (1994) (theorems 2.3 and 4.3). A version of this kind of result for scalar diffusion was established by Lindvall (1983), which we now state.[10] Let τ_K denote the first time that the process $\{z_t\}$ hits the point K in the interior of the state space conditioned on being in state z at date zero:

$$\tau_K(z) = \inf\{t \geq 0,\, z_t = K | z_0 = z\}.$$

Theorem 5.1. *Suppose Assumptions 3.1, 3.2 and 3.3 hold. Suppose that there is a function $\xi \geq 0$ non-decreasing on $[0,\infty)$ such that $E[\xi(\tau_K)] < \infty$. Then*

$$\lim_{t\uparrow\infty} \xi(t)\beta_t = 0,$$

if further, ξ is absolutely continuous with respect to Lebesgue measure and has a density ξ', then

$$\int_0^\infty \xi(t)'\beta_t < \infty.$$

In particular, if $\xi(t) = t^\delta$ for some $\delta > 0$ and $E\left[(\tau_K)^\delta\right] < \infty$, then:

$$\int_0^\infty t^{\delta-1}\beta_t < \infty,$$

$$\lim_{t\to\infty} t^\delta \beta_t = 0. \tag{7}$$

Proof. See Appendix. □

5.2 Natural Scale Case

One strategy for establishing the hitting time moment bounds of Theorem 5.1 is to follow Lindvall (1983) and study natural scale diffusion processes. Transforming the scale does not alter the hitting time distribution. Lindvall (1983) derives a sufficient condition for $E(\tau_K)^\delta < \infty$ based on a moment restriction expressed using the natural scale stationary distribution. The following result is based on Lemma 1 and Proposition 2 of Lindvall (1983) and our Theorem 5.1, but is stated in terms of growth rates of θ. It is the counterpart of Theorem 4.1, but for slower growth rates.

Theorem 5.2. *Suppose that Assumptions 3.1 and 3.2 hold, and there exists some constant $\frac{1}{2} < \eta < 1$ such that:*

$$\liminf_{|z|\to\infty} \frac{\theta(z)}{|z|} = 0, \quad \liminf_{|z|\to\infty} \frac{\theta(z)}{|z|^\eta} > 0. \tag{8}$$

Let

$$\eta^* \equiv \sup\left\{\eta \in \left(\frac{1}{2}, 1\right) : \text{inequality} \quad (8) \quad \text{is} \quad \text{satisfied}\right\}.$$

[10] Lindvall (1983) does not make any link to β-mixing, but, as we show in the Appendix, it is easy to modify his result to obtain β-mixing coefficients decaying at a polynomial order.

Then: (i) Assumption 3.3 is satisfied; (ii) $\{z_t : t \geq 0\}$ *is* $\beta-$*mixing with* $\lim_{t\to\infty} t^{\delta}\beta_t = 0$ *for any* $\delta < \delta^* = \frac{2\eta^*-1}{2-2\eta^*}$*, but is not* β *- mixing with exponential decay.*

Proof. Equation (8) implies the result (i) as long as $\eta > 1/2$. Equation (8) also implies that

$$\int_{-\infty}^{\infty} |z|^{1-(\alpha)^{-1}} \frac{1}{\theta^2(z)} dz < \infty \text{ provided } 2\eta + (\alpha)^{-1} - 1 > 1.$$

Let $0 < \delta \equiv \alpha - 1 < \frac{2\eta-1}{2-2\eta}$. Then by Lemma 1 or Proposition 2 of Lindvall (1983), we have $\int_{-\infty}^{\infty} (\tau_K)^{\delta} \frac{1}{\theta^2} < \infty$. The result (ii) now follows from Theorem 5.1 inequality (7). Finally, Hansen et al. (1998) (Theorems 4.2 and 4.3) show that when $\liminf_{|z|\to\infty} \frac{\theta(z)}{|z|} = 0$ there is no spectral gap (i.e., $\rho_t = 1$); hence $\{z_t : t \geq 0\}$ cannot be $\beta - mixing$ with exponential decay. □

Theorem 5.2 gives interesting results when the tail growth of θ as a function of $|z|$ exceeds $|z|^{\frac{1}{2}}$ but is less than linear. Slower growth in θ implies slower decay in the $\beta - mixing$ coefficients.

We next derive a sufficient condition for (8) in the natural scale. Although we derive this in terms of the natural scale, our interest in this sufficient condition is its counterpart in the original scale. The logarithmic derivative of a power function is proportional to $1/z$. The coefficient used in this proportionality dictates when the tail inequality (8) is satisfied. Thus we are led to compute the derivative:

$$\left(\frac{\theta}{\theta'}\right)'(z) = 1 - \frac{\theta''(z)\theta(z)}{[\theta'(z)]^2}$$

to study the tail behavior of θ. We are interested in the case in which θ' tends to zero. Define,

$$\nu^+ \equiv \limsup_{z\to+\infty} \left[1 - \frac{\theta''(z)\theta(z)}{[\theta'(z)]^2}\right]$$
$$\nu^- \equiv \liminf_{z\to-\infty} \left[1 - \frac{\theta''(z)\theta(z)}{[\theta'(z)]^2}\right].$$

Moreover, let

$$\nu^* \equiv \begin{cases} \nu^+ & \text{if } \limsup_{z\searrow-\infty} \theta'(z) > 0 \\ \nu^- & \text{if } \liminf_{z\nearrow+\infty} \theta'(z) < 0 \\ \max\{\nu^+, \nu^-\} & \text{otherwise.} \end{cases}$$

Lemma 5.3. *Suppose that* θ *is twice differentiable, and that*

$$\liminf_{z\nearrow+\infty} \theta'(z) \leq 0,$$
$$\limsup_{z\searrow-\infty} \theta'(z) \geq 0,$$

where at least one of these two limits is zero. If $1 < \nu^* < 2$*, then (8) is satisfied for any* $\eta < \frac{1}{\nu^*}$*. In particular,* $\eta^* \geq \frac{1}{\nu^*}$*.*

Proof. We prove this result when both limiting derivatives are equal to zero. The other cases can be proved using a more direct argument for one boundary and an entirely similar argument to what follows for the other boundary. Consider any $\frac{1}{\eta} > \nu^*$. For sufficiently large z^*, and $z \geq z^*$,

$$\frac{1}{(\log\theta)'(z)} \leq \frac{z}{\eta} + c_1,$$

and $z \leq -z^*$

$$\frac{1}{(\log\theta)'(z)} \geq \frac{z}{\eta} - c_1$$

where c^+ and c^- are some appropriately chosen positive constants. Taking reciprocals reverses the inequalities. Hence taking reciprocals and integrating implies that for $z \geq z^*$:

$$\log\theta(z) \geq \eta\log(z/\eta + c_1) + c_2,$$

and for $z \leq -z^*$:

$$\log\theta(z) \geq \eta\log(-z/\eta + c_1) + c_2$$

for some appropriately chosen constant c_2. Consequently, we may find a positive constant c_3 such that

$$\theta(z) \geq c_3\left(|z|/\eta + c_1\right)^\eta$$

for $|z| \geq z^*$. Therefore, (8) is satisfied. □

The following example illustrates the polynomial bounds on the $\beta - mixing$ coefficients and in particular when Lemma 5.3 produces the same bounds (*i.e.* $\eta^* = \frac{1}{\nu^*}$).

Example 5.4. *Suppose that* $\theta(z) = (1 + |z|^2)^{\frac{\gamma}{2}}$ *and* $1/2 < \gamma < 1$. *It may be shown directly that* $\eta^* = \gamma$. *Thus from Theorem 5.2,* $\lim_{t\to\infty} t^\delta\beta_t = 0$ *for any*

$$\delta < \frac{2\gamma - 1}{2 - 2\gamma}.$$

Notice also that ν^* *used in Lemma 5.3 is given by* $\nu^* = \frac{1}{\gamma} = \frac{1}{\eta^*}$.

5.3 Diffusion with a Nonzero Drift

Our main interest in Lemma 5.3 is its implication for a diffusion process in the original scale. We now transform the condition given in this lemma. Note that

$$\theta'(S(x)) = -2\left(\frac{\mu}{\sigma} - \frac{\sigma'}{2}\right)(x) = \left(\frac{\sigma\sigma' - 2\mu}{\sigma}\right)(x).$$

We are interested in cases in which θ' tends to zero. Since $\theta(S(x)) = s(x)\sigma(x)$,

$$\left(\frac{\theta[S(x)]}{\theta'[S(x)]}\right)' = \frac{1}{s}\left(\frac{s\sigma^2}{\sigma\sigma' - 2\mu}\right)'(x) = \left[\left(\frac{\sigma^2}{\sigma\sigma' - 2\mu}\right)' - \frac{2\mu}{\sigma\sigma' - 2\mu}\right](x).$$

This leads us to define:

$$\nu^+ \equiv \limsup_{x \nearrow r} \left[\left(\frac{\sigma^2}{\sigma\sigma' - 2\mu} \right)' - \frac{2\mu}{\sigma\sigma' - 2\mu} \right]$$
$$\nu^- \equiv \liminf_{x \searrow \ell} \left[\left(\frac{\sigma^2}{\sigma\sigma' - 2\mu} \right)' - \frac{2\mu}{\sigma\sigma' - 2\mu} \right].$$

Corollary 5.5. *Suppose that*

$$\limsup_{x \nearrow r} \left(\frac{\mu}{\sigma} - \frac{\sigma'}{2} \right) \leq 0$$
$$\liminf_{x \searrow \ell} \left(\frac{\mu}{\sigma} - \frac{\sigma'}{2} \right) \geq 0$$

with at least one of these limits equal to zero. Let ν^* be ν^+ if only the first limit is zero, be ν^- if only the second limit is zero, and be $\max\{\nu^+, \nu^-\}$ if both limits are zero. If $1 < \nu^* < 2$, then for any $\delta < \frac{2-\nu^*}{2\nu^*-2}$ the process $\{x_t : t \geq 0\}$ is $\beta - mixing$ and $\lim_{t\to\infty} t^\delta \beta_t = 0$.

This corollary shows how to compute a polynomial bound on the rate of decay of the β mixing coefficients when the pull measure is zero in one of the two tails.

Example 5.6. *Suppose that $\sigma = 1$ and*

$$\mu(x) \begin{cases} \leq & -\frac{\kappa}{x} \quad x \geq a \\ \geq & -\frac{\kappa}{x} \quad x \leq -a \end{cases}$$

for some positive κ and a as in Veretennikov (1997). Then

$$\nu^* \equiv 1 + \frac{1}{2\kappa} \in (1, 2) \text{ provided } \kappa > \frac{1}{2},$$

and the restriction on δ is:

$$\delta < \frac{2 - \nu^*}{2\nu^* - 2} = \kappa - \frac{1}{2}.$$

This matches the conclusion in Veretennikov (1997) for a scalar diffusion.

6 Strong dependence and spectral densities

For linear time series models, it is common to link temporal dependence to the behavior of the spectral density near frequency zero. For instance, the rate of divergence of the spectral density at frequency zero gives a way to characterize *long memory* of a stochastic process. For this reason, we now examine the implied behavior of spectral density function for test functions applied to the Markov diffusion. In what follows, we will first deduce a convenient formula for calculating the spectral density at a given frequency for transformations of a natural scale diffusion. Then we will construct diffusion processes with spectral densities that diverge at frequency zero. For this phenomenon to occur, at the very least we need the processes to fail to be $\rho - mixing$. However, even when the $\rho - mixing$ coefficients are identically one, the spectral density at frequency zero will still be finite for many (but not all) functions of the Markov state. In particular, we will use Example 5.4 as a starting point for a natural scale diffusion that fails to be $\rho - mixing$, and transform the state space to obtain Markov processes with divergent spectral density functions.

6.1 A formula for the spectral density

Let $\{z_t\}$ be a natural scale diffusion with diffusion coefficient θ^2 (and generator $\mathcal{A} = \theta^2\phi''/2$). Let $Z \equiv \{\phi \in L^2 : \int \phi dQ = 0\}$ denote the class of real-valued *test* functions with zero means and finite variances, where Q has density proportional to $\frac{1}{\theta^2}$. For any test function $\phi \in Z$, the process $\{\phi(z_t)\}$ is stationary $\beta - mixing$, hence its spectral measure is absolute continuous and the spectral density $f(\omega)$ exists satisfying $\int_{-\infty}^{\infty} \frac{|\ln f(\omega)|}{1+\omega^2} d\omega < \infty$ and can be represented as: (see e.g. Ibragimov and Rozanov (1978), pages 34-36, 112 and 138)

$$f(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp(-i\omega t)\left[\int \phi\left(\mathcal{T}_t\phi\right) dQ\right] dt \quad \text{if} \quad \int_0^{\infty}\left(\int \left[\phi\left(\mathcal{T}_t\phi\right)\right] dQ\right) dt < \infty,$$

and

$$\begin{aligned} f(\omega) &= \lim_{M\to\infty} f_M(\omega) \quad \text{in } L^2((-\infty,\infty), Leb) \quad \text{if} \quad \int_0^{\infty}\left(\int \left[\phi\left(\mathcal{T}_t\phi\right)\right] dQ\right)^2 dt < \infty, \\ f_M(\omega) &= \frac{1}{2\pi}\int_{-M}^{+M} \exp(-i\omega t)\left[\int \phi\left(\mathcal{T}_t\phi\right) dQ\right] dt. \end{aligned}$$

Since the natural scale diffusion $\{z_t\}$ is time-reversible, the autocorrelations are non-negative:

$$\int \left[\phi\left(\mathcal{T}_t\phi\right)\right] dQ \geq 0.$$

Thus the spectral density function at frequency zero dominates the spectral density at all other frequencies.

For any given $\phi \in Z$ and frequency ω, we solve the differential equation system:

$$\left(\mathcal{A} - i\omega\mathcal{I}\right)\psi = \phi \tag{9}$$

where $\psi = \psi_r + i\psi_i$ and $\psi_r, \psi_i \in D$, (notice that the solution ψ depends implicitly on ω). Thus ψ solves the pair of differential equations:

$$\begin{aligned} \frac{\theta^2}{2}\psi_r'' &= -\omega\psi_i + \phi \\ \frac{\theta^2}{2}\psi_i'' &= \omega\psi_r \end{aligned}$$

for $\psi_r, \psi_i \in D$. By construction, the solution ψ satisfies for any $\omega \neq 0$

$$\psi = (\mathcal{A} - i\omega\mathcal{I})^{-1}\phi = -\int_0^\infty \mathcal{T}_t \exp(-i\omega t)\phi dt.$$

Then $\{\phi(z_t)\}$ has a finite spectral density at frequency $\omega \neq 0$ given by:

$$\begin{aligned} f(\omega) &= \int \phi \left[\int_{-\infty}^{+\infty} (\mathcal{T}_t \exp(-i\omega t)\phi)\, dt\right] dQ \\ &= -2 \text{ real } \left(\int \phi \left[(\mathcal{A} - i\omega\mathcal{I})^{-1}\phi\right] dQ\right) \\ &= -2\left(\int \phi\psi_r dQ\right). \end{aligned}$$

For frequency $\omega = 0$, the pair of differential equations (9) becomes $\psi_i = 0$ and $\frac{\theta^2}{2}\psi_r'' = \phi$ (i.e. $\mathcal{A}\psi = \phi$), which has a solution $\psi = \mathcal{A}^{-1}\phi \in D$ if and only if $\phi \in Z$ belongs to the range of $\mathcal{A}$. In this case, an integration-by-parts argument leads to:

$$f(0) = -2\int \phi\left(\mathcal{A}^{-1}\phi\right) dQ = \frac{\int (\psi_r')^2}{\int \frac{1}{\theta^2}}. \tag{10}$$

Moreover, $f(0)$ gives the asymptotic variance for the central limit approximation for $\left\{\frac{1}{\sqrt{N}}\int_0^N \phi(x_t)dt\right\}$ (see Bhattacharya (1982) and Hansen and Scheinkman (1995)).

Notice that when the diffusion process $\{z_t\}$ is $\rho-mixing$, the range of $\mathcal{A}$ coincides with the space Z, hence any test function $\phi \in Z$ has finite $f(0)$ given by the formula (10). However, when the diffusion process is strongly dependent in the sense $\rho_t \equiv 1$ for all $t \geq 0$, the range of $\mathcal{A}$ is merely a dense subset of Z; hence there exist functions $\phi \in Z$ that are outside the range of $\mathcal{A}$. For all the test functions ϕ that belong to the range of $\mathcal{A}$, the familiar central limit approximations continue to apply to $\{\phi(x_t)\}$, and for which the spectral densities remain bounded in the vicinity of frequency zero. When ϕ is outside the range of $\mathcal{A}$, we can no longer solve the operator equation $\mathcal{A}\psi = \phi$ in D. Bhattacharya (1982)'s Central Limit Theorem may fail and the spectral density may diverge at frequency zero. We now construct examples for which the spectral density becomes unbounded at frequency zero.

6.2 Divergent Spectral Densities

Formula (10) also suggests a way to construct transformations (test functions) ϕ with finite variances but infinite spectral densities at frequency zero for strongly dependent processes. Find ϕ's with zero means and finite variances that satisfy:

$$\frac{\theta^2}{2}\psi'' = \phi$$

for which the corresponding ψ's are outside the domain D and in particular:

$$\int (\psi')^2 = +\infty. \tag{11}$$

We illustrate such a construction by developing further Example 5.4:

$$\theta(z) = \left(1+z^2\right)^{\gamma/2}$$

for $\frac{1}{2} < \gamma < 1$. For this range of γ's we have already argued that the $\rho - mixing$ coefficients are all one. To find a test function ϕ with zero mean, finite variance and infinite spectral density at frequency zero, we use the construction:

$$\phi = \frac{\theta^2}{2}\,\psi''$$

and find a function ψ that satisfies (11) along with

$$\int \theta^2 \left(\psi''\right)^2 < \infty, \tag{12}$$

$$\lim_{z\to\pm\infty} \psi'(z) = 0. \tag{13}$$

An example of such a function is:

$$\psi(z)' = \left(1+z^2\right)^{-\eta/2}$$

for

$$\gamma - \frac{1}{2} < \eta \leq \frac{1}{2}. \tag{14}$$

The resulting test function ϕ has a finite variance and a zero mean against the stationary distribution by virtue of (12) and (13). The infinite spectral density of the process $\{\phi(z_t)\}$ at frequency zero is suggested by (11).

The process $\{\phi(z_t)\}$ is itself Markov for $\frac{1}{2} < \gamma \leq \frac{3}{4}$ (since $\phi(z)$ is decreasing), and also a scalar diffusion (since $\phi \in C^2$), but with a nonzero drift. To illustrate the divergence of the spectral density, we plot the spectral density for alternative choices of γ and η. First, we compute the spectral density functions $f(\omega)$ for frequencies ω in the vicinity of frequency zero for three values of $(\gamma,\eta) : \left(.51, \frac{1}{4}\right), \left(\frac{3}{4}, \frac{1}{3}\right), \left(.99, \frac{1}{2}\right)$ in Figure 1. The γ values .51 and .99 were chosen because they are near the endpoints of the interval $(\frac{1}{2}, 1)$. Recall that when γ is greater than or equal to one, the process is $\rho - mixing$ with exponential decay, and when γ is less than or equal to $\frac{1}{2}$, the process fails to be stationary. The γ value of $\frac{3}{4}$ is chosen because the $\beta - mixing$ coefficients are integrable for $\gamma > \frac{3}{4}$, and the mean time to hit a compact set is infinite when

$\gamma \leq \frac{3}{4}$. The corresponding η's were chosen to be close to the midpoint of the interval $(\gamma - \frac{1}{2}, \frac{1}{2}]$ in (14).[11]

Long memory processes including stationary versions of fractional Brownian motion (see Mandelbrot and Ness (1968)) have spectral densities that behave like:

$$\log f(\omega) \approx c_0 - c_1 \log(\omega) \tag{15}$$

in the vicinity of frequency zero for $0 < c_1 < 1$. For this reason, we also depict the spectral densities in a $\log - \log$ scale in Figure 2. Notice that while the log spectral density is distinctly concave in $\log(\omega)$ for $\gamma = .99$, it is almost linear for $\gamma = .51$. This latter behavior imitates closely the spectral density of long memory time series. We also study how η alters the slope:

$$slope = \frac{d \log f}{d \log \omega}$$

in Figure 3. Notice that decreasing η increases the magnitude of the slope, but slope never exceeds one, which is the upper bound on the parameter c_1 in (15).

Remark 6.1. *In the preceding example, our calculations were based primarily on the tail properties of the natural scale diffusion coefficient θ and of the function ψ. We are free to alter the behavior of these functions on compact subsets and to modify accordingly the transient dynamics for Markov states on compact sets without changing the divergence of the resulting spectral density functions.*

Remark 6.2. *Stationary versions of fractional Brownian motion are known to have infinite quadratic variations with probability one. As emphasized by Maheswaran and Sims (1993), this property may make such processes fail to be local martingales and as a consequence unappealing as models of arbitrage-free asset prices.*[12] *Maheswaran and Sims (1993) go on to argue that a nice feature of fractional Brownian motion that is often featured is its long range dependence and not its "finite-time unit" properties. They then study continuous-time Gaussian moving-average models that break the link between short run responses to shocks and long run dependence.*[13] *(Also, see Robinson*

[11] To compute the (normalized) spectral density for frequency ω, we first solved the differential equation (9) numerically subject to the boundary restrictions $\psi'(z^*) = \psi'(-z^*)$ for a large value of z^*. We then evaluated numerically:

$$f(\omega) = -2 \frac{\int \frac{\phi \psi_r}{\theta^2}}{\int \frac{\phi^2}{\theta^2}}$$

and checked the sensitivity of the answer to the choice of z^*.

[12] As follows from Harrison and Kreps (1979), Harrison et al. (1984) and Maheswaran and Sims (1993), local martingale or semimartingale implications for security markets are tied directly to the classes of admissible trading rules. Trading rule restrictions are required at the outset to admit even geometric Brownian motions as admissible processes. The more severe the trading rule restrictions, the larger is the class of admissible price processes.

[13] Similar properties have also been investigated by Comte and Renault (1996) for a multivariate continuous-time moving-average models driven by fractional Brownian Motion.

(1995) for a semiparametric estimation method of models that break this link in a discrete-time setting.) In a similar vein, our nonlinear diffusion examples show how to maintain the local Gaussian structure while inducing nonlinearities and long-run dependence. In particular, the local martingale property is preserved by construction.

7 Endogenous sampling

We conclude the paper by considering discrete-time processes obtained by sampling a diffusion in a manner that is state dependent. Following Duffie and Glynn (2004) we construct an endogenous sampling scheme built from a Poisson process with a state dependent intensity. Let $\{x_t\}$ denote a Markov process with generator $\mathcal{A}$, and let $\{N_t\}$ denote a Poisson process with state-dependent intensity λ. The event times $\tau_j^* = \inf\{t : N_t = j\}$ are the times at which x_t is observed. We denote the discrete-time process as $\left\{y_j^* : j = 0, 1, 2, ...\right\}$, where $y_j^* = x_{\tau_j^*}$, with $\tau_0^* = 0$.

While the intensity λ can depend on the Markov state, there is an equivalent way to depict the process $\left\{y_j^*\right\}$ with an intensity that is state independent. This construction first alters the time clock of the diffusion $\{x_t\}$ in a manner analyzed by Ethier and Kurtz (1986), and then uses a Poisson sampling process with a unit intensity applied to the diffusion with a distorted time clock.

7.1 Altering the time clock

The continuous-time Markov process $\{\widehat{y}_t : t \geq 0\}$ with a distorted time clock may be constructed as follows:

$$\widehat{y}_t = x_{\tau_t}$$

with the increasing process:

$$\tau_t = \int_0^t \frac{1}{\lambda(x_u)} du$$

Assumption 7.1. *λ is continuous, strictly positive on (ℓ, r).*

Ethier and Kurtz (1986) (pages 308-309) show that we may construct the Markov process $\{\widehat{y}_t\}$ by using the generator:

$$\widehat{\mathcal{A}} = \frac{1}{\lambda}\mathcal{A}$$

motivated heuristically by the chain rule. Since λ can be state dependent, the domain of $\widehat{\mathcal{A}}$ may differ from that of $\mathcal{A}$, but the intersection of the domains will typically contain a dense (in L^2) set of functions. While the Ethier and Kurtz (1986) construction is applicable to a general class of Markov processes, we are interested in the case in which the original process is a scalar diffusion with continuous drift and diffusion coefficients, $\mathcal{A}$ and hence $\widehat{\mathcal{A}}$ is a second-order differential operator that is at least well defined on the space C^2. The time-deformed process $\{\widehat{y}_t : t \geq 0\}$ is still a diffusion, where the drift and the diffusion coefficients are obtained by multiplying the original drift and diffusion coefficients of $\{x_t\}$ by the reciprocal of the intensity λ. The scale function of the new process remains unchanged.

The stationary distribution $\widehat{Q}$ of the process $\{\widehat{y}_t\}$ may be constructed as the Radon-Nikodym derivative proportional to:

$$d\widehat{Q} = \frac{\frac{\lambda}{s\sigma^2}}{\int_\ell^r \frac{\lambda}{s\sigma^2}}.$$

Given the state dependent intensity λ, $\widehat{Q}$ is different from the stationary distribution Q of $\{x_t\}$. In fact one may not even exist while the other one is a positive finite measure. The next assumption ensures the existence of $\widehat{Q}$:

Assumption 7.2. $\int_\ell^r \frac{\lambda(x)}{s(x)\sigma^2(x)} dx < \infty$.

Now all the results in previous sections apply to the time-altered diffusion $\{\widehat{y}_t : t \geq 0\}$. In particular, the pull measure for the altered process is given by:

$$\frac{1}{\lambda^{1/2}} \left[\frac{\mu}{\sigma} - \frac{\sigma'}{2} + \frac{\sigma \lambda'}{4\lambda} \right]$$

when λ and σ are differentiable.

7.2 Poisson sampling

We can now form the discrete-time process $\left\{y_j^* : j = 0, 1, 2, ...\right\}$ by taking a Poisson sample of $\{\widehat{y}_t : t \geq 0\}$ with a unit intensity. The resulting discrete-time process is still stationary with distribution $\widehat{Q}$, and is an aperiodic Markov chain with one-period transition operator:

$$\widehat{\mathcal{T}}\phi(y) = \left(\mathcal{I} - \widehat{\mathcal{A}}\right)^{-1} \phi(y). \tag{16}$$

The discrete-time transition operator is a special case of what is referred to as a resolvent operator for the generator $\widehat{\mathcal{A}}$.

The next result states that $\left\{y_j^* : j = 0, 1, 2, ...\right\}$ preserves all the temporal dependence properties of $\{\widehat{y}_t : t \geq 0\}$.

Theorem 7.3. *Suppose that Assumptions 3.1, 3.2 and 7.1 are satisfied. If*

$$\liminf_{x \nearrow r} \frac{s\sigma}{|S|\lambda^{1/2}} \quad > \quad 0 \tag{17}$$

$$\limsup_{x \searrow \ell} \frac{s\sigma}{|S|\lambda^{1/2}} \quad < \quad 0, \tag{18}$$

then: Assumption 7.2 holds and $\{y_j^* : j = 0, 1, 2, ...\}$ *is stationary,* $\rho-$*mixing and* $\beta-$*mixing with exponential decay rates.*

This theorem is a special case of Theorem A.1, which is stated and proved in the appendix. The latter theorem includes a characterization of the $\beta - mixing$ coefficients when the resulting process is strongly dependent.

Remark 7.4. *When σ and λ are smooth, the sufficient conditions for inequalities (17) and (18) are:*

$$\limsup_{x \nearrow r} \lambda^{-1/2} \left[\frac{\mu}{\sigma} - \frac{\sigma'}{2} + \frac{\sigma \lambda'}{4\lambda} \right] < 0,$$
$$\liminf_{x \searrow \ell} \lambda^{-1/2} \left[\frac{\mu}{\sigma} - \frac{\sigma'}{2} + \frac{\sigma \lambda'}{4\lambda} \right] > 0.$$

Consider now the case in which the subordinated process $\{\widehat{y}_t\}$ is stationary (Assumption 7.2 is satisfied), but its pull measure:

$$\lambda^{-1/2} \left[\frac{\mu}{\sigma} - \frac{\sigma'}{2} + \frac{\sigma \lambda'}{4\lambda} \right]$$

is zero at one of the two boundaries. By the arguments in Hansen et al. (1998), there exists a sequence of functions $\{\phi_j\}$ with norm one such that:

$$\lim_{j \to \infty} \int \phi_j \left(\widehat{\mathcal{A}} \phi_j \right) d\widehat{Q} = 0.$$

It follows that there exists a sequence of functions $\{\psi_j\}$ with mean zero and unit norm such that:

$$\lim_{j \to \infty} \int \psi_j \left(\widehat{\mathcal{T}} \psi_j \right) d\widehat{Q} = 1,$$

which implies that all of the discrete-time $\rho - mixing$ coefficients are unity. Hence the dependence properties of the Poisson sampled discrete-time process $\{y_j^*\}$ mirror that of the deformed continuous-time process $\{\widehat{y}_t\}$.

Our next examples illustrate how subordination can alter the unconditional distribution as well as the temporal dependence of a scalar diffusion:

Example 7.5. *Let $\{x_t\}$ be a stationary diffusion process on $(-\infty, +\infty)$ with $\mu(x) = -\gamma x$ for some $\gamma > -\frac{1}{2}$ and $\sigma^2(x) = 1 + x^2$. The stationary density $q(x)$ is proportional to $(1 + x^2)^{-\gamma - 1}$ (Example "E" in Wong (1964)). Clearly,*

$$\lim_{x \searrow -\infty} \frac{\mu}{\sigma} - \frac{\sigma'}{2} = \gamma + \frac{1}{2} > 0 \ , \ \lim_{x \nearrow +\infty} \frac{\mu}{\sigma} - \frac{\sigma'}{2} = -\gamma - \frac{1}{2} < 0.$$

Thus the $\rho - mixing$ and $\beta - mixing$ coefficients decay exponentially. Let $\{\widehat{y}_t\}$ be the time-deformed diffusion with $\lambda(x) = 1 + x^2$. Then for $\gamma \in (-\frac{1}{2}, \frac{1}{2}]$, this diffusion does not have a stationary distribution. While for $\gamma > \frac{1}{2}$ this diffusion has a stationary density proportional to $(1 + x^2)^{-\gamma}$, and has a pull measure equal to zero *at both boundaries. Thus the $\rho-$mixing coefficients are unity and the $\beta-$mixing coefficients decay slowly. When we take a Poisson sample of the process $\{\widehat{y}_t\}$ with a unit intensity, the discrete-time process $\left\{y_j^*\right\}$ remains non-stationary for $\gamma \in (-\frac{1}{2}, \frac{1}{2}]$, and stationary but strongly dependent for $\gamma > \frac{1}{2}$.*

Example 7.6. *Suppose that $\{x_t\}$ has the same drift and diffusion coefficient as $\{\widehat{y}_t\}$ had in the previous example: $\mu = \frac{-\gamma x}{1+x^2}$ and $\sigma^2 = 1$. As we just argued, this process is stationary and strongly dependent for $\gamma > \frac{1}{2}$, and fails to be stationary when $\gamma \in (-\frac{1}{2}, \frac{1}{2}]$. Let $\lambda(x) = \frac{1}{1+x^2}$, then the resulting time-deformed $\{\widehat{y}_t\}$ process coincides with the $\{x_t\}$ process of the previous example. Hence $\{\widehat{y}_t\}$ and the associated Poisson-sampled discrete-time process $\left\{y_j^*\right\}$ with a unit intensity are stationary, $\beta-$mixing and $\rho-$mixing with exponential decay.*

8 Multivariate diffusions

We now explore an alternative convenient formulation of a Markov diffusion process and give some extensions of our previous results. We use Ω to denote the set of hypothetical Markov states, and we restrict Ω to be open and connected subset of $\mathbb{R}^n$. Let y denote an element of Ω, or equivalently a possible realized value of the Markov state.[14] We will model a multivariate diffusion by first constructing a quadratic form on the domain C_K^2 of twice continuously differentiable functions with compact support. This form is built using a multivariate diffusion matrix and a stationary density. We study mixing as a restriction on tail behavior of these two objects.

For any pair of functions ϕ and ψ in C_K^2, we construct the positive semidefinite quadratic form:

$$f_o(\phi, \psi) = \frac{1}{2} \int \sum_{i,j} \sigma_{ij} \frac{\partial \phi}{\partial y_j} \frac{\partial \psi}{\partial y_i} q.$$

where

$$\Sigma = [\sigma_{ij}]$$

is a positive definite matrix for each Markov state y and q is a positive density that integrates to unity.

Assumption 8.1. *Σ is a continuously differentiable, positive definite matrix function on Ω.*

Assumption 8.2. *$q = \exp(-2h)$, and h is twice continuously differentiable satisfying $\int_\Omega \exp(-2h) = 1$.*

[14]Here and elsewhere in the section we define new notation with a distinct usage from that in previous sections.

We construct the generator for the semigroup of conditional expectation operators for Markov diffusion from the differential operator associated with the form by solving:

$$f_o(\phi, \psi) = -\langle \mathcal{B}\phi, \psi \rangle = -\langle \phi, \mathcal{B}\psi \rangle .$$

Applying integration by parts:

$$\mathcal{B}\phi = \frac{1}{2}\sum_{i,j} \sigma_{ij} \frac{\partial^2 \phi}{\partial y_i \partial y_j} + \frac{1}{2q}\sum_{i,j} \frac{\partial (q\sigma_{ij})}{\partial y_i}\frac{\partial \phi}{\partial y_j}.$$

With this formula, Σ is interpreted as the diffusion matrix and q as the stationary density. The implicit drift can be constructed from Σ and q via: $\mu = (\mu_1, ..., \mu_n)'$ satisfies

$$\mu_j = \frac{1}{2q}\sum_{i=1}^{n} \frac{\partial (q\sigma_{ij})}{\partial y_i}$$

for $j = 1, ..., n$ when $\{x_t\}$ satisfies the stochastic differential equation:

$$dx_t = \mu(x_t)dt + \Lambda(x_t)dW_t$$

with appropriate boundary restrictions. The process $\{W_t : t \geq 0\}$ is an n-dimensional, standard Brownian motion, and $\Sigma = \Lambda\Lambda'$. Notice that we start with Σ and q and infer a *unique* generator for a Markov diffusion. Under this construction the process $\{x_t\}$ is time *reversible*, although later we will also consider irreversible diffusions.

We extend the form f_o to a larger space H using the weak notion of a derivative.

$$\begin{aligned} H &= \{\phi \in L^2 : \text{ there exists } g \text{ measurable, with } \int g'\Sigma g q < \infty, \\ &\quad \text{and } \int \phi \nabla \psi = -\int g\psi, \text{for all } \psi \in \mathcal{C}_K^\infty\} \end{aligned}$$

The Borel measurable function g is unique (for each ϕ) and is referred to as the *weak derivative* of ϕ. From now on, for each ϕ in H we write $\nabla\phi = g$. Then H is a Hilbert space under the inner product

$$< \phi, \psi >_* = < \phi, \psi > + \frac{1}{2}\int (\nabla\phi)'\Sigma(\nabla\psi)q.$$

For any pair of functions ψ and ϕ in H, we define a quadratic form f

$$f(\phi, \psi) = \frac{1}{2}\int (\nabla\phi)'\Sigma(\nabla\psi)q,$$

as a closed form extension of f_o to H.

For the purpose of approximation, we maintain:

Assumption 8.3. *For any ϕ_o in H, there exists a sequence $\{\phi_j\}$ in C_K^2 such that*

$$\lim_{j\to\infty} < \phi_j - \phi_o, \phi_j - \phi_o >_* = 0.$$

Fukushima et al. (1994) and Chen et al. (2009) give sufficient conditions for this assumption to be satisfied. Under this assumption we can focus our attention on the form f_o defined on C_K^2 instead of the extension f.

Hansen and Scheinkman (1995) give the following necessary and sufficient condition for the Markov process to be ρ-mixing with exponential decay. (See Proposition 8.)

Condition 8.4. *There exists a $\delta > 0$ such that*

$$f_o(\phi, \phi) \geq \delta \left[< \phi, \phi > - \left(\int \phi q \right)^2 \right]$$

for all $\phi \in C_K^2$.

In the case of a scalar diffusion with state space (ℓ, u) and $\Sigma = \varsigma^2$, Hansen and Scheinkman (1995) show that Condition 8.4 is satisfied when

$$\liminf_{y\to u} \left[\varsigma'(y) + \varsigma(y) \frac{q'(y)}{q(y)} \right] > 0$$
$$\limsup_{y\to \ell} \left[\varsigma'(y) + \varsigma(y) \frac{q'(y)}{q(y)} \right] < 0$$

where both limits are assumed to exist. This shows how ρ-mixing (and hence β and α-mixing with exponential decay) can be induced either by a stationary density with a thin tail or by a volatility specification that grows at least linearly with the Markov state. Hansen and Scheinkman (1995) show that this restriction is equivalent to the scalar drift condition introduced previously in (6).

To obtain multivariate results, we follow an approach in Chen et al. (2009) developed for a different purpose.[15] We assume that $\Omega = \mathbb{R}^n$. In what follows we will impose a lower bound on the diffusion matrix Σ.

Assumption 8.5. *Suppose that*

$$\Sigma(y) \geq \varsigma(y)^2 I \geq \epsilon I$$

where $\varsigma(y) = \exp[v(y)]$ and v is continuously differentiable.

One possibility is to set $\varsigma^2 = \epsilon$. Define a *potential* function:

$$F(y) \equiv -\sum_{i,j} \sigma_{ij} \frac{\partial^2 h}{\partial y_i \partial y_j} - \sum_{i,j} \frac{\partial \sigma_{ij}}{\partial y_i} \frac{\partial h}{\partial y_j} + (\nabla h)' \Sigma (\nabla h). \tag{19}$$

[15]Chen et al. (2009) study the existence of functional principal components, which requires more stringent restrictions.

Theorem 8.6. *Let Assumptions 8.1, 8.2, 8.3 and 8.5 be satisfied. If*

$$\lim \inf_{|y|\to\infty} F(y) > 0,$$

then Condition 8.4 is satisfied. Hence $\{x_t\}$ is $\rho-$mixing.

Proof. See Appendix. □

As in Chen et al. (2009) it is sometimes possible to construct a more refined result by exploiting even more the state dependent growth in volatility. Such a result is important to accommodate processes with stationary densities and tail behavior that is algebraic rather than exponential. Form the potential function

$$\begin{aligned}\hat{F}(y) = \varsigma(y)^2 \left(-\text{trace}\left[\frac{\partial^2 h(y)}{\partial y_i \partial y_j} - \frac{\partial^2 v(y)}{\partial y_i \partial y_j} \right] + |\nabla h(y) - \nabla v(y)|^2 \right) \\ + \epsilon \left(\nabla v(y) \cdot \nabla v(y) - \text{trace}\left[\frac{\partial^2 v(y)}{\partial y_i \partial y_j} \right] \right).\end{aligned}$$

Theorem 8.7. *Let Assumptions 8.1, 8.2, 8.3 and 8.5 be satisfied. If*

$$\lim \inf_{|y|\to\infty} \hat{F}(y) > 0,$$

then Condition 8.4 is satisfied. Hence $\{x_t\}$ is $\rho-$mixing.

Proof. See Appendix. □

Notice that derivatives of the logarithms of both the density and the state dependent bound on the diffusion matrix contribute to the construction of the potential function $\hat{F}$.

While our use of forms leads naturally to representing a Markov diffusion in terms of a stationary density q and a diffusion matrix Σ, as we noted the resulting process is time reversible. In multivariate settings there are typically many other constructions of a generator which result in diffusions with the same density and the same diffusion matrix. Provided that the generator $\mathcal{A}$ satisfies

$$f_o(\phi, \psi) = -\int \psi(\mathcal{A}\phi) q$$

on C_K^2, our results remain applicable to a diffusion process with generator $\mathcal{A}$.

Rockner and Wang (2001) use forms to study multivariate Markov processes with slower than exponential rates of convergence of the α-mixing coefficients. Let C_B^2 be given by all linear combinations of functions in C_K^2 and constant functions. These functions are necessarily in H. Rockner and Wang (2001) provide sufficient conditions for the inequality

$$(||\mathcal{T}_t\phi||_2)^2 \leq \xi(t) \left[(||\phi||_\infty)^2 + (||\phi||_2)^2 \right] \quad \text{for all } \phi \in C_B^2 \cap Z,$$

For instance, see their Theorem 2.1. From this inequality, we see that

$$\frac{||\mathcal{T}_t\phi||_1}{||\phi||_\infty} \leq \frac{||\mathcal{T}_t\phi||_2}{||\phi||_\infty} \leq \sqrt{2\xi(t)}$$

and hence

$$\alpha_t \leq \sqrt{2\xi(t)}.$$

For an example of the construction of $\xi(t)$ in terms of Σ and $h = -\frac{1}{2}\log q$, see the discussion on page 579 of Rockner and Wang (2001).[16]

9 Concluding remarks

In this paper we studied the temporal dependence of nonlinear scalar diffusion models. As we have seen, scalar diffusion models provide a convenient and pedagogically valuable platform for understanding how nonlinearities in time series models get transmitted into temporal dependence. Scalar diffusions are of course special. Nevertheless, they are often used as building blocks in more realistic empirical models of financial data.

We explored extensions in Sections 7 and 8. We studied the temporal dependence in models of subordinated diffusions. In these models the time clock is distorted in a random and temporally dependent way. Since the work of Clark (1973) and Nelson (1990), it has been known that subordination is a convenient way to model returns with unconditional distributions that have fat tails and volatility that is clustered over time. Strongly dependent diffusions provide a useful tool for studying the time clock distortions as a way of inducing highly persistent stochastic volatility.[17] Also we showed how to extend some characterizations of dependence through the use of quadratic forms as a modeling device. This approach allowed us to include multivariate diffusion models in our analysis, and it demonstrates the connection between mixing properties and the tail behavior of stationary densities and conditional volatilities.

Strongly dependent diffusions, like models of fractional integration, serve to blur the distinction between stationary and nonstationary processes. As we have seen, the strong dependence of diffusions is conveniently manifested in the pull of the diffusion at extreme values of the Markov state. As a practical matter, this pull behavior will be hard to measure accurately without using parametric restrictions on, at the very least, the tail behavior of the drift and diffusion coefficients. This practical problem, however, is no different than what occurs in attempts to detect the degree of long range dependence in times series. It is known that once we allow for flexible transient dynamics, the degree of long range dependence is hard to measure.

[16]See Veretennikov (1997) for sufficient conditions for the polynomial decay rate of the $\beta-$mixing coefficients for multivariate, continuous-time Markov processes.

[17]Mixing properties of models with other forms of subordination are explored in Carrasco et al. (1999).

A Appendix

Proof. (**Theorem 3.6**): It suffices to prove this result in the natural scale. We establish uniform ergodicity by finding a nonnegative, C^2 function V that satisfies:

$$\frac{1}{2}\theta^2 V'' \leq -c(V+1) + d\mathbf{1}_K \tag{20}$$

for some positive numbers c and d and some compact set K. We construct the function V by first solving the eigenvalue problem:

$$\frac{1}{2}\theta^2 \phi'' = -c\phi$$

for some $c > 0$ and some $\phi \in C^2$, and then we construct V from ϕ. If ϕ solves this differential equation than so does $a\phi$ for any real number a. Since the diffusion is $\rho - mixing$, we may choose a sufficiently small $c > 0$ such that ϕ has only a finite number of zeroes (*e.g.* see Weidmann (1987), page 225). Let K be a closed interval containing all of the zeroes as interior points. Notice that ϕ is concave when it is positive and convex when it is negative. Thus ϕ is bounded away from zero in both tails. We let $V + 1$ be equal to $a\phi$ to the left of K and $b\phi$ to the right of K where the scale factors are chosen so that $V + 1$ exceeds one outside K. We extend V to the interior so that it remains nonnegative and is C^2, and select d to guarantee inequality (20) on K. □

Proof. (**Theorem** 3.7): Let ϕ be a non-trivial solution of the eigenvalue problem:

$$\frac{1}{2}\theta^2 \phi'' = -c\phi. \tag{21}$$

It suffices to show that ϕ has a finite number of zeroes. In the study of second-order differential equations, it is common to use the Prufer substitution (Birkhoff and Rota (1989), page 312) to count the number of zeroes of a solution to a second-order differential equation. We recall the Prufer substitution:

$$\phi'(z) = r(z)\cos\alpha(z)\ ,\ \phi(z) = r(z)\sin\alpha(z) \tag{22}$$

where

$$(r(z))^2 = (\phi(z))^2 + (\phi'(z))^2\ \ ; \alpha(z) = \arctan\left(\frac{\phi(z)}{\phi'(z)}\right).$$

Obviously $r(z) = 0$ for a given z if and only if $\phi'(z) = 0 = \phi(z)$, which leads to a trivial solution $\phi(\cdot) \equiv 0$ for (21). Hence we can assume $r(z) > 0$ for all z. Then the second-order differential equation (21) is equivalent to the following system of first-order differential equations for (r, α) :

$$(\alpha(z))' = \frac{2c}{\theta^2(z)} \sin^2 \alpha(z) + \cos^2 \alpha(z) \tag{23}$$

$$(r(z))' = \frac{1}{2}\left[1 - \frac{2c}{\theta^2(z)}\right] r(z) \sin 2\alpha(z) \tag{24}$$

Notice that although $\arctan\left(\frac{\phi(z)}{\phi'(z)}\right)$ is not defined whenever $\phi'(z) = 0$, the Prufer system of first-order differential equations (23)-(24) are well-defined and have unique solution given any initial values say $(r(z_0), \alpha(z_0)) = (r_0, \alpha_0)$, which in turn defines a unique solution $\phi(z)$ for the second-order equation (21) via (22). Moreover, every non-trivial solution $\phi(\cdot)$ to (21) takes value zero at a point z ($\phi(z) = 0$) if and only if the solution $\alpha(\cdot)$ to (23) takes value $n\pi$ for some integer n at that point z, (i.e. $\sin \alpha(z) = 0$).

The equation (23) has a unique solution $\alpha(z)$ for any initial value say $\alpha(z_0) = a$, and the solution is an increasing (continuously differentiable) function. In particular, $\alpha(z)$ is bounded (above and below) over the compact set K. Thus ϕ has only a finite number of zeroes over the compact set K.

We now show that ϕ has at most finite many zeros outside the set K, (or equivalently, $\alpha(z)$ is bounded above and below outside the set K). Suppose ϕ has a zero to the right of K. (Otherwise the conclusion follows immediately.) Thus there exists a z^* to the right of K such that

$$\alpha(z^*) = n^*\pi \quad \text{for some integer } n^*.$$

Applying the Prufer substitution to inequality (3), and denote the corresponding new dependent variables as (r_{V+1}, α_{V+1}) (in particular $\alpha_{V+1}(z) = \arctan\left(\frac{V+1}{V'}\right)$). It may be shown that

$$(\alpha_{V+1}(z))' \geq \frac{2c}{\theta^2(z)} \sin^2 \alpha_{V+1}(z) + \cos^2 \alpha_{V+1}(z).$$

Since $V + 1$ never crosses the zero axis, the function $\alpha_{V+1}(z)$ can be initialized to be in the interval $(n^*\pi, (n^* + 1)\pi)$ for $z \geq z^*$. In particular,

$$\alpha_{V+1}(z^*) > n^*\pi = \alpha(z^*).$$

From the Comparison Theorem (*e.g.* see Birkhoff and Rota (1989), pages 29-31),

$$\alpha_{V+1}(z) \geq \alpha(z) \quad \text{for all } z \geq z^*$$

implying that

$$\alpha(z) < (n^* + 1)\pi \quad \text{for all } z \geq z^*$$

Hence ϕ has no zero values for all $z > z^*$. An analogous argument that studies the behavior of $\alpha(z), \phi(z)$ to the left of the set K.

Proof. (**Theorem 5.1**): We follow Lindvall (1983) and use a coupling argument. Consider two independent diffusions. One $\{z_t^1 : t \geq 0\}$ is initialized at z and the other is initialized according to the stationary distribution. We are interested in the stopping time $\tau \equiv \inf\{t \geq 0 : z_t^1 = z_t^2\}$. The probability distribution for z_t^1 and z_t^2 coincide from τ on. Define the conditional β-mixing coefficient:

$$\beta_t(z) \equiv \sup_{0 \leq \phi \leq 1} |E[\phi(z_t^1)|z_0^1 = z] - E\phi(z_t^2)|$$

Then from Lindvall (1983) (Section 2),

$$\beta_t(z) \leq \Pr\{\tau > t | z_0^1 = z\}.$$

As a consequence,

$$\beta_t \leq E[\beta_t(z_0^1)] \leq \Pr\{\tau > t\}.$$

To bound the tail probabilities of the hitting time τ, we follow Pitman (1974) and Lindvall (1983) by using a familiar inequality for nonnegative random variables. Suppose that $E\xi(\tau) < \infty$ for $\xi \geq 0$ non-decreasing on $[0, \infty)$. Then

$$\lim_{t \uparrow \infty} \xi(t) \Pr\{\tau > t\} = 0 \text{ hence } \lim_{\mathrm{t} \uparrow \infty} \xi(\mathrm{t}) \beta_{\mathrm{t}} = 0.$$

If further, ξ is absolutely continuous with respect to Lebesgue measure and has density ξ', then a simple integration-by-parts argument implies that

$$\int_0^\infty \xi'(t) \Pr\{\tau > t\} < \infty \text{ and hence } \int_0^\infty \xi'(\mathrm{t}) \beta_{\mathrm{t}} < \infty.$$

□

Theorem A.1. *Suppose that Assumptions 3.1, 3.2 and 7.1 are satisfied. If for some* $\eta \in (\frac{1}{2}, 1]$,

$$\begin{aligned} \liminf_{x \nearrow r} \frac{s\sigma}{|S|^\eta \lambda^{1/2}} &> 0 \\ \limsup_{x \searrow \ell} \frac{s\sigma}{|S|^\eta \lambda^{1/2}} &< 0, \end{aligned}$$

then: (i) Assumption 7.2 holds and $\{y_j^* : j = 0, 1, 2, ...\}$ *is stationary* $\beta-$*mixing for* $\eta > \frac{1}{2}$. *Denote*

$$\eta^* \equiv \sup\{\eta \in (\frac{1}{2}, 1] : \textit{inequalities (17) and (18) are satisfied}\}.$$

(ii) If $\eta^* = 1$, *then* $\{y_j^* : j = 0, 1, 2, ...\}$ *is* $\rho-$*mixing and* $\beta-$*mixing with exponential decay rates; (iii) if* $\eta^* \in (\frac{1}{2}, 1)$, *then* $\lim_{j \to \infty} j^\delta \beta_j = 0$ *for any* $\delta < \delta^*$ *where* $\delta^* = \frac{2\eta^* - 1}{2 - 2\eta^*}$.

Proof. (**Theorem A.1**): In what follows we let $\widehat{D}$ denote the domain of $\widehat{\mathcal{A}}$ constructed using the stationary distribution $\widehat{Q}$, and $\widehat{L}^2$ denote the space of functions with finite second moment (against $\widehat{Q}$).

(i) As long as $\eta > \frac{1}{2}$, Assumptions 3.1, 3.2 and 7.1 imply that the time-altered continuous time diffusion $\{\widehat{y}_t : t \geq 0\}$ is stationary, recurrent and aperiodic. Hence $\{y_j^* : j = 0, 1, 2, ...\}$ is still stationary, recurrent and aperiodic; hence it is $\beta - mixing$.

(ii) If $\eta = 1$ in inequalities (17) and (18), then $\{\widehat{y}_t : t \geq 0\}$ is $\rho - mixing$ and $\beta - mixing$ with exponential decay by Corollary 4.2.

Next, by the result of Banon (1977) (see also Hansen and Scheinkman (1995)), $\rho-$ *mixing* of $\{\widehat{y}_t : t \geq 0\}$ implies the existence of spectral gap of the negative semidefinite generator $\widehat{\mathcal{A}}$. That is, $\widehat{\mathcal{A}}$ satisfies

$$\int \phi \left(\widehat{\mathcal{A}}\phi\right) d\widehat{Q} \leq -\delta \int \phi^2 d\widehat{Q} \tag{25}$$

for all $\phi \in \widehat{Z}$ for some $\delta > 0$ where

$$\widehat{Z} = \left\{\phi \in \widehat{D} : \int \phi d\widehat{Q} = 0\right\}.$$

An implication of (25) is that

$$\int \phi \left(\mathcal{I} - \widehat{\mathcal{A}}\right) \phi d\widehat{Q} \geq (1 + \delta) \int \phi^2 d\widehat{Q}.$$

Therefore, by equation (16)

$$\int \phi \widehat{\mathcal{T}} \phi d\widehat{Q} \leq \frac{1}{1+\delta} \int \phi^2 d\widehat{Q}$$

for $\phi \in \widehat{Z}$. In other words, the conditional expectation operator $\widehat{\mathcal{T}}$ of the discrete time process $\left\{y_j^* : j = 0, 1, 2, ...\right\}$ is a strong contraction on $\phi \in \widehat{Z}$. Since $\widehat{D}$ is dense in $\widehat{L}^2$, it follows from Rosenblatt (1971) that $\left\{y_j^* : j = 0, 1, 2, ...\right\}$ is $\rho - mixing$ with exponential decay provided that inequality (25) is satisfied, which holds given inequalities (17) and (18) with $\eta = 1$.

Finally, by Theorem 3.6, the $\rho - mixing$ of $\{\widehat{y}_t : t \geq 0\}$ also implies that there exists a non-negative Lyapunov function $V \in C^2$ with $V \geq 1$, a compact set K, positive constants c and d such that:

$$\widehat{\mathcal{A}} V \leq -cV + d\mathbf{1}_K$$

Thus:

$$\left(\mathcal{I} - \widehat{\mathcal{A}}\right) V \geq (1 + c)V - d\mathbf{1}_K$$

Take inverses and obtain:

$$V \geq (1 + c) \left(\mathcal{I} - \widehat{\mathcal{A}}\right)^{-1} V - d\mathbf{1}_K,$$

or

$$\widehat{\mathcal{T}}V(y) \leq \frac{1}{1+c}V(y) + \frac{d}{1+c}\mathbf{1}_K.$$

We may now apply theorem 2.1 of Down et al. (1995) to justify that $\left\{y_j^* : j = 0, 1, 2, ...\right\}$ is $\beta - mixing$ with exponential decay.

(iii) If $\eta \in (\frac{1}{2}, 1]$ in inequalities (17) and (18), then by Theorem 5.2, $\{\widehat{y}_t : t \geq 0\}$ is $\beta - mixing$ with $\lim_{t\to\infty} t^\delta \beta_t = 0$. To establish the result for $\left\{y_j^* : j = 0, 1, 2, ...\right\}$, we apply the theorems 2.3 and 4.3 of Tuominen and Tweedie (1994), which is for discrete-time Markov processes. □

Proof. (**Theorem 8.6**): First we consider a closed quadratic form

$$\widetilde{f}(\phi, \psi) = \frac{1}{2}\int (\nabla\phi)'\Sigma(\nabla\psi) + \frac{1}{2}\int F\phi\psi$$

on the domain

$$\begin{aligned} D(\widetilde{f}) \;\; = \;\; & \{\psi \in L^2(leb) : \psi \text{ has a weak derivative and} \\ & \int F(\psi)^2 + \int (\nabla\psi)'\Sigma(\nabla\psi) < \infty\}. \end{aligned}$$

As shown in Chen et al. (2009) , under our assumption on Σ and F, the form $\widetilde{f}$ will be positive semidefinite because f is. Moreover, the spectrum of f is the same as that of $\widetilde{f}$.

Notice that

$$\widetilde{f}(\psi, \psi) \geq \frac{\epsilon}{2}\int (\nabla\psi)'(\nabla\psi) + \frac{1}{2}\int F\psi\psi \quad \text{for all } \psi \in D(\widetilde{f}). \tag{26}$$

The essential spectrum for the form $\tilde{f}$ is necessarily to the right of the essential spectrum for the form on the right-hand of (26). See Davies (1989) Section 1.1.11. The essential spectrum for the form on the right-hand side is in turn to the right of

$$\lim \inf_{|y|\to\infty} F(y)$$

which is positive by assumption. See exercise 8.2 in Davies (1995). Thus the spectrum for $\widetilde{f}$ is discrete to the right of zero and zero is not an accumulation point. Therefore Condition 8.4 is satisfied. □

Proof. (**Theorem 8.7**): We first construct a lower bound for the form $\tilde{f}$ in the same way as in the proof of Theorem 8.6 except that we use $\hat{F}$ in place of F. See Chen et al. (2009) for a detailed construction and justification for $\hat{F}$. The remainder of the proof follows from that of Theorem 8.6. □

References

Andersen, T., Bollerslev, T., Diebold, F. X., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of American Statistical Association 96, 42–55.

Banon, G., 1977. Estimation non paramétrique de densité de probabilité pour les processus de markov. Ph.D. thesis, Université Paul Sabatier, Toulouse, France.

Bhattacharya, R. N., 1982. On the functional central limit theorem and the law of the iterated logarithm. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 60, 185–201.

Birkhoff, G., Rota, G. C., 1989. Ordinary differential equations. Wiley & Sons, New York.

Bollerslev, T., Mikkelsen, H. O., 1996. Modeling and pricing long memory in stock market volatility. Journal of Econometrics 73, 151–184.

Bradley, R., 1986. Basic properties of strong mixing conditions. In: Eberlein, E., Taqqu, M. S. (Eds.), Dependence in probability and statistics. Birkhauser, Boston.

Carrasco, M., Chen, X., Hansen, L. P., 1999. Time deformation and dependence, unpublished: University of Montreal, University of Chicago and Yale University.

Chen, X., Hansen, L. P., Scheinkman, J., 2009. Principal components and long run implications of multivariate diffusions. Annals of Statistics forthcoming.

Clark, P. K., 1973. A subordinated stochastic process model with finite variance for speculative prices. Econometrica 41, 135–155.

Comte, F., Renault, E., 1996. Long memory continuous time models. Journal of Econometrics 73, 101–149.

Conley, T. G., Hansen, L. P., Luttmer, E. G. J., Scheinkman, J. A., 1997. Short-term interest rates as subordinated diffusions. The Review of Financial Studies 10, 525–577.

Davies, E. B., 1989. Heat kernels and spectral theory. Cambridge University Press, New York.

Davies, E. B., 1995. Spectral theory and differential operators. Cambridge University Press, Cambridge, UK.

Davydov, Y., 1973. Mixing conditions for markov chains. Theory of Probability and Its Applications, 312–328.

Diebold, F., Inoue, A., 2001. Long memory and regime switching. Journal of Econometrics 105, 131–159.

Ding, Z., Granger, C. W. J., Engle, R. F., 1993. A long-memory property of stock returns and a new model. Journal of Empirical Finance 1, 83–106.

Down, D., Meyn, S. P., Tweedie, R. L., 1995. Exponential and uniform ergodicity of markov processes. The Annals of Probability 23, 1671–1691.

Duffie, D., Glynn, P., 2004. Estimation of continuous-time markov processes sampled at random time intervals. Econometrica 72 (6), 1773–1808.

Ethier, S., Kurtz, T., 1986. Markov processes. Wiley & Sons, New York.

Fukushima, M., Oshima, Y., Takeda, M., 1994. Dirichlet forms and symmetric markov processes. de Gruyter Studies in Mathematics 19. Walter De Gruyter, Berlin.

Genon-Catalot, V., Jeantheau, T., Laredo, C., 2000. Stochastic volatility models as hidden markov models and statistical applications. Bernoulli 6 (6), 1051–1079.

Granger, C. W. J., Teräsvirta, T., 1999. A simple nonlinear time series model with misleading linear properties. Economics Letters 62, 161–165.

Hansen, L. P., Scheinkman, J., 1995. Back to the future: Generating moment implications for continuous-time markov processes. Econometrica 63, 767–804.

Hansen, L. P., Scheinkman, J., Touzi, N., 1998. Spectral methods for identifying scalar diffusions. Journal of Econometrics 86, 1–32.

Harrison, J. M., Kreps, D., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20, 381–408.

Harrison, J. M., Pitbladdo, R., Schaefer, S. M., 1984. Continuous price processes in frictionless markets have infinite variation. Journal of Business 57, 353–365.

Has'minskii, R. Z., 1980. Stochastic stability of differential equations. Sijthoff & Noordhoff, Rockville, MD.

Hidalgo, J., Robinson, P., 1996. Testing for structural change in a long-memory environment. Journal of Econometrics 70, 159–174.

Ibragimov, I. A., Rozanov, Y. A., 1978. Gaussian Random Processes. Springer-Verlag, New York.

Lindvall, T., 1983. On coupling of diffusion processes. Journal of Applied Probability 20, 82–93.

Maheswaran, S., Sims, C. A., 1993. Empirical implications of arbitrage-free asset markets. In: Phillips, P. C. B. (Ed.), Models, methods, and applications of econometrics: essays in honor of A. R. Bergstrom. Basil Blackwell, Cambridge, MA, pp. 301–316.

Mandelbrot, B. B., Ness, J. W. V., October 1968. Fractional brownian motions, fractional noises, and applications. SIAM Review 10 (4), 422–437.

Meyn, S. P., Tweedie, R. L., 1993. Stability of markovian processes iii: Foster-lyapunov criteria for continuous-time processes. Advances in Applied Probability 25, 518–548.

Nelson, D., 1990. Arch models as diffusion approximations. Journal of Econometrics 45, 7–39.

Pitman, J. W., 1974. Uniform rates of convergence for markov chain transition probabilities. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 29, 193–227.

Robinson, P. M., 1995. Log-periodogram regression for time series with long-range dependence. Annals of Statistics 23, 1048–1072.

Rockner, M., Wang, F.-Y., 2001. Weak poincare inequalities and **L2**-convergence rates of markov semigroups,. Journal of Functional Analysis 185 (2), 564 – 603.

Rosenblatt, M., 1971. Markov processes, structure and asymptotic behavior. Springer-Verlag, New York.

Tuominen, P., Tweedie, R. L., 1994. Subgeometric rates of convergence of f-ergodic markov chains. Advances in Applied Probability 26, 775–798.

Veretennikov, A. Y., 1987. Bounds for the mixing rate in the theory of stochastic equations. Theory of Probability and its Applications 32, 273–281.

Veretennikov, A. Y., 1997. On polynomial bounds for stochastic differential equations. Stochastic Processes and their Applications 70, 115–127.

Volkonskii, Y. A., Rozanov, Y., 1959. Some limit theorems for random functions i. Theory of Probability and its Applications 4, 178–197.

Weidmann, J., 1987. Spectral theory of ordinary differential operators. Springer-Verlag, New York.

Wong, E., 1964. The construction of a class of stationary markoff processes. In: Bellman, R. (Ed.), Sixteenth symposium in applied Mathematics-stochastic processes in mathematical physics and engineering. Americal Mathematical Society, Providence, RI, pp. 264–276.

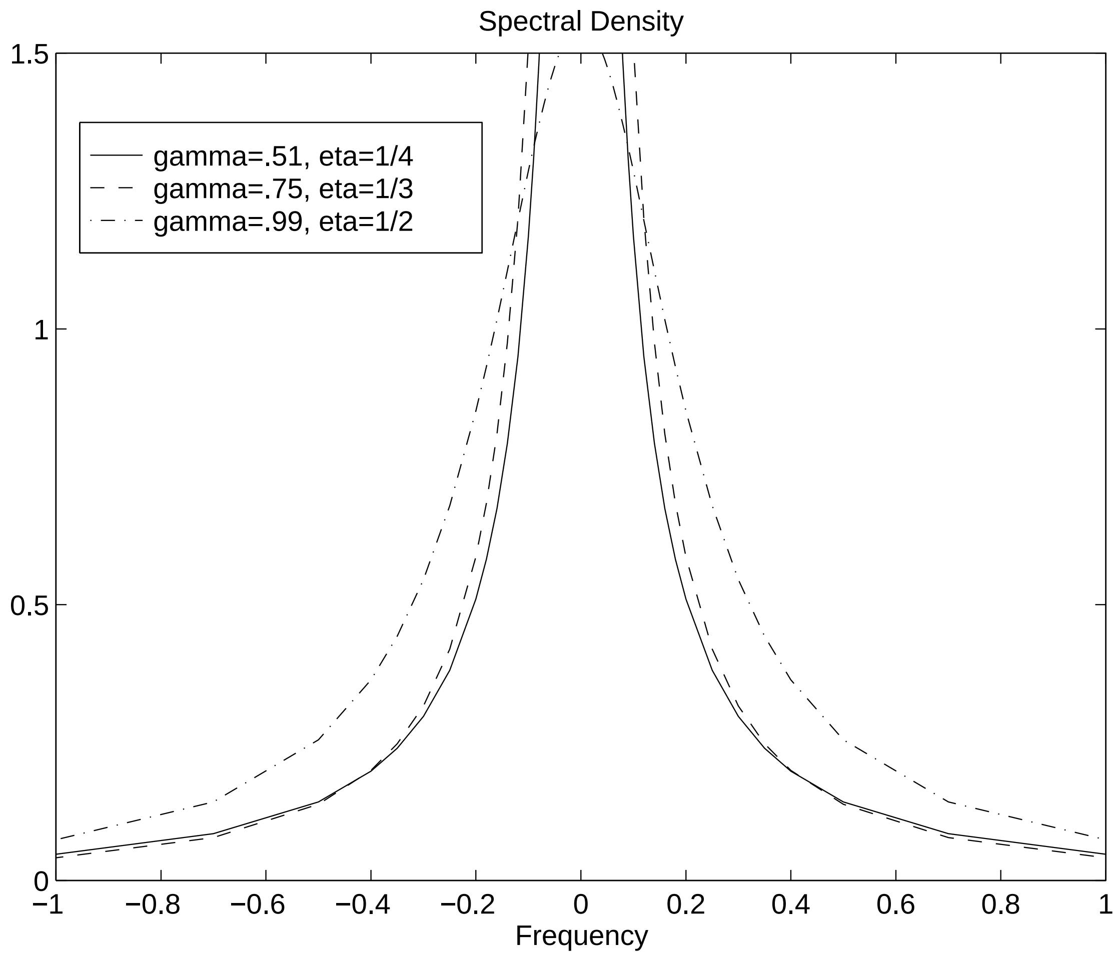

Figure 1: Spectral density functions for different pairs (γ, η). Spectral densities are rescaled to integrate to one.

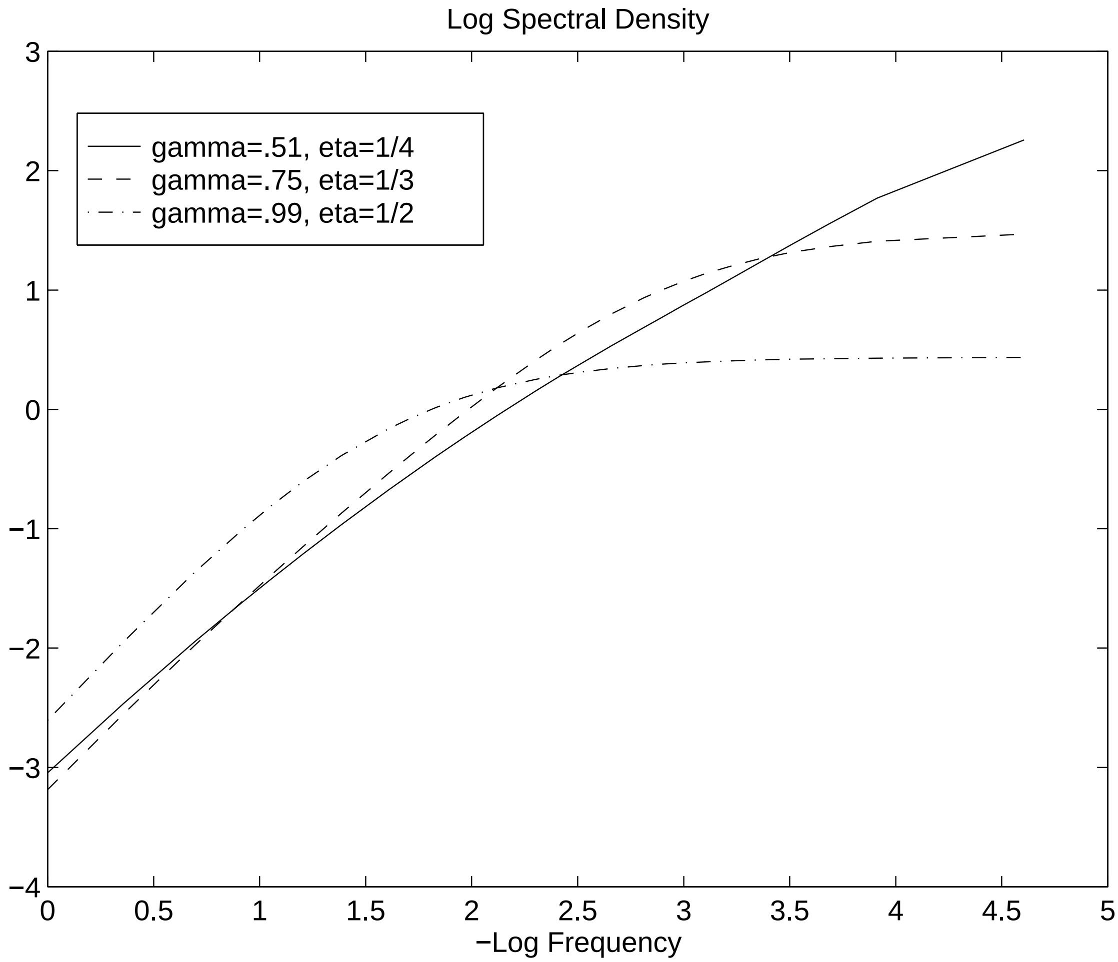

Figure 2: Spectral density functions for different pairs (γ, η) plotted on a $\log - \log$ scale. Spectral densities are rescaled to integrate to one.

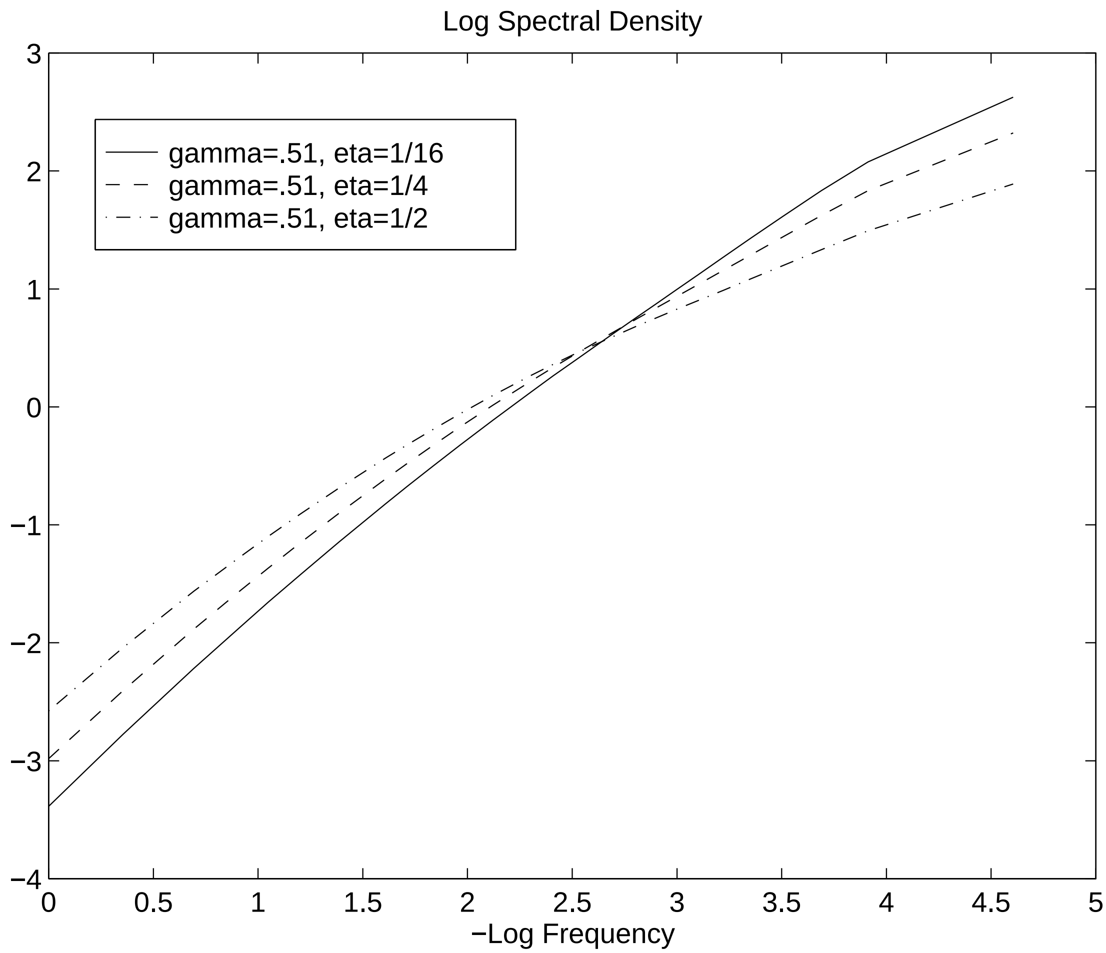

Figure 3: Spectral density functions for different values of η plotted on a $\log-\log$ scale. Spectral density functions are rescaled to integrate to one.