cemmap

centre for microdata methods and practice

Long memory via networking

Susanne M. Schennach

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP13/13

Long memory via networking

Susanne M. Schennach*
Brown University

This version: March 27, 2013
First draft: June 18, 2012.

Abstract

Many time-series data are known to exhibit "long memory", that is, they have an autocorrelation function that decays very slowly with lag. This behavior has traditionally been attributed to either aggregation of heterogenous processes, nonlinearity, learning dynamics, regime switching, structural breaks, unit roots or fractional Brownian motion. This paper identifies an entirely different mechanism for long memory generation by showing that it can naturally arise when a large number of simple linear homogenous economic subsystems with a short memory are interconnected to form a network such that the outputs of each of the subsystem are fed into the inputs of others. This networking picture yields a type of aggregation that is not merely additive, resulting in a collective behavior that is richer than that of individual subsystems. Interestingly, the long memory behavior is found to be almost entirely determined by the geometry of the network while being relatively insensitive to the specific behavior of individual agents.

1 Introduction

It widely recognized that many economic and financial time-series data exhibit "long memory", that is, posses an autocorrelation function that decays very slowly with lag (e.g., Mandelbrot and Ness (1968), Granger and Ding (1996), Comte and Renault (1996), Baillie (1996)). Explaining and modeling this feature has led to very active literature on structural breaks and/or regime switching (e.g., Diebold and Inoue (2001), Perron (1989), Perron and

*The author would like to thank Lars Hansen, Eric Renault and seminar participants at the Cemmap Workshop "Measurement Error and Related Topics", at the University of Cambridge and at the Harvard/MIT Econometrics seminar for useful comments and acknowledges support from NSF grant SES-1061263/1156347.

Qu (2007), Davidson and Sibbertsen (2005), Granger and Ding (1996)) unit roots (e.g. Hall (1978), Nelson and Plosser (1982), Perron (1988), Phillips (1987)) aggregation (e.g., Granger (1980), Zafaroni (2004), Abadir and Talmain (2002), Chambers (1998)) learning dynamics (e.g., Alfarano and Lux (2005), Chevillon and Mavroeidis (2011)) nonlinearity (e.g. Chen, Hansen, and Carrasco (2010), Miller and Park (2010)), fractional brownian motion (e.g. Mandelbrot and Ness (1968), Granger and Ding (1996), Comte and Renault (1996), Baillie (1996)), multifractal models (e.g., Calvet and Fisher (2002)), as well as other mechanisms (e.g. Parke (1999)). These approaches all identify plausible mechanisms generating a long memory behavior. The goal of this paper is to identify a different and arguably more universal mechanism that is active, regardless of the specific dynamic behavior of any one agent in the economy.

We demonstrates that long memory can naturally arise when a large number of simple linear homogenous economic subsystems with a short memory are interconnected to form a network such that the outputs of each of the subsystem are fed into the inputs of others. This networking picture yields a type of aggregation that is not merely additive, resulting in a collective behavior that is richer than that of individual subsystems. Interestingly, the long memory behavior is found to be almost entirely determined by the geometry of the network while being relatively insensitive to the specific behavior of individual agents.

Our results are the consequence of two main observations. First, we make a direct connection between a geometric property a network (namely, the fraction of pathways of a given length that reach a specific destination in the network) and the behavior of the spectral response of the network near the origin (specifically, its rate of divergence as frequency goes to zero). Second, we show how this geometric factor can be calculated for general classes of networks. In particular, drawing from the literature studying random walks on fractals (Havlin and Ben-Avraham (1987)), we show that a wide range of geometric factors can be naturally obtained, thus enabling us to generate power spectra matching the divergent behavior of any fractionally integrated process.

The network structure of an economy and its implications on aggregate fluctuations has received considerable and ongoing attention (e.g., Long and Plosser (1983), Horvath (1998), Dupor (1999), Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012)). However, this strand of literature centers on the conditions needed for preventing micro-level noise to simply average out in the aggregate and does not seek to explain long memory behavior. In fact, the question of aggregate fluctations can be studied entirely within a static model (e.g., Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012)), while investigating the long memory behavior clearly demands a dynamic model. Even existing models that do include dynamics (e.g., Long and Plosser (1983)) do not generate general classes of long memory behavior (except perhaps unit root-type behavior for specific values of the model parameters, although this is not discussed in Long and Plosser (1983)). These limitations are circumvented in the present paper by considering a general dynamic model in the limit of large networks characterized by scaling laws (including, but not limited to, fractal networks).

In this limit, the effect of network geometry on the small-frequency spectrum dominates the effect of individual subsystems, a feature that could not be captured by earlier network models. As such, we make a direct link between so far distinct literatures, the study of long memory and the study of economic networks structure.

2 Model

2.1 General ideas

We construct the generating process via a collection of elementary short-memory subsystems interconnected as a network (see Figure 1). Each subsystem takes a number of "input" variables as given (e.g. prices) and decides the value of output variables (e.g. quantity produced). The terms "input" or "output" do not necessarily refer to goods being purchased or sold. "Input" denotes information the system takes as given and cannot change while an "output" denotes variables the subsystem can decide and that provides information that can propagate to other subsystems. We place no fundamental restrictions in the direction of the flow of information (except when considering specific examples). If the "ouput" of subsytem A goes to subsystem B, the output could be sent to another subsystem C or to the same subsystem A.

In the absence of noise, this network is assumed to adopt a nonrandom steady-state equilibrium. This equilibrium may involve time-dependent quantities with deterministic trends. We then consider how this equilibrium is perturbed by introducing stationary short-memory noise at one point in the network (hereafter called the "origin") and by measuring its impact at another arbitrary point in the network (hereafter called the "destination"). The effect of noise at the origin on the destination can occur through a number of potential pathways of different geometries and lengths.

We consider networks consisting of linear subsystems (that is, their ouput is linear in the input history). If we further assume that the dynamic response of each system to noise is invariant with respect to time shifts, we can then model the response of each subsystem via a convolution operation.[1] Working in the linear limit not only makes the problem analytically tractable but also offers the advantage of illustrating that nonlinearity is not necessary to generate long memory within our framework. One can also interpret our linear approach as a linearization of the network's nonlinear subsystems that is justified in the limit of small noise.[2] A subsystem behavior that can be represented as a convolution arises naturally as the

[1]Satisfying this translation invariance assumption may involve working with some deterministic transformation of the model, e.g. woking with discounted present-value of monetary values or working with logarithms of some variables. Nevertheless, this assumption is plausible since we focus on deviations from any deterministic trends characterizing the equilibrium.

[2]Note that "small noise" is not incompatible with long memory since many long memory processes are still stationary and even more are mean-reverting (Baillie (1996)).

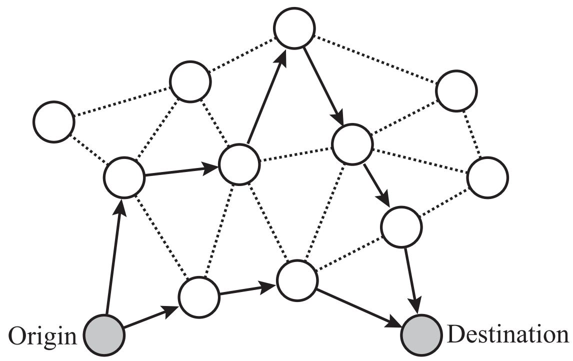

Figure 1: General ideas underying of the approach. Exogenous short-memory noise is fed in to a network of short-memory subsystems at one point called the “origin”. This noise is then propagated, through numerous paths of various geometries and lengths, to the “destination”. It is the sum of all of these indirect effects that generates the long memory property of the noise monitored at the “destination”.

solution to numerous utility maximization problems or as the linearization of such solutions around an equilibrium. A typical example can be found in Long and Plosser (1983), where the behavior of the various sectors of the economy are modeled by a vector autoregressive process in log levels that can easily be cast into a convolution operation, e.g., by calculating the impulse response function. Even some models of learning (Chevillon and Mavroeidis (2011)) take the form of convolutions.

Since our building blocks are stationary processes and translation-invariant operators, it is natural to state our results entirely in terms of spectral representations. This has the additional advantages that (i) all results can be stated in terms of deterministic functions and obtained via conventional algebraic manipulations and (ii) both discrete and continuous processes can be simultaneously handled transparently.

Mathematical technicalities aside, the general idea is to note that the exogenous noise Y fed into the network at the origin, after is has traversed n identical subsystems characterized by a convolution operation R, can be written in the form R^nY (where R^n denotes n application of the operation R). The operation R is a convolution with a rapidly decaying function (e.g., exponentially decaying tail or even compact support), so that each subsystem on its own has a “short” memory.

In general, there are many pathways of different geometries and lengths that connect the origin to the destination, so we sum over all possible pathway lengths, weighted by a geometric factor c_n that is related to the fraction of paths of length n starting at the origin that actually reach the destination:

$$\sum_{n=0}^{\infty} c_n R^n Y. \tag{1}$$

It turns out that the asymptotic decay of the coefficients c_n determines if the limiting process (1) has long memory for any sufficiently regular (and, in fact, short-memory) process Y and convolution R. More specifically, the asymptotic rate of decay of the coefficients c_n very directly determines the rate of divergence of the power spectrum at the origin. Power spectrum divergence is often considered as a signature of the various types of long memory process (see, e.g., Lobato and Robinson (1996), Baillie (1996), Granger and Ding (1996)). For instance, fractionally integrated processes are known to exhibit a power spectrum of the form $|\lambda|^{-2\alpha}$ near the origin, with α being a nonintegral positive number. To show how such spectra can be obtained, we calculate the asymptotic behavior for the c_n for simple periodic and finite networks. We also borrow powerful results from the literature studying random walks on fractals to show that one can obtain coefficients c_n with a wide range of behaviors, thus yielding power spectra exhibiting the same divergent behavior at the origin as a fractionally integrated process of any given order. We further show that our results are robust to various forms of heterogeneity.

2.2 Definitions and Preliminaries

Definition 1 *Let*

1. $\mathcal{L}_{12c}^{+}$ denote the set of all Lebesgue-measurable functions in $\mathcal{L}_1(\mathbb{R}^+) \cap \mathcal{L}_2(\mathbb{R}^+)$, endowed with the usual $\mathcal{L}_2$ norm.

2. $\mathcal{Y}_c$ denote the set of stochastic processes $Y(t)$ admiting a moving average representation

$$Y(t) = \int_{-\infty}^{t} y(t-s)\, dW(s)$$

where $W(s)$ is a standard Wiener process, $y \in \mathcal{L}_{12c}^{+}$, and where the equalities and integrals involving continuous-time stochastic processes are understood in the usual mean square sense (see Doob (1953), Chap. XI).

3. $\mathcal{R}_c$ be the set of linear operators R admitting the representation

$$[RY](t) = \int_{-\infty}^{t} r(t-s)\, Y(s)\, ds$$

with $r \in \mathcal{L}_{12c}^{+}$ for any $Y \in \mathcal{Y}_c$

For discrete processes, we have similar assumptions.

Definition 2 *Let*

1. $\mathcal{L}_{12d}^{+}$ *denote the set of all real-valued sequence (denoted* $Y(t)$ *with* $t = 0, 1, \ldots$*) that are absolutely summable (note that such sequences are automatically square summable as well),*[3] *endowed with the usual* ℓ_2 *norm.*

2. $\mathcal{Y}_d$ *denote the set of stochastic processes* $Y(t)$ *admitting a moving average representation*

$$Y(t) = \sum_{s=-\infty}^{t} y(t-s)\, G(s)$$

where $G(s)$ *are independent* $N(0,1)$ *random variables (indexed by* s*) and* $y \in \mathcal{L}_{12d}^{+}$.

3. $\mathcal{R}_d$ *be the set of linear operators* R *admitting the representation*

$$[RY](t) = \sum_{s=-\infty}^{t} r(t-s)\, Y(s)$$

with $r \in \mathcal{L}_{12d}^{+}$ *for any* $Y \in \mathcal{Y}_d$.

The requirement of both integrability and square integrability may be surprising, but both restrictions are important. On the one hand, square integrability is central to the theory of stochastic integrals (Doob (1953)). On the other hand, integrability is important because, by Young's inequality,[4] among the $\mathcal{L}_p$ spaces, only $\mathcal{L}_1$ has the property of being closed under convolution.

Although it is not necessary for the applicability of our approach, we focus on Gaussian processes for simplicity of exposition. To avoid duplication, we adopt the following conventions in the sequel:

Definition 3 *For a discrete process, the notation*[5] $\int_0^\infty \ldots dt$ *stands for* $\sum_{t=0}^{\infty} \ldots$*. Let* $(\mathcal{R}, \mathcal{Y}, \mathcal{L}_{12}^{+})$ *stand for either* $(\mathcal{R}_c, \mathcal{Y}_c, \mathcal{L}_{12c}^{+})$ *or* $(\mathcal{R}_d, \mathcal{Y}_d, \mathcal{L}_{12d}^{+})$.

The following Lemmas summarize well-known results from the theory of stochastic process (e.g., Doob (1953), Chap. XI, Section 9):

[3] Note that each element $y(t)$ of a sequence is bounded by $\|y\|_1$, the ℓ_1-norm of that sequence. Then note that $\sum_{t=0}^{\infty} |y(t)|^2 = (\|y\|_1)^2 \sum_{t=0}^{\infty} \left(\frac{|y(t)|}{\|y\|_1}\right)^2 \leq (\|y\|_1)^2 \sum_{t=0}^{\infty} \frac{|y(t)|}{\|y\|_1} = (\|y\|_1)^2 < \infty$.

[4] Young's inequality states that the convolution of a function in $\mathcal{L}_p$ and a function in $\mathcal{L}_q$ belongs to $\mathcal{L}_r$, with $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Requiring $r = p = q$ leaves 1 as the only possibility.

[5] A more sophisticated approach would have been to write all integrals as $\int \ldots d\mu(t)$ with μ set to eigher the Lebesgue measure or a sum point masses on all the integers, but this would have made the text less accessible.

Lemma 1 *If $Y \in \mathcal{Y}$ then it also admits a spectral representation $\tilde{y}(\lambda) \equiv \int_0^\infty y(t) e^{i\lambda t} dt$ and an associated power spectrum $|\tilde{y}(\lambda)|^2$. Moreover, $\tilde{y}(\lambda)$ is a bounded and square-integrable function defined for any $\lambda \in \mathbb{R}$. A corresponding result hold with $\mathcal{Y}, Y, y, \tilde{y}$ replaced by $\mathcal{R}, R, r, \tilde{r}$, respectively, with $\tilde{r}(\lambda) = \int_0^\infty r(t) e^{i\lambda t} dt$.*

For conciseness, we often call the "spectral representation" simply the "spectrum", reserving the term "power spectrum" for its modulus square.

Lemma 2 *Let $Y_0 \in \mathcal{Y}$ and let $Y_n = R_n \cdots R_1 Y_0$ with $R_1, \ldots, R_n \in \mathcal{R}$ for some $n \in \mathbb{N}$. Then $Y_n \in \mathcal{Y}$ and the spectral representation of these quantities are related through $\tilde{y}_n(\lambda) = \tilde{r}_n(\lambda) \cdots \tilde{r}_1(\lambda) \tilde{y}_0(\lambda)$.*

Note that Lemma 2 does not let us conclude that $\lim_{n\to\infty} Y_n \in \mathcal{Y}$. In fact, it is precisely the fact that $\lim_{n\to\infty} Y_n \notin \mathcal{Y}$ in general that allows us to consider long memory processes via a limiting process (since processes in $\mathcal{Y}$ necessarily have short memory).

Following standard practice (see, e.g., Lobato and Robinson (1996), Baillie (1996), Granger and Ding (1996)) we consider a divergence of the spectrum at the origin as a signature of a process exhibiting long memory. To circumvent well-known difficulties in defining the power spectrum of nonstationary processes (Mandelbrot and Ness (1968), Flandrin (1989), Loyne (1968)), we view a long memory process as a limiting case of a sequence of stationary processes. Accordingly, for a given sequence Z_n of stationary processes with corresponding well-defined power spectrum $|\tilde{z}_n(\lambda)|^2$, we study the behavior of $\lim_{n\to\infty} |\tilde{z}_n(\lambda)|^2 \equiv |\tilde{z}_\infty(\lambda)|^2$. We will consider that the sequence Z_n exhibits "long memory of order α" if $|\tilde{z}_\infty(\lambda)|^2$ behaves as $|\lambda|^{-2\alpha}$ as $\lambda \to 0$.

2.3 Main results

With these tools in place, we are ready to state our main assumptions and corresponding results.

Assumption 1 *$R \in \mathcal{R}$ with its associated $r(t)$ satisfying (i) $\int_0^\infty |r(t)| (1+t^2) dt < \infty$, (ii) $\int_0^\infty r(t) dt = 1$, (iii) $\int_0^\infty r(t) t dt \neq 0$ and (iv) $\int_0^\infty r(t) t^2 dt > \left(\int_0^\infty r(t) t dt\right)^2$.*

Assumption 1(i) is a standard constraint on the tail behavior of $r(t)$ that implies that the corresponding spectrum $\tilde{r}(\lambda)$ is twice continuously differentiable. Assumption 1(ii) is just a normalization, since deviations from it can be absorbed in multiplicative prefactors (denoted c_n below). Assumptions 1(iii) and (iv) are automatically satisfied if $r(t) \geq 0$ and $r(t)$ is supported on a nondegenerate interval (or more than one point in the discrete case), but hold more generally as well. Assumption 1(iv) implies that the spectrum $\tilde{r}(\lambda)$ does not exceed 1 in magnitude near the origin (and, in fact, can be replaced by that later condition without affecting the results).

Assumption 2 *$Y \in \mathcal{Y}$ with its associated $y(t)$ satisfying $\int_0^\infty y(t)\,dt \neq 0$.*

This assumption rules out the degenerate case where any divergence in the spectrum at the origin would be made impossible due to the fact that the input noise has no zero-frequency component.

Definition 4 *Let*

$$Z_{\bar{n}} = \sum_{n=0}^{\bar{n}} c_n R^n Y \tag{2}$$

where R satisfies Assumption 1 and Y satisfying Assumption 2.

Intuitively, Y is the exogenous noise input at the origin of the network, while $R^n Y$ is the effect of this noise after it has gone through n identical subsystems R of the network. We first consider a network of identical subsystems (although we shall consider extensions allowing for heterogeneity in Section 2.5) to emphasize the point that heterogeneity is not necessary to generate long memory.

Since we are interested in the overall effect of this noise at the origin on the destination via every possible pathway, we sum of pathways of any length n, after weighting (via c_n) proportionally to the effective number of pathways of length n and the "strength" of their contribution. As such, the sequence c_n may incorporate information about both the network topology and the overall average "gain" or "amplification" of the pathways of length n. We will now investigate the behavior of the power spectrum of $Z_{\bar{n}}$ as $\bar{n} \to \infty$ as a function of the asymptotic behavior of the sequence of weights c_n.

Theorem 1 *Let Assumptions 1-2 hold. Let $c_0 = 1$ and $c_n = n^{-(1-\alpha)}$ for $\alpha \in \,]-\infty, 1]$ and $n = 1, 2, \ldots$ (i) If $\alpha \geq 0$, then there exists a neighborhood $\mathcal{N}$ of the origin such that for all $\lambda \in \mathcal{N} \setminus \{0\}$, the power spectrum of $Z_{\bar{n}}$ satisfies*

$$\lim_{\bar{n} \to \infty} |\tilde{z}_{\bar{n}}(\lambda)|^2 \equiv |\tilde{z}_\infty(\lambda)|^2 = A\,|\lambda|^{-2\alpha} + o\left(|\lambda|^{-2\alpha}\right) \tag{3}$$

for some $A \in \mathbb{R} \setminus \{0\}$ (with the convention that $|\lambda|^{-2\alpha} \equiv |\ln|\lambda||^2$ for $\alpha = 0$) (ii) If $\alpha < 0$ (or, more generally, whenever $\sum_{n=0}^\infty |c_n| < \infty$), then $\lim_{\bar{n} \to \infty} |\tilde{z}_{\bar{n}}(\lambda)|^2 \equiv |\tilde{z}_\infty(\lambda)|^2 = A + o(1)$ for some $A \in \mathbb{R}$.

The proof of this Theorem, given in the Appendix, can be informally outlined as follows: The spectal representation of the series $\sum_{n=0}^\infty c_n R^n$ is $\sum_{n=0}^\infty c_n (\tilde{r}(\lambda))^n$, which is very closely related to a Taylor series of the function $(1-x)^{-\alpha}$, for our choice of sequence c_n. Since $\tilde{r}(\lambda) = 1 - C\lambda + o(\lambda)$ for some $C \neq 0$ under our assumptions, combining these results yields a spectral representation of the form $(1 - 1 + C\lambda + o(\lambda))^{-\alpha} = C^{-\alpha}\lambda^{-\alpha} + o\left(\lambda^{-\alpha}\right)$, i.e., a power spectrum of the form $|\lambda|^{-2\alpha}$.

The fact that the overall response is an infinite sum over the response of various path lengths makes it possible for the overall impulse response to not be summable, although individual agents have a summable impulse response function. Intuitively, long memory arises because each additional convolution lengthens the tail of the impulse response to the input noise. Note that the lengthing of the tail can occur even if the onset of the agents' response is instantaneous (i.e. the support of the function $r(\cdot)$ include 0). Of course, long memory cannot arise if the agents only have an instantaneous response, but that case is ruled out by Assumption 1(iii).

The assumed behavior for c_n may seem specific, but other natural possibilities either yield uninteresting or implausible results. One obvious generalization is $c_n = e^{\beta n} n^{-(1-\alpha)}$ for $\alpha, \beta \in \mathbb{R}$. However, the $\beta < 0$ case falls under case (ii) of Theorem 1 and yields a short memory process. The case $\beta > 0$ yields a spectrum that diverges at all λ such that $|\tilde{r}(\lambda)| > e^{-\beta}$ and not just at $\lambda = 0$. In that case, even a perturbation of a finite duration would magnified by the network to such an extent that the overall economy would leave the local equilibrium considered in a finite time and visit another equilibrium. The process would then presumably repeat itself until a stable equilibrium (with non-explosive c_n) is found. In a sense, the economy should plausibly self-organize to rule out cases where $\tilde{z}_\infty(\lambda)$ diverges for $\lambda \neq 0$. In this sense, $\beta = 0$ is the only nontrivial and plausible case.

One may be concerned that the long memory result of Theorem 1 appears to hold only under very specific circumstances. However, one has to realize that, realistically, the economy consists of many networks with different characteristics and the overall response to noise will be determined by a sum of the response of all of these networks. Perhaps many of them will have rapidly decaying c_n that yield short memory processes. But it only takes a few subsets of the networks with a slowly decaying c_n to obtain long memory behavior overall, since the long memory contributions will dominate all other short memory contributions at small λ. The following theorem makes this idea of negligibility of short memory processes relative to long memory processes more precise. It also shows that the conclusion of Theorem 1 is robust to various deviations from a specific power law behavior for the c_n.

Theorem 2 *Let Assumptions 1-2 hold. Let c_n and c'_n be two sequences such that $\sum_{n=0}^{\infty} |c_n - c'_n| < \infty$. Then, the corresponding $\tilde{z}_\infty(\lambda)$ and $\tilde{z}'_\infty(\lambda)$ are such that (i) $|\tilde{z}_\infty(\lambda) - \tilde{z}'_\infty(\lambda)|$ is continuous and uniformly bounded in a neighborhood $\mathcal{N}$ of the origin and (ii) whenever $|\tilde{z}_\infty(\lambda)| = A|\lambda|^{-2\alpha} + o\left(|\lambda|^{-2\alpha}\right)$ (for $A \in \mathbb{R}$ and $\alpha \in \mathbb{R}^+$) we also have $|\tilde{z}'_\infty(\lambda)|^2 = \tilde{A}|\lambda|^{-2\alpha} + o\left(|\lambda|^{-2\alpha}\right)$ for some $\tilde{A} \in \mathbb{R}$ (with $\tilde{A} = A$ if $\alpha > 0$).*

Since we have so far considered the effect of the introducing noise at one point of the network, it is natural to next consider what happens when multiple sources of noise are introduced at various points in the network. Let us consider independent noise input at various point of the network. This is without fundamental loss of generality since dependent noise could, in principle, simply be modeled by constructing a new network describing how a common shock is fed to different point of the orginal network. That is, for any network affected by dependent noise, there exists a bigger network (including the original network plus a parallel network propagating the noise to different points of the original network) with independent noise sources that is equivalent to it.

The power spectrum of a linear combination of multiple independent processes is simply a linear combination of the power spectrum of the individual processes (with squared weights). It follows that the result of Theorem 1 can be trivially extended to conclude that the power spectrum resulting from independent noise introduced at any number of points in the network would also behave as $|\lambda|^{-2\alpha}$ near the origin. We can also relax the assumption of a network of homogenous agents (i.e. sharing a common R). This extension is discussed in Section 2.5.

In the case where the limiting power spectrum $|\tilde{z}_\infty(\lambda)|^2$ is integrable (which occurs in Theorem 1 for $\alpha < 1/2$), we can also establishing a stronger form of convergence that implies the existence of a stationary limiting process $Z_\infty(t)$ with a power spectum of the form $|\lambda|^{-2\alpha}$ as $|\lambda| \to 0$.

Theorem 3 *Let the Assumptions of Theorem 1 hold. Let $\mathbb{L} = \mathbb{R}$ for continuous processes and $\mathbb{L} = [-\pi, \pi]$ for discrete processes and assume that $|\tilde{r}(\lambda)| < 1$ for $\lambda \in \mathbb{L}\backslash\{0\}$ and that $|\tilde{y}(\lambda)|$ is uniformy bounded for $\lambda \in \mathbb{L}$. If $\sum_{n=0}^{\infty} |c_n| < \infty$ or if $\alpha < 1/2$, there exists a stationary process $Z_\infty(t)$ with spectrum $\tilde{z}_\infty(\lambda) \equiv \lim_{n\to\infty} \tilde{z}_n(\lambda)$ and corresponding moving average representation $z_\infty(t)$ such that $\int_{\mathbb{L}} |\tilde{z}_n(\lambda) - \tilde{z}_\infty(\lambda)|^2 d\lambda \to 0$, $\int_0^\infty |z_n(t) - z_\infty(t)|^2 dt \to 0$ and $E\left[|Z_n(t) - Z_\infty(t)|^2\right] \to 0$ for almost any given $t \in \mathbb{R}$ and $\int_{-\infty}^{\infty} E\left[|Z_n(t) - Z_\infty(t)|^2\right] w(t)\, dt \to 0$ for a given absolutely integrable, bounded and continuous weighting function $w(t)$.*

One can also establish a similar convergence result that covers both integrable ($\alpha < 1/2$) and non integrable ($\alpha \geq 1/2$) limiting power spectra $|\tilde{z}_\infty(\lambda)|^2$ by focusing on increments of the processes. Working with increments is a standard technique (see Mandelbrot and Ness (1968) and Comte and Renault (1996), for instance) that offers the advantage of providing finite-variance quantities even in the presence of nonstationarity in the process.

Theorem 4 *Let the Assumptions of Theorem 1 hold. Let $\mathbb{L} = \mathbb{R}$ for continuous processes and $\mathbb{L} = [-\pi, \pi]$ for discrete processes and assume that $|\tilde{r}(\lambda)| < 1$ for $\lambda \in \mathbb{L}\backslash\{0\}$, that $|\tilde{y}(\lambda)|$ is uniformy bounded for $\lambda \in \mathbb{L}$ and consider the differenced process*

$$\Delta Z_n(t) \equiv Z_n(t) - Z_n(t - \Delta t)$$

for a given $\Delta t \in \mathbb{R}$ and any $n \in \mathbb{N}$ (with corresponding moving average representation $\Delta z_n(t) \equiv z_n(t) - z_n(t - \Delta t)$ and spectrum $\Delta \tilde{z}_n(\lambda) \equiv \left(1 - e^{i\lambda\Delta t}\right) \tilde{z}_n(\lambda)$). Then, there exists a stationary process $\Delta Z_\infty(t)$ with moving average representation[6] *$\Delta z_\infty(t) \equiv z_\infty(t) - z_\infty(t - \Delta t)$ and spectrum $\Delta \tilde{z}_\infty(\lambda) \equiv \left(1 - e^{i\lambda\Delta t}\right) \tilde{z}_\infty(\lambda)$ satisfying $\int_{\mathbb{L}} |\Delta \tilde{z}_n(\lambda) - \Delta \tilde{z}_\infty(\lambda)|^2 d\lambda \to$*

[6] We take the convention that $z_\infty(t) = 0$ for $t < 0$.

0, $\int_0^\infty |\Delta z_n(t) - \Delta z_\infty(t)|^2 \, dt \to 0$ *and* $E\left[|\Delta Z_n(t) - \Delta Z_\infty(t)|^2\right] \to 0$ *for almost any given* $t \in \mathbb{R}$ *and* $\int_{-\infty}^\infty E\left[|\Delta Z_n(t) - \Delta Z_\infty(t)|^2\right] w(t) \, dt \to 0$ *for a given absolutely integrable, bounded and continuous weighting function* $w(t)$.

Theorems 3 and 4 thus confirm that our approach does not merely amount to studying a limiting hypothetical case, but that the limit is reachable as an actual well-defined process. These theorems also show that our conclusions hold whether we take the limit as $n \to \infty$ before or after calculating the spectrum.[7]

2.4 Network models

An important observation is that if an "output" is sent to multiple subsystems of the network, its effect tends to be diluted by the number of recipients. For instance, if an agent experiences an exogenous increase in income, its overall expenditure and savings may increase in proportion to the increase, but such increases will be distributed across multiple categories of goods and services and the increase will be proportionally less in each specific category. Similarly, an exogenous increase in the availability of some material or goods will be spead over the multiple agents that make use of it and the effect on one agent (say, via a price change) will be reduced in proportion to the number of competing agents.

Iterating this argument leads to the conclusion that one should not obtain the coefficients c_n in Equation (2) by summing the *number* of pathways of length n connecting the origin and a given destination. Instead, one should obtain c_n by taking the number of pathways of length n connecting the origin and a given destination and normalizing it by the total number of pathways of length n reaching *any* destination. In other words, the weight c_n assigned to pathways of length n should not be related to the raw *number* of such pathways, but rather to the *fraction* of such pathways that actually reach a given destination, among all pathways of length n starting at a given point. If such a normalization were not performed, any noise introduced into the network would be endlessly exponentially magnified, yielding a completely unstable system — a clearly unrealistic setup. The observation that the coefficients c_n should represent the fraction of pathways of length n reaching a given destination therefore indicates that the coefficients c_n can be obtained by evaluating the probability that a random walk on the network (started at the origin) will land on a given node of the network after n steps.

Of course, the c_n could, in principle, also be affected by the "gain" of each subsystem. Thus, in general, c_n would not exactly equal the probability of reaching the destination. For simplicity, we do not explicitly consider such deviations in the present section. Small deviations could be straightforwardly accounted for via Theorem 2 and we consider larger deviations in Sections 2.5 and 3 below.

[7] The limit as $|\lambda| \to 0$ is always taken last, because this is the only way to study the *rate* of divergence of the spectrum near the origin.

2.4.1 Periodic and finite networks

We first consider random walks on a simple periodic lattice and calculate the probability of landing on a given node of the network after n steps, starting from the origin.

Theorem 5 *Let X_n and U_n be sequences of random variables taking values in $\mathbb{Z}^d$ ($d \in \mathbb{N}\backslash\{0\}$) and satisfying $X_{n+1} = X_n + U_n$, where $X_0 = 0$ and where U_n is iid with a distribution symmetric about $U_n = 0$, supported on a finite subset of $\mathbb{Z}^d$ and such that* $\text{Var}[U_n]$ *is positive definite. Then, for any $x \in \mathbb{Z}^d$, there exists a finite $C > 0$ such that*

$$P[X_n = x] = Cn^{-d/2} + O\left(n^{-1-d/2}\right).$$

If we apply this theorem with U_n uniformly distributed on a finite subset of $\mathbb{Z}^d$, we can conclude that the fraction of paths of length n reaching a given point in the network scales as $n^{-d/2}$. The form of this result could have been easily anticipated from the central limit theorem since, informally, the density of a sum near the origin would scale as $-d/2$. However some care must taken to ensure that the result holds for probabilities at a point for a discrete distribution and to secure a bound on the remainder term, which is shown to be always absolutely integrable so that it can only contribute a short-memory component to the process.

Interestingly, $d = 1$ (a linear network) gives us c_n scaling as $n^{-1/2}$ and therefore a long memory process of order $1 - 1/2 = 1/2$ by Theorem 1. Similarly $d = 2$ gives an order of $1-2/2 = 0$ (i.e. a spectrum with a logarithmic divergence at the origin). For $d = 3, 4, \ldots$ the sequence $n^{-d/2}$ is absolutely summable, so that no long memory results. The case $d = 0$ is treated separately below and can be interpreted as a random walk on a small finite network. In this case, we essentially obtain a constant c_n corresponding to the $\alpha = 1$ case of Theorem 1, a long memory process of order (i.e. a unit root). This occurs because the noise "feeds back" an infinite number of times into each subsystem and since the network is finite, none of the noise ever "diffuses" to infinity.

Theorem 6 *Let X_n be a sequence of random variables taking value in a finite set $\mathbb{F} \subset \mathbb{N}$ and such that for any $i, j \in \mathbb{F}$ and any $n \in \mathbb{N}$, we have (i) $P[X_{n+1} = j|X_n = i, X_{n-1}, \ldots, X_0] = P[X_{n+1} = j|X_n = i]$, (ii) $P[X_{n+1} = j|X_n = i] = P[X_{n+2} = j|X_{n+1} = i]$ and (iii) for any $i, j \in \mathbb{F}$, there exists $\Delta n \in \mathbb{N}$ such that $P[X_{n+\Delta n} = j|X_n = i] > 0$. Then, for any $x_0, x_1 \in \mathbb{F}$, there exists $\varepsilon \in]0, 1[$, a finite $C_0 > 0$ and $C_1, \ldots, C_{H-1} \in \mathbb{C}$ for some $H < \#\mathbb{F}$ such that*[8]

$$P[X_n = x_1|X_0 = x_0] = C_0 + \sum_{h=1}^{H-1} C_h e^{i2\pi nh/H} + O\left((1-\varepsilon)^n\right).$$

where the C_h for $h = 1, \ldots, H-1$ satisfy $C_h = C^*_{H-h}$, where $*$ denotes complex conjugation.

[8]The complex exponential term could be written in terms of sin and cos, to emphasize that the sum is real-valued, however this yields a cumbersome expression that depends on wether H is odd or even and that complicates subsequent proofs.

By Theorem, 1, the constant term yields a $|\lambda|^{-2}$ behavior for the spectrum while the $O\left((1-\varepsilon)^n\right)$ remainder only yields a short-memory contribution. The oscillatory terms of the form $e^{i2\pi nh/H}$ are not covered by Theorem 1 and demand a separate treatment. These oscilatory terms may appear if the network contains rings with unidirectional connections. Fortunately, as the following Theorem shows, these terms also do not affect the $|\lambda|^{-2}$ behavior of the power spectrum introduced by the constant term C_0.

Theorem 7 *Let Assumptions 1-2 hold. If* $c_n = \sum_{h=1}^{H-1} C_h e^{i2\pi nh/H}$ *(with* $C_h = C^*_{H-h}$*) then*

$$\lim_{\lambda \to 0} \left|\tilde{z}_\infty(\lambda)\right|^2 < \infty.$$

With these simple networks, we can already obtain common integrated and fractionally integrated processes, but there some unsatisfactory gaps in the exponents possible. Fortunately, any fractional power of frequency can be obtained by considering a more general class of network.

2.4.2 Fractal networks

Actual social or economic networks have been observed to exhibit some self-similarity across scales (Song, Havlin, and Makse (2005), Inaoka, Ninomiya, Taniguchi, and Takayasu (2004)), suggesting that fractals provide a useful model of such networks. Thanks to the fact that there is a direct relationship between random walks and the coefficients c_n for the network, we can borrow results from the literature studying random walks on fractals (e.g., Havlin and Ben-Avraham (1987)) to identify network geometries with associated power laws for diffusion behavior that generate coefficients c_n with almost any power-law behavior, thus leading, via Theorem 1, to long memory processes of any fractional order. This literature has observed that the probability of a random walker to visit a given point after n steps scales as $n^{-\gamma}$ asymptotically, where γ is some positive real number related to the geometry of the network. There is therefore a rather direct analogy with diffusion on periodic lattices in Euclian space. This observation comes from a combination of formal analytical treatments of various self-similar fractals (e.g. the well-known Sierpinski Gasket, see Figure 2a) as well as from thorough Monte Carlo simulations on random fractals (*statistically* self-similar fractals, e.g., see Figure 2b) guided by renormalization arguments (Havlin and Ben-Avraham (1987)). We outline these arguments heuristically below.

Interestingly, the exponent γ is not merely equal to $d_f/2$, where d_f is the fractal dimension of the network, as a direct analogy the periodic lattices would have suggested. Instead, $\gamma = d_f/d_w$, where d_w is an exponent characterizing the degree of so-called *anomalous* diffusion (Havlin and Ben-Avraham (1987)). For a conventional random walk $d_w = 2$ in any dimension, but d_w generally differs from 2 for fractal networks. The source of this distinction is that, a random walk on a fractal typically lacks a characteristic property of a conventional random walk: Two jumps (say U_i,U_j) are *not* necessarily such that $E\left[U_i U_j\right] = 0$ for $i \neq j$,

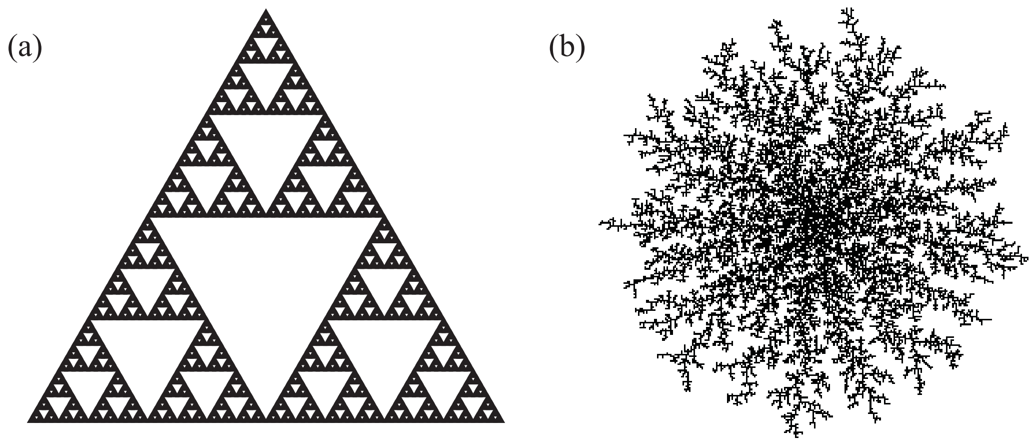

Figure 2: Examples of (a) a self-similar fractal, the well-known Sierpinski Gasket and of (b) a statistically self-similar fractal obtained by diffusion-limited aggregation (DLA).

thus leading to the failure of the usual conclusion that $\left(E\left[\|X_n\|^2\right]\right)^{1/2} = O\left(n^{1/2}\right)$. A power law of the form $\left(E\left[\|X_n\|^2\right]\right)^{1/2} = O\left(n^{1/d_w}\right)$ is nevertheless maintained, thanks to a fractal's self-similarity, albeit with a $d_w \neq 2$.

The scaling n^{-d_f/d_w} of the c_n coefficients can then be heuristically understood as follows. The (random) endpoint of an n step random walk has a distribution of a width of order $L = n^{1/d_w}$. If a steady-state distribution exists (up to an n-dependent scaling), and if it admits a density relative to the underlying fractal network of dimension d_f, this density must be normalized by a L^{-d_f} prefactor in order to have unit total probability at all n. Combining these two power laws leads to a scaling of n^{-d_f/d_w} for the density at a given point in the network.

It is well-known that one can construct subsets of D-dimensional space that are fractals of any fractal dimensions $d_f \in [0, D]$. This can be illustrated using, for instance, a self-similar fractal $\mathcal{F}$, i.e., one that has the property that

$$\mathcal{F} = \bigcup_{k=1}^{K} s_k\left(\mathcal{F}\right)$$

where $s_k\left(\mathcal{F}\right)$ denotes the set $\mathcal{F}$ transformed by the function $s_k\left(\cdot\right)$, which is an isotropic scaling (by a factor denoted by the scalar $|s_k|$), followed by a rotation and a translation. The fractal dimension of such a set, under an easy-to-satisfy "open set condition"[9] is given

[9] There exists a nonempty open set $\mathcal{S}$ such that $s_k\left(\mathcal{S}\right) \subseteq \mathcal{S}$ and $s_k\left(\mathcal{S}\right) \cap s_{k'}\left(\mathcal{S}\right) = \emptyset$ for $k \neq k'$ for $k, k' = 1, \ldots, K$.

by the solution d_f to the equation (Bandt and Graf (1992)):

$$1 = \sum_{k=1}^{K} |s_k|^{d_f} .$$

Taking, for instance, a fractal consisting of K copies of itself with identical scaling factor $|s_k| = s$ we obtain $1 = K s^{d_f}$, or

$$d_f = -\frac{\ln K}{\ln s},$$

from which it is clear that a continuum of values[10] of d_f can be obtained by suitable choices of K and $s \in [0, 1[$.

The above discussion, in conjunction with Theorem 1, then leads to the conclusion that the divergent spectrum characteristic of fractionally integrated processes of any order can be naturally obtained from the collective behavior of a population of linear homogenous agents interconnected through a (possibly) fractal network. Two interesting simple special cases stand out: The case of a power spectrum of the form $|\lambda|^{-2\alpha}$ with $\alpha = 1/2$ in fact does not require a fractal network at all: A simple linear chain suffices. Also, the unit root case just arises from a finite network.

2.5 Heterogeneity

To allow for heterogeneity, we consider a network with the following properties. Each node $i \in \mathbb{N}$ of the network is characterized by a *different* convolution operation R_i with spectral representation $\tilde{r}_i(\lambda)$. Even though a given node may have multiple inputs and outputs, we consider a single R_i per node. This assumes that the inputs have the same effect on the ouputs (which may be numerous, but identical). This assumption is actually without loss of generality because having a single node with multiple inputs with different effects on multiple outputs can merely be alternatively represented by multiple nodes with different R_i.

We view the R_i as being picked at random once at $t = -\infty$ and kept constant thereafter. We make no assumption regarding the covariance structure of $\tilde{r}_i(\lambda)$ between different λ. We do not require independence between the $\tilde{r}_i(\lambda)$ associated with different nodes i, although we do need to contraint the amount of dependence. This section provides conditions under which the conclusion of Theorem 1 actually holds with probability 1 for such randomly constructed networks. A key feature of the result is the existence of an average spectral representation $\bar{r}(\lambda)$. In essence, there are so many very long pathways that connect the origin and the destination, that the fluctuations in the $r_i(\lambda)$ across the nodes quickly average out to a

[10] Given this flexibility in d_f, the fact that fewer analytic results are known about the exponent d_w is not a real concern. Even if there were gaps in the possible values of d_w (and there is no evidence of it (Havlin and Ben-Avraham (1987))), it seems exceedingly unlikely that such gaps could not easily be "filled-in" by adjusting d_f so that any value of d_f/d_w in $[0, 1]$ can be reached.

single effective value $\bar{r}(\lambda)$ representative of the whole network. Stating this result requires us to be a bit more specific about the geometry of the network:

Definition 5 *Let $\mathcal{P}_{n,ij}$ denote the set of paths connecting node i to j in n steps (each element p of $\mathcal{P}_{n.ij}$ is a n-dimensional vector of integer specifying which sequence of nodes are visited by the path). Let $|\mathcal{P}_{n,ij}|$ denote the number of elements in $\mathcal{P}_{n,ij}$. Finally, for two given nodes i, j of the network and any $\bar{n} \in \mathbb{N}$, let*

$$Z_{\bar{n}} = \sum_{n=0}^{\bar{n}} c_n \left(\frac{1}{|\mathcal{P}_{n,ij}|} \sum_{p \in \mathcal{P}_{n,ij}} \prod_{l=1}^{n} R_{p_l} \right) Y_0 \tag{4}$$

where the inner parenthesis is conventionally 0 if $|\mathcal{P}_{n,ij}| = 0$ and where the product stands for the composition of multiple convolution operators.

This definition is necessary because each pathway of length n is associated with a different sequence of convolutions. Note that we average over paths rather than summing, in accordance with the discussion at begining of Section 2.4. Under the normalization convention used in Equation (4), the coefficients c_n still have the interpretation of the fraction of path of length n reaching the destination (because we divide by $|\mathcal{P}_{n,ij}|$, the number of pathways reaching the destination, not the total number of pathways of length n).

Theorem 8 *Let Y_0 satisfy Assumption 2 Let $\bar{r}(\lambda) \equiv \lim_{n\to\infty} \left(\frac{1}{|\mathcal{P}_{n,ij}|} \sum_{p \in \mathcal{P}_{n,ij}} E\left[\prod_{l=1}^{n} \tilde{r}_{p_l}(\lambda) \right] \right)^{1/n}$. Assume that $\bar{r}(\lambda)$ exists, satisfies Assumption 1 and is such that*

$$c_n^2 E\left[\left(\frac{1}{|\mathcal{P}_{n,ij}|} \sum_{p \in \mathcal{P}_{n,ij}} \left(\prod_{l=1}^{n} \frac{\tilde{r}_{p_l}(\lambda)}{\bar{r}(\lambda)} - 1 \right) \right)^2 \right] \leq D n^{-3-\varepsilon} \tag{5}$$

for some $D, \varepsilon > 0$ for all λ in some neighborhood of the origin. If $c_n = (\max\{1, n\})^{-(1-\alpha)}$ with $\alpha \in]-\infty, 1]$, then the conclusion of Theorem 1 holds with probability 1.

Condition (5) is stated in somewhat high-level form for maximum generality but it is relatively easy to realize that it is a weak restriction. This condition places a limit on the order of magnitude of the variance of a certain average. This average is taken over all possible pathways and effectively samples the spectral representation of a large number of nodes. Typically, the number $|\mathcal{P}_{n,ij}|$ of possible pathways of length n is an exponentially increasing function of n (because at each node there are certain number of possible ways to go and these alternative multiply to give the number of paths). Hence, unless the covariance of the summand across two pathways is extremely strong, we expect an exponential decrease of the variance of the average with n, which easily satifies the bound (5). The prefactor c_n is at worst polynomially increasing (in the cases of interest here, where the network is locally stable) and thus does not destroy the exponential rate. It is therefore natural to expect that the covariance of the summand across pathways should not to be an issue in most networks. Another key point is that (5) bounds the heterogeneity in the reponse of paths while placing only weak restrictions on the heterogeneity in the response of agents. For instance, one could argue that very large firms are very different from other, smaller, firms in the economy, thus introducing substantial heterogeneity in the agents' responses. However, it is likely the such large firms will constitute some of the links for most long paths, so that most long path are still quite similar to each other, although agents along the path are not.

3 Discussion

While our framework does not generate new classes of processes, it does provide a structural basis for popular models that have been proposed, such as fractionally integrated processes and their extensions (e.g., Mandelbrot and Ness (1968), Comte and Renault (1996), Baillie (1996), Comte and Renault (1998)). Nevertheless, in this section, we address a few potential objections to the framework proposed here.

One may argue that our results are in big part due to the assumption of an infinite network, while real networks are, in reality, finite. However, the behaviors of finite and infinite networks are similar in a way that makes then empirically difficult to distinguish for the following reason. It is obvious to see that convergence of the series $\sum_{n=0}^{\infty} c_n \left(\tilde{r}\left(\lambda\right)\right)^n$ (used in the proof of Theorem 1) occurs at an exponential rate for any λ such that $|\tilde{r}(\lambda)| < 1$. Hence, the series, truncated to a finite number of term, tends to be very close to its limiting value. A finite series represents a set of pathways that can comfortably fit within a finite network, so that a finite a network can closely emulate the behavior of an infinite one. Furthermore, we typically have $|\tilde{r}(\lambda)| < 1$ for a region of the form $|\lambda| \geq \varepsilon$ for any $\varepsilon > 0$ (our Assumption 1 implies this at least within a neighborhood of the origin, but it is often true over the whole spectrum). A region of the form $|\lambda| \geq \varepsilon$ is precisely the only portion of the spectrum that is empirically accessible, since the necessarily finite duration of recorded time series limits the smallest frequency for which the spectrum can be reliably determined.

One may also be tempted to conclude that our findings do not necessarily add much to the existing literature attempting to explain the occurence of fractionally integrated processes. Notably, it is known that fractionally integrated processes can arise from additive aggregation of an infinite number of heterogenous times series following more elementary processes (Granger (1980)). However, in this picture, long memory *only* arises if some of the original time series arbitrarily closely approach a unit-root process. That is, one needs to assume almost long memory behavior at the individual level to obtain collective long memory behavior. Moreover, the generation of a fractionally integrated process in this picture is the consequence of an assumed fractional power law behavior in the density of the autoressive parameter in the population. In contrast, in our framework, all subsystems on their own have short memory and none of them even approach a long memory behavior. Long memory only results from the network structure and fractional exponents arise from geometric arguments. Also, in our approach, heterogeneity is not necessary to yield long memory (although it is not inconsistent with it, since our results are robust to heterogeneity).

Although Theorem 1 focus on pointwise convergence of the spectrum, one can easily show that convergence is, in fact, uniform over $\mathcal{N}\setminus[-\varepsilon,\varepsilon]$ for any $\varepsilon>0$. It is not possible to show uniform convergence over a set of frequency containing the origin because the limiting spectrum is discontinuous — it is not a limitation of our approach.

One may also wonder what happens if the coefficients c_n do not follow a power law. Theorem 2 already handles "small" (i.e. absolutely summable) deviations from power laws. More generally, it is possible to extend Theorem 2 to show that the $|\lambda|^{-2\alpha}$ behavior can be robust to deviations from $n^{-(1-\alpha)}$ in the coefficients c_n than are bigger than absolutely summable. For instance, consider the case where the c_n (for $n\geq 1$) admit an expansion of the form

$$c_n=\sum_{i=1}^{\bar{\imath}}A_i n^{-(1-\alpha_i)}+c_n' \tag{6}$$

where $\alpha_1>\alpha_2>\cdots>\alpha_{\bar{\imath}}$ and $\sum_{n=1}^{\infty}|c_n'|<\infty$. One can apply Theorem 1 to each individual term of the expansion and Theorem 2 to remainder c_n' to yield the conclusion that the resulting power spectrum $|\tilde{z}_\infty(\lambda)|^2$ would then have the behavior

$$|\tilde{z}_\infty(\lambda)|^2=\sum_{i=1}^{\bar{\imath}}O\left(|\lambda|^{-2\alpha_i}\right)=O\left(|\lambda|^{-2\alpha_1}\right)\text{ as }|\lambda|\to 0,$$

since $\alpha_1>\alpha_i$ for $i=2,\ldots,\bar{\imath}$. Taking $\bar{\imath}$ finite is without much loss of generality, since eventually, for some α_i, the power law would become absolutely integrable (if consecutive exponents α_i are at least some finite distance from each other).

One could argue that our result are in part driven by the normalization of the spectrum at the origin $\tilde{r}(0)=1$ and our focus on the geometric origin of the coefficients c_n. If $\tilde{r}(\lambda)$ is everywhere (including at the origin) less than 1 in magnitude, then a short memory behavior invariably results. However, our results are still applicable even if many subsystems have the property $|\tilde{r}(\lambda)|<1$, as long as some subsystems to have the property $\tilde{r}(0)=1$. In that case, we can "prune" the network from its $|\tilde{r}(\lambda)|<1$ subsystems (since they would only contribute to a short-memory response) and focus on the "backbone" of the network containing the subsystems with $\tilde{r}(0)=1$. The fractional exponent γ (see Section 2.4) would then simply be the value characterizing the fractal structure of the backbone rather than the entire network. It should also be noted that the case $\tilde{r}(0)=1$ is not at all pathological: It merely corresponds to the case where the output is of the same magnitude as the input in the limit of slow variation in input (it is **not** a "unit root"). It is therefore a plausibly common occurence.

One may argue that our network models and associated spectral representations are too simple and regular to have anything to do with real economic networks. While real networks are admittedly more complex and less regular than assumed here, we feel that obtaining formal asymptotic results is important to illustrate that simple ingredients and a simple derivation can yield complex long memory behavior. Without this relative simplicity, one would have to rely on extensive numerical analysis of realistic complex networks and attempt to make a convincing case that the numerically obtained power spectra "look like" the ones of fractionally integrated processes.

4 Empirical Evidence

One way to empirically assess if the proposed mechanism for long memory generation is plausible is to verify if the c_n coefficients in a real economic network indeed obey a power law. For this purpose, we use the so-called "input-output accounts" database compiled by the Bureau of Economic Analysis describing interactions between ~ 460 sectors of the US economy in 2010. We construct the network following the same procedure as in Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012), considering the existence of a "link" between two sectors of the economy if one sector constitutes more than 5% of the inputs (in monetary value) of the other sector.

The c_n coefficients can be directly calculated via Monte Carlo by generating a large number of random walks on that network and by keeping track of the fraction of those walks that visit a given point of the network as a function of path length n. In our application, we simply keep track of the fraction of random walk that return to their starting point in n steps as function of n. At each n, we generate 10000 random walks with random starting points to obtain Figure 3.

This exercise reveals evidence of a power law $c_n = n^{-\gamma}$ with an exponent of $\gamma \approx 0.84$ (with a standard deviation of 0.07) that persists over an order of magnitude, as obtained with a standard linear least squares regression of the data in logarithmic form. This corresponds to a power spectrum behaving as $|\lambda|^{-0.32}$ near the origin, resulting in network behavior that is between white noise and a random walk. This finding supports the plausibility of the long memory generation mechanism we propose.

Although this is, strictly speaking, a finite network, one can still observe a behavior that would be expected from an infinite network for "short" paths, because "short" paths do not "feel" the boundary of the network. Here, we are considering paths up to a length of 50, which indicates these "short" paths can in fact be fairly long without seeing signs of the finiteness of the network. Of course, if we increased the cut off (to $n \approx 75$, in this example), the graph would flatten out, as would be expected for a finite network (since the c_n would be asymptotically constant in that case). As explained at the beginning of Section 3, since the convergence of the power spectrum is exponential for any frequency outside of

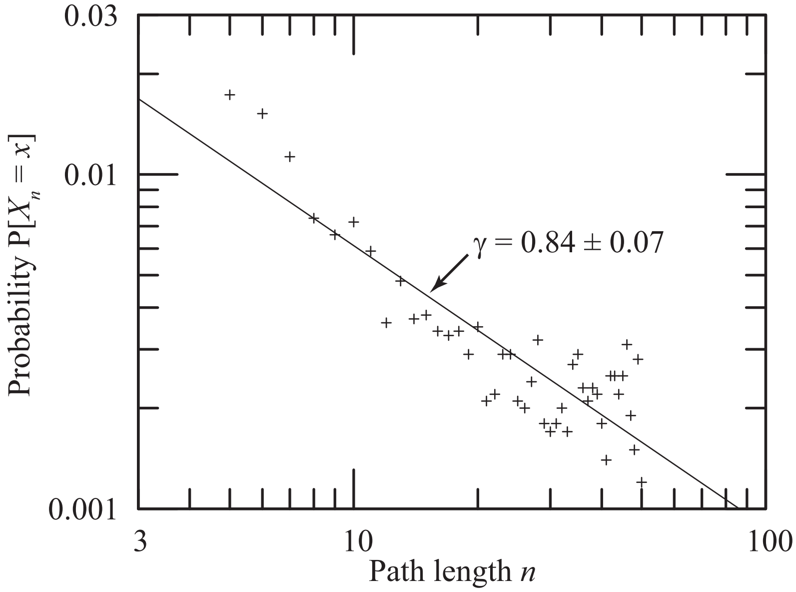

Figure 3: Evidence of power law scaling $n^{-\gamma}$ in the c_n coefficients (i.e. the probability of reaching a given point x of the network after n steps of a random walk X_n) in a network representing the US economy as ~ 460 interconnected sectors.

a small neighborhood of the origin, obtaining a power law behavior for small n is sufficient to emulate the response of an infinite network. Another potential concern is the presence of fluctuations around the power law that appear large and roughly constant over path length n. However, since the figure is plotted on a logarithmic scale, deviations that appear constant in magnitude are in fact exponentially decaying in n and are therefore summable, so that they would not affect the power law behavior of the spectrum near the origin.

5 Conclusion

While numerous plausible mechanisms generating a long memory behavior have been proposed, this paper identifies a different and arguably more universal mechanism that may be active, regardless of the specific dynamic behavior of any one agent in the economy. We show that long memory can naturally arise when a large number of simple linear homogenous economic subsystems with a short memory are interconnected to form a network such that the outputs of each of the subsystem are fed into the inputs of others. This networking picture yields a type of aggregation that is not merely additive, resulting in a collective behavior that is richer than that of individual subsystems. Interestingly, the long memory behavior is found to be almost entirely determined by the geometry of the network while being relatively insensitive to the specific behavior of individual agents. Specifically, we find that, under weak regularity conditions, the power spectrum of the network's response to exogenous short-memory noise can mimic a fractionally integrated processes $I(d)$, with d related to the scaling properties of the network (e.g. its fractal dimension).

As such, this work not only provides a plausible structural model for the generation of fractionally integrated long memory processes but also demonstrates that long memory is possible without nonlinearity, heterogeneity, unit roots or near unit roots, learning or structural breaks (although these mechanisms can obviously play a role as well). The proposed approach also make a direct connection between the literatures focusing on long memory processes, economic networks, diffusion on fractals, and applications of fractals in economics.

A Proofs

Definition 6 *To avoid ambiguities due to the multivalued nature of the fractional power function, we define:*

$$(i\lambda)^\alpha \equiv \begin{cases} |\lambda|^\alpha e^{\alpha i\pi/2} & \text{if } \lambda > 0 \\ |\lambda|^\alpha e^{-\alpha i\pi/2} & \text{if } \lambda < 0 \end{cases}$$

Moreover, the following convention is useful to avoid special cases: If $\alpha = 0$, then

$$(i\lambda)^\alpha \equiv -\ln(i\lambda) \equiv \begin{cases} \ln|\lambda| + i\pi/2 & \text{if } \lambda > 0 \\ \ln|\lambda| - i\pi/2 & \text{if } \lambda < 0 \end{cases}$$

Lemma 3 *Assumption 1 implies that (i) for some finite $A, B \in \mathbb{R}^+ \setminus \{0\}$, $\tilde{r}(\lambda) = 1 + Ai\lambda + o(\lambda)$ and $|\tilde{r}(\lambda)|^2 = 1 - B\lambda^2 + o(\lambda^2)$ as $\lambda \to 0$.and (ii) there exists a neighborhood $\mathcal{N}$ of the origin such that $|\tilde{r}(\lambda)| < 1$ for all $\lambda \in \mathcal{N} \setminus \{0\}$.*

Proof. Assumption 1(i) implies that $\tilde{r}(\lambda)$ is everywhere twice continuously differentiable. Thus, in particular, near the origin, we have the expansion $\tilde{r}(\lambda) = A_0 - A_1 i\lambda - \frac{1}{2}A_2\lambda^2 + o(\lambda^2)$ with A_0, A_1, A_2 finite. Assumption 1(i) also implies that the moment theorem applies up to order 2, so that $A_j = \int_0^\infty t^j r(t)\,dt$. By Assumption 1(ii) $A_0 = 1$. Since $r(t)$ is real, the real part of $\tilde{r}(\lambda)$ is symmetric while its imaginary part is anti-symmetric. Therefore, A_1 and A_2 must be real. Assumption 1(iii) implies that $A_1 \in \mathbb{R} \setminus \{0\}$ and the first conclusion of the lemma follows. Next, we note that $|\tilde{r}(\lambda)|^2 = \left(1 - \frac{1}{2}A_2\lambda^2\right)^2 + A_1^2\lambda^2 + o(\lambda^2) = 1 - A_2\lambda^2 + \frac{1}{4}A_2^2\lambda^4 + A_1^2\lambda^2 + o(\lambda^2) = 1 - (A_2 - A_1^2)\lambda^2 + o(\lambda^2)$, where $A_2 - A_1^2 > 0$ by assumption 1(iv). It follows that $|\tilde{r}(\lambda)| < 1$ in some neighborhood of the origin. ■

Lemma 4 *Assumption 2 implies that $\tilde{y}(\lambda) = B + o(1)$ for some $B \in \mathbb{R} \setminus \{0\}$.*

Proof. Assumption 2 requires that $y \in \mathcal{Y}$, which implies that the Fourier transform $\tilde{y}(\lambda)$ is continuous, thus implying an expansion of the form $B + o(1)$. Moreover $y(t)$ is real, so $B = \tilde{y}(0)$ is real as well and nonzero by Assumption 2. ■

Proof of Theorem 1. By Assumption 1 and Lemma 3, (i) for some finite $A \neq 0$, $\tilde{r}(\lambda) = 1 - Ai\lambda + o(\lambda)$ as $\lambda \to 0$.and (ii) there exists a neighborhood $\mathcal{N}$ of the origin such that $|\tilde{r}(\lambda)| < 1$ for all $\lambda \in \mathcal{N} \setminus \{0\}$. Also, by Lemma 4, $\tilde{y}(\lambda) = B + o(1)$ as $\lambda \to 0$ for some $B \in \mathbb{R} \setminus \{0\}$.

Consider first the special case[11] $\alpha = 1$, so that $c_n = 1$. By Lemma 2, the spectrum of $R^n Y$ is given by $(\tilde{r}(\lambda))^n \tilde{y}(\lambda)$ and thus the spectrum of Z_n is $\sum_{m=0}^{n} (\tilde{r}(\lambda))^m \tilde{y}(\lambda)$ (and the corresponding power spectrum is $|\sum_{m=0}^{n} (\tilde{r}(\lambda))^m \tilde{y}(\lambda)|^2$). For all $\lambda \in \mathcal{N} \setminus \{0\}$, the series $\sum_{m=0}^{\infty} (\tilde{r}(\lambda))^m \equiv \lim_{n\to\infty} \sum_{m=0}^{n} (\tilde{r}(\lambda))^m$ is convergent because $|\tilde{r}(\lambda)| < 1$ and we can directly evaluate this geometric series:

$$\begin{aligned}
\tilde{z}_\infty(\lambda) &= \tilde{y}(\lambda) \sum_{n=0}^{\infty} (\tilde{r}(\lambda))^n = \tilde{y}(\lambda) \frac{1}{1 - \tilde{r}(\lambda)} \\
&= (B + o(1)) \frac{1}{1 - 1 + Ai\lambda + o(\lambda)} = (B + o(1)) \frac{1}{Ai\lambda + o(\lambda)} \\
&= B(1 + o(1)) \frac{A^{-1}}{i\lambda} \frac{1}{1 + o(\lambda)/\lambda} = B \frac{A^{-1}}{i\lambda} \frac{1 + o(1)}{1 + o(1)} \\
&= \frac{BA^{-1}}{i\lambda}(1 + o(1)) = \frac{BA^{-1}}{i\lambda} + o\left(\lambda^{-1}\right)
\end{aligned}$$

Next, we consider the more general cases where $\alpha \in]0, 1[$. Consider the Taylor series $C(1-x)^{-\alpha} = \sum_{n=0}^{\infty} c'_n x^n$ for $|x| < 1$ for any nonzero constant C, where

$$c'_n = C \frac{1}{n!} \prod_{j=1}^{n} (\alpha + j - 1)$$

(with $c'_0 \equiv C$ by convention) and note that, for $\lambda \in \mathcal{N} \setminus \{0\}$,

$$\begin{aligned}
\tilde{z}'_\infty(\lambda) &\equiv \tilde{y}(\lambda) \sum_{n=0}^{\infty} c'_n (\tilde{r}(\lambda))^n = C\tilde{y}(\lambda)(1 - \tilde{r}(\lambda))^{-\alpha} = C(B + o(1))(1 - 1 + Ai\lambda + o(\lambda))^{-\alpha} \\
&= C(B + o(1)) A^{-\alpha} (i\lambda + o(\lambda))^{-\alpha} = CBA^{-\alpha}(i\lambda)^{-\alpha}(1 + o(1))(1 + o(\lambda)/\lambda)^{-\alpha} \\
&= CBA^{-\alpha}(i\lambda)^{-\alpha}(1 + o(1)) = \frac{CBA^{-\alpha}}{(i\lambda)^\alpha} + o\left(|\lambda|^{-\alpha}\right)
\end{aligned}$$

There remains to show that c_n is sufficiently close to c'_n so that $\tilde{z}_\infty(\lambda)$ has the same asymptotic behavior as $\tilde{z}'_\infty(\lambda)$. By Theorem 2 (below), it is sufficient to show that $\sum_{n=0}^{\infty} |c_n - c'_n| < \infty$. To this effect, note that

$$c'_n = C \prod_{j=1}^{n} \frac{(\alpha + j - 1)}{j} = C \prod_{j=1}^{n} \left(1 - \frac{\bar{\alpha}}{j}\right)$$

[11] This case could be combined with the more general case $\alpha \in]0, 1[$ below, but this simple case illustrates the idea of the proof with the least technical complications.

where $\bar{\alpha} \equiv 1 - \alpha$. Let $a_n = \ln\left(n^{\bar{\alpha}} c'_n / C\right)$ and observe that

$$
\begin{aligned}
a_n &= \bar{\alpha} \ln n + \sum_{k=1}^{n} \ln\left(1 - \frac{\bar{\alpha}}{k}\right) \\
&= \ln(1-\bar{\alpha}) + \bar{\alpha} \ln n + \sum_{k=2}^{n} \ln\left(1 - \frac{\bar{\alpha}}{k}\right) \\
&= \ln(1-\bar{\alpha}) + \bar{\alpha} \sum_{k=2}^{n} \left(\ln k - \ln(k-1)\right) + \sum_{k=2}^{n} \ln\left(1 - \frac{\bar{\alpha}}{k}\right) \\
&= \ln(1-\bar{\alpha}) - \bar{\alpha} \sum_{k=2}^{n} \ln \frac{k-1}{k} + \sum_{k=2}^{n} \ln\left(1 - \frac{\bar{\alpha}}{k}\right) \\
&= \ln(1-\bar{\alpha}) - \sum_{k=2}^{n} \bar{\alpha} \ln\left(1 - \frac{1}{k}\right) + \sum_{k=2}^{n} \ln\left(1 - \frac{\bar{\alpha}}{k}\right) \\
&= \ln(1-\bar{\alpha}) + \sum_{k=2}^{n} \left(\ln\left(1 - \frac{\bar{\alpha}}{k}\right) - \bar{\alpha} \ln\left(1 - \frac{1}{k}\right)\right) \qquad (7)
\end{aligned}
$$

Note that since $\ln(1-x) = -x - \frac{1}{2}x^2 + O\left(x^3\right)$ as $x \to 0$, the summand in (7) is such that

$$
\begin{aligned}
\ln\left(1 - \frac{\bar{\alpha}}{k}\right) - \bar{\alpha} \ln\left(1 - \frac{1}{k}\right) &= -\frac{\bar{\alpha}}{k} - \frac{1}{2}\left(\frac{\bar{\alpha}}{k}\right)^2 - \bar{\alpha}\left(-\frac{1}{k} - \frac{1}{2}\frac{1}{k^2}\right) + O\left(k^{-3}\right) \\
&= -\frac{\bar{\alpha}}{k} - \frac{\bar{\alpha}^2}{2} k^{-2} + \frac{\bar{\alpha}}{k} + \frac{\bar{\alpha}}{2} k^{-2} + O\left(k^{-3}\right) \\
&= \frac{\bar{\alpha}(1-\bar{\alpha})}{2} k^{-2} + O\left(k^{-3}\right)
\end{aligned}
$$

Since k^{-2} is a summable sequence, it follows that the series (7) converges, i.e. $a_\infty \equiv \lim_{n\to\infty} a_n$ is well-defined and finite. We can also conclude that

$$
\begin{aligned}
a_n - a_\infty &= \sum_{k=n+1}^{\infty} \left(\ln\left(1 - \frac{\bar{\alpha}}{k}\right) - \bar{\alpha} \ln\left(1 - \frac{1}{k}\right)\right) \\
&= \sum_{k=n+1}^{\infty} \left(\frac{\bar{\alpha}(1-\bar{\alpha})}{2} k^{-2} + O\left(k^{-3}\right)\right) \\
&\leq \int_{n}^{\infty} \left(\frac{\bar{\alpha}(1-\bar{\alpha})}{2} k^{-2} + O\left(k^{-3}\right)\right) dk \\
&= O\left(n^{-1}\right).
\end{aligned}
$$

Now, set the constant $C = \exp(-a_\infty)$ and consider $c_n = n^{-\bar{\alpha}}$. We have

$$\begin{aligned}
c'_n - c_n &= c'_n - n^{-\bar{\alpha}} \\
&= n^{-\bar{\alpha}} \left(C c'_n n^{\bar{\alpha}} / C - 1 \right) \\
&= n^{-\bar{\alpha}} \left(C \exp \ln \left(c'_n n^{\bar{\alpha}} / C \right) - 1 \right) \\
&= n^{-\bar{\alpha}} \left(C \exp(a_n) - 1 \right) \\
&= n^{-\bar{\alpha}} \left(\exp(a_n - a_\infty) - 1 \right) \\
&= n^{-\bar{\alpha}} \left(\exp\left(O\left(n^{-1}\right)\right) - 1 \right) \\
&= n^{-\bar{\alpha}} \left(1 + O\left(n^{-1}\right) - 1 \right) \\
&= n^{-\bar{\alpha}-1}
\end{aligned}$$

Since $\sum_{n=1}^{\infty} n^{-\bar{\alpha}-1} < \infty$, we have $\sum_{n=1}^{\infty} |c'_n - c_n| < \infty$ and the result follows.

For $\alpha = 0$, consider $-\ln(1-x) = \sum_{n=1}^{\infty} c'_n x^n$ with $c'_n = \frac{1}{n}$ for $n \geq 1$ and $c'_0 = 0$. Note that, for $\lambda \in \mathcal{N} \setminus \{0\}$,

$$\begin{aligned}
\tilde{z}'_\infty(\lambda) &\equiv \sum_{n=0}^{\infty} c'_n \left(\tilde{r}(\lambda)\right)^n = -\ln\left(1 - \tilde{r}(\lambda)\right) = -\ln\left(1 - 1 + Ai\lambda + o(\lambda)\right) = -\ln\left(Ai\lambda + o(\lambda)\right) \\
&= -\ln\left(Ai\lambda\left(1 + o(1)\right)\right) = -\ln(Ai\lambda) + \ln(1 + o(1)) \\
&= -\ln(Ai\lambda) + o(1) = -\ln(i\lambda) - \ln(A) + o(1) = -\ln(i\lambda) + O(1) = -\ln(i\lambda) + o(|\ln|\lambda||)
\end{aligned}$$

The same conclusion holds for $\tilde{z}_\infty(\lambda)$ since c'_n and c_n differ only for $n = 0$, implying that $\sum_{n=0}^{\infty} |c_n - c'_n| < \infty$ and enabling the use of Theorem 2.

We now consider the final case where either $\alpha < 0$ or $\sum_{n=0}^{\infty} |c_n| < \infty$. Since $\alpha < 0$ implies $\sum_{n=0}^{\infty} |c_n| < \infty$, we focus on the latter condition. Conclusion (i) of Theorem 2 with $c'_n = 0$ delivers the desired result: $\tilde{z}_\infty(\lambda) = B + o(1)$. ■

Proof of Theorem 2. Let $\Delta c_n \equiv c_n - c'_n$ and let $\Delta \tilde{z}_n(\lambda) = \tilde{z}_n(\lambda) - \tilde{z}'_n(\lambda)$ denote the corresponding spectrum. To prove the result, we exploit the fact that a uniformly convergent sequence of continuous functions converges to a continuous function. Since, by Assumption 1 and Lemma 3, $|\tilde{r}(\lambda)| \leq 1$ for $\lambda \in \mathcal{N}$, some neighborhood of the origin and since $\sum_{n=0}^{\infty} |\Delta c_n| < \infty$ by assumption, we can write, for $\lambda \in \mathcal{N}$,

$$|\Delta \tilde{z}_{\bar{n}}(\lambda) - \Delta \tilde{z}_\infty(\lambda)| = \left| \sum_{n=\bar{n}+1}^{\infty} \Delta c_n \left(\tilde{r}(\lambda)\right)^n \right| \leq \sum_{n=\bar{n}+1}^{\infty} |\Delta c_n| \, |\tilde{r}(\lambda)|^n \leq \sum_{n=\bar{n}+1}^{\infty} |\Delta c_n| \to 0$$

as $\bar{n} \to \infty$. Therefore, $\Delta \tilde{z}_{\bar{n}}(\lambda)$ converges uniformly to $\Delta \tilde{z}_\infty(\lambda)$ as $\bar{n} \to \infty$ for $\lambda \in \mathcal{N}$. This, combined with the fact that $\Delta \tilde{z}_{\bar{n}}(\lambda)$ is continuous in λ for any finite $\bar{n}$ and $\lambda \in \mathcal{N}$ (since it is a finite sum of continuous functions) implies that $\Delta \tilde{z}_\infty(\lambda)$ is continuous in $\mathcal{N}$ and we also have $\Delta \tilde{z}_\infty(\lambda) = B + o(1)$ as $\lambda \to 0$. It follows that, for $\alpha \geq 0$, and as $\lambda \to 0$,

$$\tilde{z}'_\infty(\lambda) = \tilde{z}_\infty(\lambda) + \Delta \tilde{z}_\infty(\lambda) = A|\lambda|^{-\alpha} + o\left(|\lambda|^{-\alpha}\right) + B + o(1) = \tilde{A}|\lambda|^{-\alpha} + o\left(|\lambda|^{-\alpha}\right)$$

for some finite nonzero $\tilde{A}$ (that equals A if $\alpha > 0$). $\blacksquare$

Proof of Theorem 3. First note that $|\tilde{r}(\lambda)| < 1$ for $\lambda \in \mathbb{L} \backslash \{0\}$ implies that $\tilde{z}_n(\lambda) \to \tilde{z}_\infty(\lambda)$ pointwise for any $\lambda \in \mathbb{L} \backslash \{0\}$, since

$$\begin{aligned} |\tilde{z}_n(\lambda) - \tilde{z}_\infty(\lambda)| &= \left| \tilde{y}(\lambda) \sum_{m=n+1}^{\infty} c_m \tilde{r}(\lambda)^m \right| \leq |\tilde{y}(\lambda)| \sum_{m=n+1}^{\infty} |c_m| |\tilde{r}(\lambda)|^m \\ &\leq |\tilde{y}(\lambda)| \left(\sup_m |c_m| \right) \sum_{m=n+1}^{\infty} |\tilde{r}(\lambda)|^m = \left(\sup_m |c_m| \right) |\tilde{y}(\lambda)| (1 - |\tilde{r}(\lambda)|)^{-1} |\tilde{r}(\lambda)|^{n+1}, \end{aligned}$$

where $|\tilde{r}(\lambda)|^{n+1} \to 0$ as $|\tilde{r}(\lambda)| < 1$ for $\lambda \in \mathbb{L} \backslash \{0\}$ and where all the prefactors are finite by assumption.[12]

The proof then proceeds by first showing that $\int_{\mathbb{L}} |\tilde{z}_\infty(\lambda)|^2 d\lambda < \infty$, thus implying that $\int_0^\infty |z_\infty(t)|^2 dt < \infty$, which in turn implies, that there exists some stationary process $Z_\infty(t)$ with moving average representation $z_\infty(t)$ and with spectrum $\tilde{z}_\infty(\lambda)$. Then, we show that there exists some $\bar{z}(\lambda)$ also satisfying $\int_{\mathbb{L}} (\bar{z}(\lambda))^2 d\lambda < \infty$ such that

$$|\tilde{z}_n(\lambda) - \tilde{z}_\infty(\lambda)|^2 \leq (\bar{z}(\lambda))^2$$

for all n, so that, by Lebesgue dominated convergence theorem, $\lim_{n\to\infty} \int_{\mathbb{L}} |\tilde{z}_n(\lambda) - \tilde{z}_\infty(\lambda)|^2 d\lambda = \int_{\mathbb{L}} \lim_{n\to\infty} |\tilde{z}_n(\lambda) - \tilde{z}_\infty(\lambda)|^2 d\lambda = 0$. This implies that $\int_0^\infty |z_n(t) - z_\infty(t)|^2 dt \to 0$, from which the mean square convergence of $Z_n(t)$ to $Z_\infty(t)$ follows by standard arguments (e.g., Doob (1953), Chap. XI, Section 9).

The $\sum_{n=0}^\infty |c_n| \equiv C_1 < \infty$ case (including the $\alpha < 0$ case) is simple:

$$\begin{aligned} |\tilde{z}_\infty(\lambda)| &\leq |\tilde{y}(\lambda)| \sum_{n=0}^{\infty} |c_n| |\tilde{r}(\lambda)|^n \leq |\tilde{y}(\lambda)| \sum_{n=0}^{\infty} |c_n| 1^n = |\tilde{y}(\lambda)| C_1 \equiv \bar{z}(\lambda) \\ |\tilde{z}_n(\lambda) - \tilde{z}_\infty(\lambda)| &= \left| \tilde{y}(\lambda) \sum_{m=n+1}^{\infty} c_m \tilde{r}(\lambda)^m \right| \leq |\tilde{y}(\lambda)| \sum_{m=n+1}^{\infty} |c_m| |\tilde{r}(\lambda)|^m \\ &\leq |\tilde{y}(\lambda)| \sum_{m=0}^{\infty} |c_m| |\tilde{r}(\lambda)|^m \leq |\tilde{y}(\lambda)| C_1 \equiv \bar{z}(\lambda) \end{aligned}$$

where $\int_{\mathbb{L}} |\tilde{y}(\lambda)|^2 d\lambda < \infty$.

For the $\alpha \in]0, 1/2[$ case, we consider some small cutoff $\bar{\lambda} > 0$ and compute a separate bound for large ($|\lambda| \geq \bar{\lambda}$) and small ($|\lambda| \leq \bar{\lambda}$) frequencies.

[12] Note that if there existed sequence m_n such that $|c_{m_n}| \to \infty$, then we would have $\sum_{n=0}^\infty |c_n| \geq \sum_{n=0}^\infty |c_{m_n}| \to \infty$. Having $c_n = n^{-(1-\alpha)}$ with $a < 1/2$ also rules out $|c_{m_n}| \to \infty$.

To find a bound on $|\tilde{r}(\lambda)|$ for $|\lambda| \geq \bar{\lambda}$, we note that, by Assumption 1, and Lemma 3, $|\tilde{r}(\lambda)|^2 = 1 - C_2\lambda^2 + o\left(\lambda^2\right)$ for some $C_2 > 0$ as $\lambda \to 0$ and thus

$$|\tilde{r}(\lambda)| \leq 1 - C_3\lambda^2 \tag{8}$$

for some $C_3 \in \left]0, C_2/2\right[$ for all $|\lambda| \leq \bar{\lambda}$ sufficiently small. We can then show that for $\bar{\lambda}$ sufficiently small, the maximum of $|\tilde{r}(\lambda)|$ over the set $\left[\bar{\lambda}, \infty\right[\cap \mathbb{L}$ is reached at $\lambda = \bar{\lambda}$. In the case where $\mathbb{L}$ is infinite, we can rule out a maximum "at infinity" as follows. We must have $\limsup_{|\lambda|\to\infty} |\tilde{r}(\lambda)| = 0$, because any other limiting value, combined with the fact that $\tilde{r}(\lambda)$ has a uniformly bounded first derivative (by Assumption 1(i)) would imply that $\int_{\mathbb{L}} |\tilde{r}(\lambda)|^2 d\lambda$ diverges (contradicting that $\tilde{r} \in \mathcal{L}_2(\mathbb{L})$). Since the maximum cannot be at infinity, we can limit the search to bounded sets. But the maximum of $|\tilde{r}(\lambda)|$ in any set of the form $\left[\bar{\lambda}, \bar{\lambda}^*\right]$ for $\bar{\lambda}, \bar{\lambda}^* > 0$ is reached at some λ^*, by compactness of the set and continuity of $\tilde{r}(\lambda)$ (by Assumption 1(i)) and by Assumption (iv), $\tilde{r}(\lambda^*) < 1$. Such a $\tilde{r}(\lambda^*)$ would eventually be exceed by $\left|\tilde{r}\left(\bar{\lambda}\right)\right|$ for $\bar{\lambda}$ sufficiently small since $\tilde{r}\left(\bar{\lambda}\right) \to 1$ as $\bar{\lambda} \to 0$. This contradiction is avoided only if $\lambda^* = \bar{\lambda}$ for all $\bar{\lambda}$ sufficiently small. Hence $|\tilde{r}(\lambda)| \leq 1 - C_3\bar{\lambda}^2$ for $|\lambda| \geq \bar{\lambda}$ for sufficiently small $\bar{\lambda}$.

Letting $\bar{\alpha} = 1 - \alpha$, we can then write, for $|\lambda| \geq \bar{\lambda}$,

$$\begin{aligned}
|\tilde{z}_\infty(\lambda)| &= |\tilde{y}(\lambda)| \left| 1 + \sum_{m=1}^{\infty} m^{-\bar{\alpha}} (\tilde{r}(\lambda))^m \right| \leq |\tilde{y}(\lambda)| \left(1 + \sum_{m=1}^{\infty} m^{-\bar{\alpha}} |\tilde{r}(\lambda)|^m \right) \\
&\leq |\tilde{y}(\lambda)| \left(\sum_{m=0}^{\infty} \left(1 - C_3\bar{\lambda}^2\right)^m \right) = \frac{|\tilde{y}(\lambda)|}{1 - \left(1 - C_3\bar{\lambda}^2\right)} \\
&= \frac{|\tilde{y}(\lambda)|}{C_3\bar{\lambda}^2} \leq C_4 |\tilde{y}(\lambda)| \equiv \bar{z}(\lambda)
\end{aligned}$$

and

$$\begin{aligned}
|z_n(\lambda) - z_\infty(\lambda)| &= |\tilde{y}(\lambda)| \left| \sum_{m=n+1}^{\infty} m^{-\bar{\alpha}} (\tilde{r}(\lambda))^m \right| \leq |\tilde{y}(\lambda)| \sum_{m=n+1}^{\infty} m^{-\bar{\alpha}} |\tilde{r}(\lambda)|^m \\
&\leq |\tilde{y}(\lambda)| \sum_{m=1}^{\infty} |\tilde{r}(\lambda)|^m \leq |\tilde{y}(\lambda)| \sum_{m=0}^{\infty} \left(1 - C_3\bar{\lambda}^2\right)^m \\
&= |\tilde{y}(\lambda)| \frac{1}{1 - \left(1 - C_3\bar{\lambda}^2\right)} = \frac{|\tilde{y}(\lambda)|}{C_3\bar{\lambda}^2} \leq C_4 |\tilde{y}(\lambda)| \equiv \bar{z}(\lambda)
\end{aligned}$$

for some $C_3, C_4 > 0$ and where $\tilde{y}$ is such that $\int_{|\lambda|\geq\bar{\lambda}} |\tilde{y}(\lambda)|^2 d\lambda \leq \int |\tilde{y}(\lambda)|^2 d\lambda < \infty$ since $\tilde{y} \in \mathcal{L}_2(\mathbb{R})$ because $y \in \mathcal{L}_2(\mathbb{R}^+)$.

For $|\lambda| \leq \bar{\lambda}$, since $\tilde{z}_\infty(\lambda) = C|\lambda|^{-\alpha} + o\left(|\lambda|^{-\alpha}\right)$, we have

$$|\tilde{z}_\infty(\lambda)| \leq C_4 |\lambda|^{-\alpha}$$

which satisfies $\int_{|\lambda|\leq\bar{\lambda}} |\lambda|^{-2\alpha} d\lambda < \infty$ for $\alpha \in [0, 1/2[$. Also, since $\tilde{r}(\lambda) = 1 + Ai\lambda + o(\lambda)$ (from Lemma 3), we have, by Lemma 5 (below),

$$\begin{aligned}
|\tilde{z}_n(\lambda) - \tilde{z}_\infty(\lambda)| &= |\tilde{y}(\lambda)| \left| \sum_{m=n+1}^{\infty} m^{-\bar{\alpha}} (\tilde{r}(\lambda))^m \right| = |\tilde{y}(\lambda)| |\tilde{r}(\lambda)|^n \left| \sum_{m=1}^{\infty} (m+n)^{-\bar{\alpha}} (\tilde{r}(\lambda))^m \right| \\
&\leq |\tilde{y}(\lambda)| |\tilde{r}(\lambda)|^n \sum_{m=1}^{\lceil C_5/|\lambda| \rceil} (m+n)^{-\bar{\alpha}} |\tilde{r}(\lambda)|^m \leq |\tilde{y}(\lambda)| |\tilde{r}(\lambda)|^n \sum_{m=1}^{\lceil C_5/|\lambda| \rceil} (m+n)^{-\bar{\alpha}} |\tilde{r}(\lambda)|^m \\
&\leq \frac{\pi}{2} \sum_{m=1}^{\lceil C_5/|\lambda| \rceil} m^{-\bar{\alpha}} \leq 2 \left(1 + \int_1^{C_5/|\lambda|} m^{-\bar{\alpha}} d\lambda \right) = \frac{\pi}{2} \left(1 + \left[m^{1-\bar{\alpha}} \right]_1^{C_5/|\lambda|} \right) \\
&= \frac{\pi}{2} \left(1 + \left(\frac{C_5}{|\lambda|} \right)^{1-\bar{\alpha}} - 1 \right) = \frac{\pi}{2} \left(\frac{C_5}{|\lambda|} \right)^{1-\bar{\alpha}} = C_6 |\lambda|^{-\alpha}
\end{aligned}$$

for some finite $C_5, C_6 > 0$. Hence, we can set $\bar{z}(\lambda) = C_6 |\lambda|^{-\alpha}$ for $|\lambda| \leq \bar{\lambda}$, which is square integrable over $|\lambda| \leq \bar{\lambda}$ for $\alpha \in [0, 1/2[$. ■

Lemma 5 *Let $r_m \to 0$ be a real, positive and decreasing sequence and $\theta \in [-\pi, \pi]$, then for any $n \in \mathbb{N}$,*

$$\left| \sum_{m=1}^{n} r_m e^{i\theta m} \right| \leq \frac{\pi}{2} \sum_{m=1}^{\bar{m}} r_m,$$

where $\bar{m} \equiv \lceil 2\pi/|\theta| \rceil$ (where $\lceil \cdot \rceil$ denotes the "round up" operation).

Proof. Let $s(t) = \frac{i\theta}{1-e^{-i\theta}} \int_0^t r_{\lceil\tau\rceil} e^{i\theta\tau} d\tau$ for $t \in \mathbb{R}^+$ and note that $s(n)$ for $n \in \mathbb{N}^*$ matches the partial sum $\sum_{m=1}^{n} r_m e^{i\theta m}$:

$$\begin{aligned}
s(n) &= \frac{i\theta}{1-e^{-i\theta}} \int_0^n r_{\lceil\tau\rceil} e^{i\theta\tau} d\tau = \frac{i\theta}{1-e^{-i\theta}} \sum_{m=1}^{n} \int_{m-1}^{m} r_{\lceil\tau\rceil} e^{i\theta\tau} d\tau \\
&= \frac{i\theta}{1-e^{-i\theta}} \sum_{m=1}^{n} r_m \int_{m-1}^{m} e^{i\theta\tau} d\tau = \frac{i\theta}{1-e^{-i\theta}} \sum_{m=1}^{n} r_m \frac{e^{i\theta m} - e^{i\theta(m-1)}}{i\theta} = \sum_{m=1}^{n} r_m e^{i\theta m}.
\end{aligned}$$

Now, observe that $s(t)$ traces out a spiral in the complex plane as t increases and let $\mathcal{D}$ be the closed and finite region bounded by the curve $s(t)$ for $t \in [0, \bar{m}]$ and the segment joining $s(\bar{m})$ with the origin. That is, $\mathcal{D}$ contains the first complete "turn" of the spiral (which corresponds to terms 1 to $\bar{m}$ of the series). Since r_m is decreasing, the region $\mathcal{D}$ will also encloses all subsequent "turns" of the spiral and we can write, for any $n \in \mathbb{N}$,

$$\begin{aligned}\left|\sum_{m=1}^{n} r_m e^{i\theta m}\right| &\leq \max_{z\in\mathcal{D}} \|z\| = \sup_{t\in[0,\bar{m}]} |s(t)| \leq \sup_{t\in[0,\bar{m}]} \left|\frac{i\theta}{1-e^{-i\theta}}\right| \int_0^t \left|r_{\lceil\tau\rceil}\right| \left|e^{i\theta\tau}\right| d\tau \\ &= \left(\sup_{\theta\in[-\pi,\pi]} \left|\frac{i\theta}{1-e^{-i\theta}}\right|\right) \sup_{t\in[0,\bar{m}]} \int_0^t r_{\lceil t\rceil} dt \leq \frac{\pi}{2} \int_0^{\bar{m}} r_{\lceil t\rceil} dt = \frac{\pi}{2} \sum_{m=1}^{\bar{m}} r_m.\end{aligned}$$

■

Proof. The proof is similar to the one of Theorem 4 and we focus here on the differences. It is clear that the differenced process $\Delta Z_n(t)$ admits the moving average representation:

$$\Delta Z_n(t) = \int_{-\infty}^{t} (z_n(t-s) - z_n(t-\Delta t - s))\, dW(s)$$

where the kernel $z_n(t-s) - z_n(t-\Delta t - s)$ is absolutely integrable/summable since it is a difference of two absolutely integrable/summable terms. Its Fourier transform is thus well-defined and equal to:

$$\Delta\tilde{z}_n(\lambda) = \int_0^\infty (z_n(t) - z_n(t-\Delta t))\, e^{i\lambda t} dt = \tilde{z}_n(\lambda) - e^{i\lambda\Delta t}\tilde{z}_n(\lambda) = \left(1 - e^{i\lambda\Delta t}\right)\tilde{z}_n(\lambda)$$

The pointwise limit of $\Delta\tilde{z}_n(\lambda)$ also poses no problem (as in Theorem 3):

$$\Delta\tilde{z}_\infty(\lambda) \equiv \lim_{n\to\infty} \left(1 - e^{i\lambda\Delta t}\right)\tilde{z}_n(\lambda) = \left(1 - e^{i\lambda\Delta t}\right)\tilde{z}_\infty(\lambda),$$

with the aditionaly advantage that $\Delta\tilde{z}_n(0) = 0$ and therefore $\Delta\tilde{z}_\infty(0) = 0$ (so the $\lambda = 0$ point is no longer exceptional).

Now observe that, for some sufficiently small $\bar{\lambda} > 0$,

$$\begin{aligned}\int_{\mathbb{L}} \left|\Delta\tilde{z}_\infty(\lambda)\right|^2 d\lambda &= \int_{|\lambda|\leq\bar{\lambda}} \left|\left(1 - e^{i\lambda\Delta t}\right)\tilde{z}_\infty(\lambda)\right|^2 d\lambda + \int_{\lambda\in\mathbb{L}:\left|\lambda\geq\bar{\lambda}\right|} \left|\left(1 - e^{i\lambda\Delta t}\right)\tilde{z}_\infty(\lambda)\right|^2 d\lambda \\ &\leq \int_{|\lambda|\leq\bar{\lambda}} C_1 \left|\Delta t\lambda\right|^2 |\lambda|^{-2\alpha} d\lambda + \int_{\lambda\in\mathbb{L}:\left|\lambda\geq\bar{\lambda}\right|} 2\left|\tilde{z}_\infty(\lambda)\right|^2 d\lambda \\ &\leq \int_{|\lambda|\leq\bar{\lambda}} C_1 |\lambda|^{2(1-\alpha)} d\lambda + \int_{\lambda\in\mathbb{L}:\left|\lambda\geq\bar{\lambda}\right|} 2\left|\tilde{y}(\lambda)\right|^2 d\lambda < \infty\end{aligned}$$

for some finite constant $C_1 > 0$ and where $1-\alpha \geq 0$. Hence $\Delta\tilde{z}_\infty \in \mathcal{L}_2(\mathbb{R})$ and therefore the corresponding Δz_∞ is also in $\mathcal{L}_2(\mathbb{R}^+)$ and the corresponding process $\Delta Z_\infty(t)$ is stationary.

Next, we again make use of Lebesgue's dominated convergence theorem to show that $\int_{\mathbb{L}} |\Delta\tilde{z}_n(\lambda) - \Delta\tilde{z}_\infty(\lambda)|^2 d\lambda \to 0$, which requires the existence of a square integrable $\bar{z}(\lambda)$ such that $|\Delta\tilde{z}_n(\lambda) - \Delta\tilde{z}_\infty(\lambda)| \leq \bar{z}(\lambda)$. For $|\lambda| \geq \bar{\lambda}$, we proceed as in Theorem 3 after noting that the prefactor $\left(1 - e^{i\lambda\Delta t}\right)$ is bounded in magnitude by 2. For $|\lambda| \leq \bar{\lambda}$, we proceed as in Theorem 3, after noting that the prefactor $\left(1 - e^{i\lambda\Delta t}\right)$ is bounded in magnitude by $C_2|\lambda|$ for some finite $C_2 > 0$. This leads to a $\bar{z}(\lambda)$ that has the form $|\lambda|^{1-\alpha}$ (instead of $|\lambda|^{-\alpha}$), which is clearly square integrable for $|\lambda| \leq \bar{\lambda}$ for any $\alpha \in [0,1]$. ■

Proof of Theorem 5. Let F denote the distribution of U_n (the same for any n). Note that the distribution of X_n (denoted $F^{\otimes n}$, the n-fold convolution of F with itself) is supported on $\mathbb{Z}^d$, so that $P[X_n = x_0]$ can be written in the form

$$P[X_n = x_0] = \int_{\mathbb{R}^d} g(x - x_0)\, dF^{\otimes n}(x) \tag{9}$$

where $g : \mathbb{R}^d \mapsto \mathbb{R}$ is a continuous function such that $g(0) = 1$ and $g(x) = 0$ for $x \in \mathbb{Z}^d \setminus \{0\}$ (its value for $x \in \mathbb{R}^d \setminus \mathbb{Z}^d$ is not restricted, other than to satisfy continuity). A convenient choice of $g(x)$ is

$$g(x) = \prod_{j=1}^{d} \frac{\sin(\pi x_j)}{\pi x_j}.$$

Note that $g(x)$ is continuous (even at $x = 0$), $\sin(\pi x_j) = 0$ for any integer x_j and $g(0) = 1$ (as defined via a limit). The function $g(x)$ is the inverse Fourier transform of a rectangular function on $[-\pi, \pi]^d$:

$$g(x) = (2\pi)^{-d} \int_{\xi \in [-\pi,\pi]^d} e^{-i\xi x} d\xi.$$

Using Parseval's identity, we can write (9) in terms of Fourier transforms:

$$P[X_n = x_0] = (2\pi)^{-d} \int_{\xi \in [-\pi,\pi]^d} e^{i\xi \cdot x_0} \left(\tilde{f}(\xi)\right)^n d\xi,$$

where $\tilde{f}(\xi)$ is the characteristic function of the probability measure F and, by the Convolution Theorem, $\left(\tilde{f}(\xi)\right)^n$ is the characteristic function of the probability measure $F^{\otimes n}$.

We can further decompose $P[X_n = x_0]$ as

$$P[X_n = x_0] = (2\pi)^{-d} \int_{\xi \in \mathcal{B}\left(n^{-1/2+\varepsilon}\right)} e^{i\xi \cdot x_0} \left(\tilde{f}(\xi)\right)^n d\xi + R_1 \tag{10}$$

where $\mathcal{B}(r)$ denotes an open ball of radius r centered at the origin, $\varepsilon \in]0, 1/8[$ and where R_1 is a remainder:

$$R_1 = (2\pi)^{-d} \int_{\xi \in [-\pi,\pi]^d \setminus \mathcal{B}\left(n^{-1/2+\varepsilon}\right)} e^{i\xi \cdot x_0} \left(\tilde{f}(\xi)\right)^n d\xi. \tag{11}$$

To bound R_1, we observe that, since U_n is supported on a finite subset of $\mathbb{Z}^d$, the characteristic function $\tilde{f}(\xi)$ is a sum of a finite number of terms of the form $e^{i\xi\cdot x}$, with $x \in \mathbb{Z}^d$. As a result, $\left|\tilde{f}(\xi)\right|$ can only reach the value 1 if $\xi/(2\pi) \in \mathbb{Z}^d$. Hence, in the set $[-\pi,\pi]^d$, $\left|\tilde{f}(\xi)\right|$ can only reach 1 at $\xi = 0$. Since U_n is supported on a bounded set, any of its moments are finite and thus $\tilde{f}(\xi)$ is differentiable (any number of times) and, in particular, it admits a Taylor expansion about $\xi = 0$:

$$\tilde{f}(\xi) = 1 + \frac{1}{2}\xi' \tilde{f}^{(2)}(0)\,\xi + O\left(\|\xi\|^4\right) \tag{12}$$

where we exploit the facts that $\tilde{f}(0) = 1$ and that the distribution of U_n is symmetric about 0, so all odd terms vanish. Also the second derivative $\tilde{f}^{(2)}(0)$ is a negative-definite $d\times d$ matrix by the moment theorem, since $\operatorname{Var}[U_n]$ is positive-definite by assumption. The expansion (12) implies that there exists $\eta_1 > 0$ such that $\left|\tilde{f}(\xi)\right| \leq 1 - C_1\|\xi\|^2$ for any $\xi \in \mathcal{B}(\eta_1)$ for some $C_1 > 0$. Let $\xi_0 \equiv \arg\max_{\xi\in[-\pi,\pi]^d\backslash\mathcal{B}(\eta_1)}\left|\tilde{f}(\xi)\right|$, which exists since $\tilde{f}(\xi)$ is continuous and $[-\pi,\pi]^d\backslash\mathcal{B}(\eta)$ is compact. Since $\left|\tilde{f}(\xi)\right|$ only reaches 1 at $\xi = 0$, we must have $\left|\tilde{f}(\xi_0)\right| < 1$. Let $f_1 = \left(1 + \left|\tilde{f}(\xi_0)\right|\right)/2$ and pick $\eta_2 \in\,]0,\eta_1]$ such that for any $\xi \in \mathcal{B}(\eta_2)$ we have $\left|\tilde{f}(\xi)\right| > f_1$. Such an η_2 always exists since $\left|\tilde{f}(\xi)\right| \leq 1 - C_1\|\xi\|^2$ for $\xi \in \mathcal{B}(\eta_1)$. It follows that for any n such that $n^{-1/2+\varepsilon} < \eta_2$, we have $\left|\tilde{f}(\xi)\right| \leq 1 - C_1\left(n^{-1/2+\varepsilon}\right)^2 = 1 - C_1 n^{-1+2\varepsilon}$ for any $\xi \in [-\pi,\pi]^d\backslash\mathcal{B}\left(n^{-1/2+\varepsilon}\right)$. We can now bound the $\left(\tilde{f}(\xi)\right)^n$ term in (11) as:

$$\begin{aligned}
\sup_{\xi\in[-\pi,\pi]^d\backslash\mathcal{B}\left(n^{-1/2+\varepsilon}\right)}\left|\left(\tilde{f}(\xi)\right)^n\right| &= \sup_{\xi\in[-\pi,\pi]^d\backslash\mathcal{B}\left(n^{-1/2+\varepsilon}\right)}\left|\exp\left(n\ln\tilde{f}(\xi)\right)\right| \\
&\leq \exp\left(n\ln\left(1 - C_1 n^{-1+2\varepsilon}\right)\right) = \exp\left(n\left(-C_1 n^{-1+2\varepsilon} + O\left(n^{-2+4\varepsilon}\right)\right)\right) \\
&= \exp\left(-C_1 n^{2\varepsilon} + O\left(n^{-1+4\varepsilon}\right)\right) \leq \exp\left(-C_2 n^{2\varepsilon}\right)
\end{aligned}$$

for some $C_2 \in\,]0, C_1[$ for all n sufficiently large. We then have

$$\begin{aligned}
|R_1| &\leq (2\pi)^{-d}\int_{\xi\in[-\pi,\pi]^d\backslash\mathcal{B}\left(n^{-1/2+\varepsilon}\right)^d}\left|e^{i\xi\cdot x_0}\right|\exp\left(-C_2 n^{2\varepsilon}\right)d\xi \\
&\leq (2\pi)^{-d}\exp\left(-C_2 n^{2\varepsilon}\right)\int_{\xi\in[-\pi,\pi]^d} d\xi = \exp\left(-C_1 n^{2\varepsilon}\right),
\end{aligned}$$

which goes to 0 faster than any negative power of n.

We now come back to $P[X_n = x_0]$ given by Equation (10), in which we now write $\left(\tilde{f}(\xi)\right)^n$ as $\exp\left(n\tilde{F}(\xi)\right)$ with $\tilde{F}(\xi) = \ln\tilde{f}(\xi)$. Note that since $\tilde{f}(\xi)$ is differentiable (any number of times) and since $\tilde{f}(\xi)$ is nonvanishing in a neighborhood of $\xi = 0$ (because we established above that $\tilde{f}(\xi) = 1 + O\left(\xi^2\right)$), $\tilde{F}(\xi)$ admits a Taylor expansion about $\xi = 0$:

$$\begin{aligned}
\tilde{F}(\xi) &= \tilde{F}(0) + \tilde{F}^{(1)}(0)\,\xi + \frac{1}{2}\tilde{F}^{(2)}(0)\,\xi^2 + \frac{1}{6}\tilde{F}^{(3)}(0)\,\xi^3 + \frac{1}{24}\tilde{F}^{(4)}\left(\bar{\xi}\right)\xi^4 \\
&= \tilde{F}^{(2)}(0)\,\xi^2 + \tilde{F}^{(4)}\left(\bar{\xi}\right)\xi^4
\end{aligned}$$

where $\bar{\xi} \in [0, \xi]$ is a mean value and, for simplicity, we let an expression such as $\tilde{F}^{(k)}\left(\bar{\xi}\right)\xi^k$ stand for $\sum_{j_1,\ldots,j_k} \tilde{F}^{(k)}_{j_1,\ldots,j_k}\left(\bar{\xi}\right)\xi_{j_1}\cdots\xi_{j_k}$. We used symmetry of the distribution of U_n to obtain the second expression. Note that $\tilde{F}^{(2)}(0)$ is negative-definite by the moment theorem and the positivity of the variance of U_n. We then have

$$P\left[X_n = x_0\right] = (2\pi)^{-d} \int_{\xi \in \mathcal{B}\left(n^{-1/2+\varepsilon}\right)} e^{i\xi\cdot x_0} \exp\left(n\tilde{F}^{(2)}(0)\,\xi^2 + n\tilde{F}^{(4)}\left(\bar{\xi}\right)\xi^4\right) d\xi + R_1$$

Next, we make the change of variable $\xi = n^{-1/2}\tilde{\xi}$

$$\begin{aligned}
P\left[X_n = x_0\right] &= (2\pi)^{-d} \int_{\tilde{\xi} \in \mathcal{B}(n^{\varepsilon})} e^{in^{-1/2}\tilde{\xi}\cdot x_0} \exp\left(n\tilde{F}^{(2)}(0)\,n^{-1}\tilde{\xi}^2 + n\tilde{F}^{(4)}\left(n^{-1/2}\bar{\xi}\right)n^{-2}\tilde{\xi}^4\right) n^{-d/2} d\tilde{\xi} + R_1 \\
&= (2\pi)^{-d}\, n^{-d/2} R_0 + R_1
\end{aligned}$$

where

$$R_0 = \int_{\tilde{\xi} \in \mathcal{B}(n^{\varepsilon})} e^{in^{-1/2}\tilde{\xi}\cdot x_0} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2 + n\tilde{F}^{(4)}\left(n^{-1/2}\bar{\xi}\right)n^{-2}\tilde{\xi}^4\right) d\tilde{\xi}$$

in which the mean value $\bar{\xi}$ lies in $\left[0, \tilde{\xi}\right]$. We then have

$$\begin{aligned}
R_0 &= \int_{\tilde{\xi} \in \mathcal{B}(n^{\varepsilon})} e^{in^{-1/2}\tilde{\xi}\cdot x_0} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right) d\tilde{\xi} + R_2 \\
&= \int_{\tilde{\xi} \in \mathcal{B}(n^{\varepsilon})} \left(1 + in^{-1/2}\tilde{\xi}\cdot x_0 - \frac{n^{-1}}{2} e^{i\tilde{\xi}\cdot x_0}\left(\tilde{\xi}\cdot x_0\right)^2\right) \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right) d\tilde{\xi} + R_2 \\
&= \int_{\tilde{\xi} \in \mathcal{B}(n^{\varepsilon})} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right) d\tilde{\xi} + R_2 + R_3 + R_4 \\
&= C_0 + R_2 + R_3 + R_4 + R_5
\end{aligned}$$

where we have introduced the remainder terms:

$$\begin{aligned}
R_2 &= \int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} e^{in^{-1/2}\tilde{\xi}\cdot x_0} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)\left(\exp\left(\tilde{F}^{(4)}\left(n^{-1/2}\bar{\xi}\right)n^{-1}\tilde{\xi}^4\right)-1\right)d\tilde{\xi} \\
R_3 &= -\frac{n^{-1}}{2}\int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} e^{i\tilde{\xi}\cdot x_0}\left(\tilde{\xi}\cdot x_0\right)^2 \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi} \\
R_4 &= in^{-1/2}\int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} \tilde{\xi}\cdot x_0 \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi} \\
R_5 &= \int_{\tilde{\xi}\in\mathbb{R}^d\backslash\mathcal{B}(n^{\varepsilon})} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi}
\end{aligned}$$

and the constant $C_0 = \int_{\tilde{\xi}\in\mathbb{R}^d} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi} > 0$. Considering each term in turn, we have

$$\begin{aligned}
|R_2| &\leq \int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} \left|e^{in^{-1/2}\tilde{\xi}\cdot x_0}\right| \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)\left(\exp\left(\tilde{F}^{(4)}\left(n^{-1/2}\bar{\xi}\right)n^{-1}\tilde{\xi}^4\right)-1\right)d\tilde{\xi} \\
&= \int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)\left(\exp\left(\tilde{F}^{(4)}\left(n^{-1/2}\bar{\xi}\right)n^{-1}\tilde{\xi}^4\right)-1\right)d\tilde{\xi}.
\end{aligned}$$

Let $\bar{F}^{(4)} \equiv \sup_{\xi\in\mathcal{B}(\eta_3)}\left|\tilde{F}^{(4)}(\xi)\right|$ for some $\eta_3 > 0$. For n sufficiently large, we eventually have $n^{-1/2}\bar{\xi} \leq \eta_3$ and we can write

$$\begin{aligned}
|R_2| &\leq \int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)\left(\exp\left(\bar{F}^{(4)}n^{-1}n^{4\varepsilon}\right)-1\right)d\tilde{\xi} \\
&= \left(\exp\left(\bar{F}^{(4)}n^{-1+4\varepsilon}\right)-1\right)\int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi} \\
&= \left(1+\bar{F}^{(4)}n^{-1+4\varepsilon}+o\left(n^{-1+4\varepsilon}\right)-1\right)\int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi} \\
&= O\left(n^{-1+4\varepsilon}\right)\int_{\mathbb{R}^d} \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi}
\end{aligned}$$

where the last integral is finite since $\tilde{F}^{(2)}(0)$ is negative-definite. Next,

$$\begin{aligned}
|R_3| &\leq \frac{n^{-1}}{2}\int_{\tilde{\xi}\in\mathcal{B}(n^{\varepsilon})} \left|e^{i\tilde{\xi}\cdot x_0}\right|\left(\tilde{\xi}\cdot x_0\right)^2 \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi} \\
&\leq \frac{n^{-1}}{2}\int_{\tilde{\xi}\in\mathbb{R}^d} \left(\tilde{\xi}\cdot x_0\right)^2 \exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)d\tilde{\xi} \\
&= O\left(n^{-1}\right)
\end{aligned}$$

where the last integral is finite since $\tilde{F}^{(2)}(0)$ is negative-definite.

Next, R_4 vanishes by the symmetry of $\exp\left(\tilde{F}^{(2)}(0)\,\tilde{\xi}^2\right)$ (in $\tilde{\xi}$). Finally,

$$\begin{aligned} |R_5| &\leq \int_{\tilde{\xi}\in\mathbb{R}^d\setminus\mathcal{B}(n^{\varepsilon})} \exp\left(\lambda\left\|\tilde{\xi}\right\|^2\right) d\tilde{\xi} \\ &= S_d \int_{n^{\varepsilon}}^{\infty} \rho^d \exp\left(-\lambda\rho^2\right) d\rho \\ &\leq S_d \int_{n^{\varepsilon}}^{\infty} \exp\left(-C_2\rho\right) d\rho = \frac{S_d}{C_2}\exp\left(-C_2 n^{\varepsilon}\right) \end{aligned}$$

where λ is the smallest eigenvalue of $-\tilde{F}^{(2)}(0)$. In the second line, we have expressed the integral in polar coordinates with ρ being the radius and S_d is the $(d-1)$-dimensional "surface" of a hypersphere of radius 1. The third line holds for some $C_2 > 0$ for n sufficiently large and yields an expression that decays faster than any power of n.

Collecting the order of the remainders, we have, with $C = (2\pi)^{-d} C_0 > 0$,

$$\begin{aligned} P[X_n = x_0] &= (2\pi)^{-d} n^{-d/2}\left(C_0 + O\left(n^{-1+4\varepsilon}\right) + O\left(n^{-1}\right) + 0 + O\left(\exp\left(-C_2 n^{2\varepsilon}\right)\right)\right) + \\ &\quad +O\left(\exp\left(-C_1 n^{\varepsilon}\right)\right) \\ &= Cn^{-d/2} + O\left(n^{-1-d/2}\right) \end{aligned}$$

■

Proof of Theorem 6. Assumptions (i) and (ii) imply that X_n is a discrete-time finite-state time-homogenous Markov chain characterized by the transition matrix $T_{ij} \equiv P[X_{n+1} = j | X_n = i]$. Since Assumption (iii) implies that T is irreducible, the Perron-Frobenius Theorem (Meyer (2000), Chap. 8) lets us conclude that (i) all eigenvalues λ_h of T satisfy $|\lambda_h| \leq 1$, (ii) one eigenvalue of T satisfies $\lambda_0 = 1$, its associated 1×1 "Jordan block" is $J_0 = 1$ and has an associated eigenvector v_0 with strictly positive entries (and none of the other eigenvectors have all nonnegative entries), (iii) the other eigenvalues λ_h of T satisfying $|\lambda_h| = 1$ (if any), have the form $\lambda_h = e^{i2\pi h/H}$ for $h = 1,\ldots,H-1$ for some $0 \leq H < \#\mathbb{F}$, each have multiplicity 1, and each have an associated 1×1 Jordan block $J_h = \lambda_h$ and an eigenvector v_h and (iv) all other eigenvalues λ_h, $h \geq H$, of T are such that $|\lambda_h| < 1$ and have Jordan blocks that may, in general, be larger than 1×1.

Consider an arbitrary probability vector p, which can be expressed as the linear combination

$$p = \sum_{h=0}^{\#\mathbb{F}-1} a_h v_h$$

where v_h denote the columns of the matrix V yielding the Jordan form $J = V^{-1}TV$, with entries of $J = \operatorname{diag}(J_0, J_1, \ldots)$ ordered as above. Consider the effect of iterated applications of T to p:

$$\begin{aligned} T^n p &= a_0 T^n v_0 + \sum_{h=1}^{H-1} a_h T^n v_h + \sum_{h'=H}^{\#\mathbb{F}-1} T^n a_{h'} v_{h'} \\ &= a_0 v_0 + \sum_{h=1}^{H-1} e^{i2\pi hn/H} a_h v_h + \sum_{h'=H}^{\#\mathbb{F}-1} a_{h'} T^n v_{h'}, \end{aligned}$$

where, by convention the sum $\sum_{h=1}^{H-1}$ vanishes if $H = 0$. By the properties of the Jordan form, $T^n v_{h'} = A v_{h'}$, where A is a block-diagonal matrix with entries of the form $\binom{n}{m} \lambda_{h'}^{n-m}$ with k ranging from 0 to $m-1$, where m is the size of the largest Jordan block and where $\lambda_{h'}$ is any one of the eigenvalues such that $|\lambda_{h'}| < 1$. Since m is bounded and $\binom{n}{m} \leq n^m/m!$ only grows as a power of n, we have that, for any $\varepsilon_1 > 0$, there exists $C_1 > 0$ such that $\binom{n}{m} \leq C_1 (1+\varepsilon_1)^n$. Hence, there exist $C_2 > 0$ such that $\binom{n}{m} \lambda_{h'}^{n-m} \leq C_2 \bar{\lambda}^n$, for some $\bar{\lambda} \in \left]\max_{h' \geq H} |\lambda_{h'}|, 1\right[$ and it follows that $\left|\sum_{h'=H}^{\#\mathbb{F}-1} a_{h'} T^n v_{h'}\right| \leq C_3 \bar{\lambda}^n$ for some $C_3 > 0$. Note that, since $T^n p$ must have nonnegative entries, we must have that $a_0 > 0$ (otherwise, the $\sum_{h=1}^{H-1} e^{i2\pi hn/H} a_h v_h$ term would eventually yield some negative entries, since the $\sum_{h'=H}^{\#\mathbb{F}-1} a_{h'} T^n v_{h'} \to 0$ as $n \to \infty$).

Now, $P[X_n = x_1 | X_0 = x_0]$ is given by $[T^n p]_{x_1}$ for p with elements $p_i = \mathbf{1}[i = x_0]$ and it follows that, for some $\varepsilon > 0$,

$$P[X_n = x_1 | X_0 = x_0] = C_0 + \sum_{h=1}^{H-1} C_h e^{i2\pi nh/H} + O\left((1-\varepsilon)^n\right)$$

where $C_0 = a_0 [v_0]_{x_1} > 0$ and $C_h = a_h [v_h]_{x_1}$. Requiring $P[X_n = x_1 | X_0 = x_0]$ to be real, forces $C_h = C_{H-h}^*$. ■

Proof of Theorem 7. Each of the term of the form $C_h e^{i2\pi nh/H}$ gives a spectrum of the form:

$$\sum_{n=0}^{\infty} C_h e^{i2\pi nh/H} (\tilde{r}(\lambda))^n = C_h \sum_{n=0}^{\infty} \left(e^{i2\pi h/H} \tilde{r}(\lambda)\right)^n = \frac{C_h}{1 - e^{i2\pi h/H} \tilde{r}(\lambda)}$$

for any λ such that $|\tilde{r}(\lambda)| < 1$. Note that, since $h/H \neq 1$, the denominator never vanishes and the function $C_h / \left(1 - e^{i2\pi h/H} x\right)$ is continuous in x and bounded for $|x| \leq 1$. By Assumption 1(iv), $|\tilde{r}(\lambda)| \leq 1$ in a neighborhood of the origin and therefore

$$\lim_{\lambda \to 0} \frac{C_h}{1 - e^{i2\pi h/H} \tilde{r}(\lambda)} = \frac{C_h}{1 - e^{i2\pi h/H} \lim_{\lambda \to 0} \tilde{r}(\lambda)} = \frac{C_h}{1 - e^{i2\pi h/H}} < \infty.$$

where we have used continuity and the fact that $\lim_{\lambda \to 0} \tilde{r}(\lambda) = 1$ by Assumption 1. It follows that $\lim_{\lambda \to 0} \tilde{z}_\infty(\lambda)$ exists and is bounded and so is $\lim_{\lambda \to 0} |\tilde{z}_\infty(\lambda)|^2$. Note that this proof does not establish that $\tilde{z}_\infty(0)$ is finite (it may not be). However, it is the limit $\lim_{\lambda\to 0}\tilde{z}_\infty(\lambda)$ and not $\tilde{z}_\infty(0)$ that determine whether oscillatory terms can affect the rate of divergence of the spectrum as $\lambda \to 0$. ■

Corollary 9 *Let c_n be a deterministic sequence and let the corresponding $\tilde{z}_\infty(\lambda)$ satisfy $\tilde{z}_\infty(\lambda) = A(i\lambda)^{-\alpha} + o\left((i\lambda)^{-\alpha}\right)$ (for $A \in \mathbb{R}$ and $\alpha \in \mathbb{R}^+$). Let c_n' be a random sequence such that $E\left[(c_n' - c_n)^2\right] \le D(1+n)^{-3-\varepsilon}$ for some $\varepsilon, D > 0$, then the corresponding $\tilde{z}_\infty'(\lambda)$ satisfies $\tilde{z}_\infty'(\lambda) = A(i\lambda)^{-\alpha} + o\left((i\lambda)^{-\alpha}\right)$ with probability one.*

Proof of Corollary 9. To simplify the notation let the sequence start at index $n = 1$ instead of 0. By Theorem 2, it suffice to show that $\sum_{n=1}^\infty |c_n' - c_n|$ is finite with probability one, i.e. $P\left[\sum_{n=1}^\infty |c_n' - c_n| \ge C\right] \to 0$ as $C \to \infty$. Let $\Delta c_n = c_n' - c_n$ and for a given C, let $c = C\left(\sum_{n=1}^\infty n^{-1-\varepsilon/3}\right)^{-1}$. Note that $\sum_{n=1}^\infty n^{-1-\varepsilon/3} < \infty$ and that $C \to \infty \implies c \to \infty$. Then note that $|\Delta c_n| \le cn^{-1-\varepsilon/3}$ for all $n \in \mathbb{N}^*$ implies that $\sum_{n=1}^\infty |\Delta c_n| \le C$. Taking the contrapositive of that statement yields that the event $\sum_{n=1}^\infty |\Delta c_n| \ge C$ implies the event $|\Delta c_n| \ge cn^{-1-\varepsilon/3}$ for some $n \in \mathbb{N}^*$. Then write

$$\begin{aligned}
P\left[\sum_{n=1}^\infty |\Delta c_n| \ge C\right] &\le P\left[|\Delta c_n| \ge cn^{-1-\varepsilon/3} \text{ for some } n \in \mathbb{N}^*\right] \\
&\le \sum_{n=1}^\infty P\left[|\Delta c_n| \ge cn^{-1-\varepsilon/3}\right] \\
&= \sum_{n=1}^\infty P\left[|\Delta c_n|^2 \ge c^2 n^{-2-2\varepsilon/3}\right] \\
&\le \sum_{n=1}^\infty \frac{E\left[|\Delta c_n|^2\right]}{c^2 n^{-2-(2/3)\varepsilon}} \\
&\le \sum_{n=1}^\infty \frac{Dn^{-3-\varepsilon}}{c^2 n^{-2-(2/3)\varepsilon}} = \frac{D}{c^2}\sum_{n=1}^\infty n^{-1-\varepsilon/3}
\end{aligned}$$

where we have used, in turn, (i) the fact that if two events are such that $A \implies B$ then $P[B] \ge P[A]$, (ii) for any sequence of events A_i, we have $P[\cup_i A_i] \le \sum_i P[A_i]$, (iii) monotonicity of the function x^2 for $x \ge 0$ (iv) Markov's inequality $P[X \ge x] \le E[X]/x$ applied to the random variable $X = |\Delta c_n|^2$, (v) the assumption $E\left[|\Delta c_n|^2\right] \le Dn^{-3-\varepsilon}$. Since $\sum_{n=1}^\infty n^{-1-\varepsilon/3} < \infty$, it follows that, as $C \to \infty$, $c \to \infty$ and $P\left[\sum_{n=1}^\infty |X_n| \ge C\right] \to 0$, as desired. ■

Proof of Theorem 8. The spectral representation of Equation (4) is

$$\begin{aligned}
\tilde{z}_{\bar{n}}(\lambda) &= \tilde{y}_0(\lambda)\sum_{n=0}^{\bar{n}} c_n \frac{1}{|\mathcal{P}_{n,ij}|}\sum_{p\in\mathcal{P}_{n,ij}}\prod_{l=1}^{n}\tilde{r}_{p_l}(\lambda) \\
&= \tilde{y}_0(\lambda)\sum_{n=0}^{\bar{n}} c_n (\bar{r}(\lambda))^n \frac{1}{|\mathcal{P}_{n,ij}|}\sum_{p\in\mathcal{P}_{n,ij}}\prod_{l=1}^{n}\frac{\tilde{r}_{p_l}(\lambda)}{\bar{r}(\lambda)} \\
&= \tilde{y}_0(\lambda)\sum_{n=0}^{\bar{n}}\left(c_n + \frac{c_n}{|\mathcal{P}_{n,ij}|}\sum_{p\in\mathcal{P}_{n,ij}}\left(\prod_{l=1}^{n}\frac{\tilde{r}_{p_l}(\lambda)}{\bar{r}(\lambda)} - 1\right)\right)(\bar{r}(\lambda))^n \\
&= \tilde{y}_0(\lambda)\sum_{n=0}^{\bar{n}}(c_n + \Delta c_n)(\bar{r}(\lambda))^n
\end{aligned}$$

where

$$\Delta c_n = \frac{c_n}{|\mathcal{P}_{n,ij}|}\sum_{p\in\mathcal{P}_{n,ij}}\left(\prod_{l=1}^{n}\frac{\tilde{r}_{p_l}(\lambda)}{\bar{r}(\lambda)} - 1\right)$$

Hence, Corollary 9 applies directly. ■

References

Abadir, K., and G. Talmain (2002): "Aggregation, persistence and volatility in a macro model," *Review of Economic Studies*, 69, 749–779.

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, forthcoming.

Alfarano, S., and T. Lux (2005): "A Noise Trader Model as a Generator of Apparent Financial Power Laws and Long Memory," Working Paper, University of Kiel.

Baillie, R. T. (1996): "Long memory processes and fractional integration in econometrics," *Journal of Econometrics*, 73, 5–59.

Bandt, C., and S. Graf (1992): "Self-Similar Sets 7. A Characterization of Self-Similar Fractals with Positive Hausdorff Measure," *Proceedings of the American Mathematical Society*, 114, 995–1001.

Calvet, L., and A. Fisher (2002): "Multifractality in asset returns: Theory and evidence," *Review Of Economics and Statistics*, 84, 381–406.

Chambers, M. J. (1998): "Long memory and aggregation in macroeconomic time series," *International Economic Review*, 39, 1053–1072.

Chen, X., L. P. Hansen, and M. Carrasco (2010): "Nonlinearity and Temporal Dependence," *Journal of Econometrics*, 155, 155–169.

Chevillon, G., and S. Mavroeidis (2011): "Learning generates Long Memory," Working Paper, University of Oxford.

Comte, F., and E. Renault (1996): "Long memory continuous time models," *Journal of Econometrics*, 73, 101–149.

——— (1998): "Long memory in continuous-time stochastic volatility models," *Mathematical Finance*, 8, 291–323.

Davidson, J., and P. Sibbertsen (2005): "Generating schemes for long memory processes: regimes, aggregation and linearity," *Journal of Econometrics*, 128, 253–282.

Diebold, F. X., and A. Inoue (2001): "Long memory and regime switching," *Journal of Econometrics*, 105, 131–159.

Doob, J. L. (1953): *Stochastic Processes*. Wiley, New York.

Dupor, B. (1999): "Aggregation and irrelevance in multi-sector models," *Journal of Monetary Economics*, 43, 391–409.

Flandrin, P. (1989): "On the Spectrum of Fractional Brownian Motions," *IEEE Transactions on information theory*, 35, 197–199.

Granger, C. W. J. (1980): "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227–238.

Granger, C. W. J., and Z. Ding (1996): "Varieties of long memory models," *Journal of Econometrics*, 73, 61–77.

Hall, R. E. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971–987.

Havlin, S., and D. Ben-Avraham (1987): "Diffusion in disordered media," *Advances in Physics*, 36, 695–798.

Horvath, M. (1998): "Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks," *Review of Economic Dynamics*, 1, 781–808.

Inaoka, H., T. Ninomiya, K. Taniguchi, and T. S. H. Takayasu (2004): "Fractal Network derived from banking transaction — An analysis of network structures formed by financial institutions," Working Paper 04-E-04, Bank of Japan.

Lobato, I., and P. M. Robinson (1996): "Averaged periodogram estimation of long memory," *Journal of Econometrics*, 73, 303–324.

Long, J. B., and C. I. Plosser (1983): "Real Business Cycles," *Journal of Political Economy*, 91, 39–69.

Loyne, R. M. (1968): "On the Concept of the Spectrum for Non-stationary Processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, 30, 1–30.

Mandelbrot, B. B., and J. W. V. Ness (1968): "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, 10, 422–437.

Meyer, C. D. (2000): *Matrix Analysis and Applied Linear Algebra Book and Solutions Manual*. SIAM, Philadelphia.

Miller, J. I., and J. Y. Park (2010): "Nonlinearity, nonstationarity, and thick tails: How they interact to generate persistence in memory," *Journal of Econometrics*, 155, 83–89.

Nelson, C., and C. I. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 10, 139–162.

Parke, W. R. (1999): "What is fractional integration?," *Review of Economics and Statistics*, 81, 632–638.

Perron, P. (1988): "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach," *Journal of Economic Dynamics and Control*, 12, 297–332.

——— (1989): "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57, 1361–1401.

Perron, P., and Z. Qu (2007): "An analytical evaluation of the log-periodogram estimate in the presence of level shifts," Working Paper, Boston University.

Phillips, P. C. B. (1987): "Time Series Regression with a Unit Root," *Econometrica*, 55, 277–301.

Song, C., S. Havlin, and H. A. Makse (2005): "Self-similarity of complex networks," *Nature*, 433, 392–395.

Zafaroni, P. (2004): "Contemporaneous aggregation of linear dynamic models in large economies," *Journal of Econometrics*, 120, 75–102.