Long-range correlation properties of stationary linear models with mixed periodicities

Tomomichi Nakamura,[1,*] Michael Small,[2,3,4,†] and Toshihiro Tanizawa[5,‡]

[1]*Graduate School of Simulation Studies, University of Hyogo, 7-1-28 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan*

[2]*School of Mathematics and Statistics, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

[3]*Complex Data Modelling Group, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

[4]*Mineral Resources, CSIRO, Kensington, WA 6151, Australia*

[5]*Kochi National College of Technology, Monobe-Otsu 200-1, Nankoku, Kochi 783-8508, Japan*

(Received 23 April 2018; revised manuscript received 14 January 2019; published 19 February 2019)

We consider the problem of (stationary and linear) source systems which generate time series data with long-range correlations. We use the discrete Fourier transform (DFT) and build stationary linear models using artificial time series data exhibiting a $1/f$ spectrum, where the models can include only terms that contribute significantly to the model as assessed by information criteria. The result is that the optimal (best) model is only composed of mixed periodicities [that is, the model does not include all (continuous) periodicities] and the time series data generated by the model exhibit a clear $1/f$ spectrum in a wide frequency range. It is considered that as the $1/f$ spectrum is a consequence of the contributions of all periods, consecutive periods are indispensable to generate such data by stationary linear models. However, the results indicate that there are cases where this expectation is not always met. These results also imply that although we can know linear features of time series data using the DFT, we always cannot substantially infer the type of the source system, even if the system is stationary linear.

DOI: 10.1103/PhysRevE.99.022128

I. INTRODUCTION

Time series of natural phenomena usually show irregular fluctuations. One of the major reasons for the appearance of such data is periodicities (periodic and nearly periodic behavior) in systems [1], and periodicities are a common feature of many biological and physical systems [2]. Generally speaking, as it is considered that periodicities composing the source system correspond to periodicities in the time series generated by the system [3], we often wish to know the underlying periodicities in the data. Since it is believed that the discrete Fourier transform (DFT) can accurately detect the underlying periodicities in the data, one usually applies the DFT in the hope of obtaining such information. We argue that in many cases there are a relatively small number of representative (characteristic) periods in the data (or the number of the periods is not so many). On the other hand, there is an important biological and physical phenomenon which is considered that there is no characteristic periodicity in the data (in other words, all periodicities contribute the data behavior). The phenomenon is $1/f$ noise, an attractive symptom of complexity in biological and physical systems [4–9]. It is widely considered that the difference between these two cases is obvious—we can easily distinguish between them by the DFT, and the DFT has been the standard method for a long time. Nonetheless, while pathological counterexamples are known to exist, we show that data generated by stationary linear models with only some mixed periodicities can exhibit clear $1/f$ spectra in a wide frequency range, where the power spectra are estimated by the DFT [in other words, stationary linear models with discontinuous (separate) periods can produce data with long-range correlations].

Periodicity is one of the important clues to understand dynamical phenomena. When data are generated by a stationary linear process, the information required for modeling and prediction is encapsulated in the periodicities of the system—either via the power spectrum, or, equivalently, the autocorrelation function. It is true that nonlinear (and deterministic) dynamical systems might also exhibit periodicities. In such systems the periodic structure is usually insufficient to uniquely characterize the underlying dynamical system. Nonetheless, the power spectrum still remains an important clue to understand the underlying characteristics both in the data and the original system. Obviously, the primary method to access information of the spectral content of a time series is the Fourier power spectrum (power spectra estimated by the DFT)—typically, the fast Fourier transform or FFT is usually applied to detect the periodicities in the data [10]. It is widely considered and expected that strong peaks in a power spectrum at some periods (frequencies) infer the underlying periods in the data, and the locations of the peaks are directly related to the principal periods in the system dynamics [3].

Conversely, there has been increasing interest in fascinating phenomena where the power spectra estimated by the DFT do not exhibit clear peaks at any frequency. The archetypal phenomenon of this type of spectra is the so-called $1/f$ (erstwhile *pink* or *flicker*) noise which has been widely known and well studied for many years [4–9]. In $1/f$ noise, the power spectrum $P(f)$ varies in inverse proportion to the frequency f in a wide range from low to high frequencies as $P(f) = 1/f^{\alpha}$, with a power-law exponent α close to 1. The lack of

*tomo@sim.u-hyogo.ac.jp

†michael.small@uwa.edu.au

‡tanizawa@ee.kochi-ct.ac.jp

peaks in the power spectrum as a function of frequency is interpreted as a lack of a characteristic timescale in $1/f$ noise and the $1/f$ spectrum is a typical feature which represents long-range correlations [11,12]. To investigate whether data can be treated as $1/f$ noise, one typically estimates the power spectrum and to do this we usually apply the DFT directly to the data.

The primary purpose of this paper is to consider a problem of source systems which generate time series data with long-range correlations, where the data and the systems are stationary and linear. A simple example using data with short-range correlations has been investigated where the data are generated by a simple stationary linear model composed of two periods [3]. To obtain further understanding of the relationship between periodicities in data and periodicities composing the source system, we consider that it is necessary to examine a case of long-range correlations, because we expect that the data with long-range correlations should involve numerous periodicities. It is considered that the $1/f$ spectrum is a consequence of the contributions of all periods over the $1/f$ spectrum, or that equivalently all timescales are contained in the data. Hence, we investigate whether stationary linear models with some mixed periodicities can generate time series data exhibiting a $1/f$ spectrum in a wide frequency range.

We emphasize that the periodicities which we investigate in this paper are all strictly linear and stationary. Tangential to our main argument, there are, of course, more complex nonlinear systems with no periodicities in the data even when they are generated from deterministic dynamics. It is clear that such systems, when treated with the methods of a linear time series analysis, will yield misleading results. The prime example of this type is the logistic map [13]. The equation of the logistic map is $x(t) = 4x(t-1)[1-x(t-1)] = 4x(t-1) - 4x(t-1)^2$. The equation contains the terms with an explicit time delay, $x(t-1)$ and $x(t-1)^2$. However, it is well known that the randomness is linearly equivalent to that of independent and identically distributed (IID) random variables and the power spectrum estimated by the DFT exhibits white noise (a $1/f^0$-type spectrum). By limiting periodicities to strictly linear and stationary cases, we exclude the possibility that the example we will discuss in this paper also belongs to the class of nonlinear and nonstationary complexity.

In Sec. II we describe the manifestation of the $1/f$ spectrum and that there are various behaviors of the $1/f$ spectrum. In Sec. III we introduce a linear model composed of consecutive time delays (periods) that can generate data exhibiting the $1/f$ spectrum. In Sec. IV we show that the data that are not composed of continuous periods still exhibit a clear $1/f$ spectrum in the whole frequency range. We summarize the results in Sec. V.

II. VARIOUS BEHAVIORS OF A $1/f$ SPECTRUM

There are various behaviors of the $1/f$ spectrum [14–20]. Some power spectra exhibit $1/f$ within a restricted range, while others exhibit $1/f$ in the whole range; there are spikes in the middle, and there are ranges as white noise on the low-frequency side and the high-frequency side. We do not target all these peculiarities in this paper. To examine the correspondence between periodicities in the time series and periodicities composing the source system more directly and to clarify our argument, we focus our attention on the power spectrum exhibiting only a $1/f$ spectrum in the whole range. When using a linear model the range where the $1/f$ spectrum appears is largely dependent on the maximum time delay term included in the model. To lengthen the range of the $1/f$ spectrum, it is not sufficient to merely increase the number of data—a longer time delay term is required. Hence, we use the data number by which only the $1/f$ spectrum appears, and treat situations where all frequencies corresponding to all time delays except $t-1$ in the linear models fall within the spectrum range.

There are approaches to generate data which can exhibit a $1/f$ spectrum [21–23]. Among them, there is a particularly simple model. The model is $x(t) = a\,x(t-1) + \eta(t)$, where $\eta(t)$ is IID Gaussian random variables with mean zero and a standard deviation. Broadly speaking, when a is smaller than one, it is shown in some cases that data generated by the model can exhibit a $1/f$ spectrum in the partial range [22]. However, we need time series data which can exhibit a $1/f$ spectrum in the whole range, and the model is composed of only one time period $t-1$ that cannot be detected by the DFT. Hence, the model and data generated by the model are not suitable for our purpose.

As mentioned above, we want to examine the correspondence between periodicities of the data behavior and periodicities composing the source system. Hence, we need linear models with some distinct periods to generate data. However, an arbitrary time delay is not acceptable for the linear models. It is necessary to use time delays reflecting the periodical characteristics of the time series. A reduced autoregressive (RAR) model has already been proposed as such a linear model [3,24,25]. The time delay terms included in a RAR model are evaluated to be indispensable by an information criterion (the information theory). Furthermore, it is indicated that the time delay terms in the RAR model correspond to periodicities in the time series data generated by the RAR model [3]. We will give more details on the RAR models in Sec. IV A.

The Fourier transform is linear and as such it possesses the properties of homogeneity and additivity. Hence, we use only linear models as the basis for our presentation in this paper, although the power spectra of a nonlinear time series using the DFT can exhibit a $1/f$ spectrum [26]—that is, none of the effects we present here is due to the mismatch between a nonlinear source signal and the inherently linear Fourier transform.

In the next section we introduce a linear model that can generate data exhibiting a $1/f$ spectrum, where the model is based on the typical $1/f$ noise process of the literature. In Sec. IV we build a RAR model for a $1/f$ noise time series.

III. GENERATING $1/f$ NOISE WITH CONTINUOUS PERIODICITIES

Broadly speaking, it is understood that $1/f$ noise has all periodicities. That is, $1/f$ noise has no single characteristic periodicity. Also, the longer the period (the lower the frequency), the larger is the power spectrum. An autoregressive

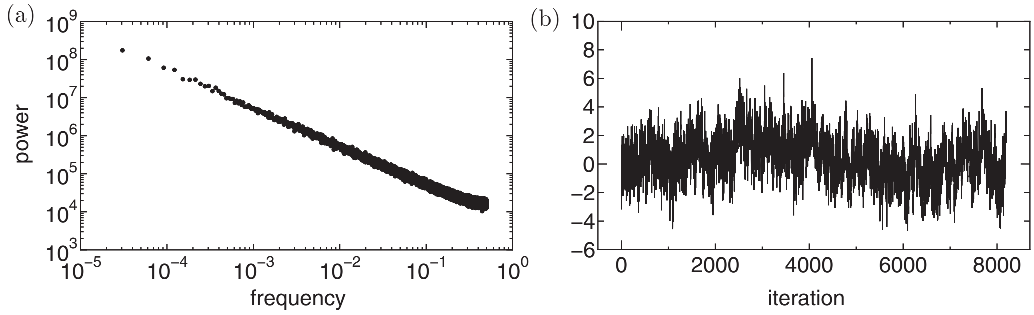

FIG. 1. Power spectrum with a double-logarithmic scale and a time series, where the data are generated by Eqs. (1) and (2), where the model size w is 2500, and $\alpha = 1.0$. (a) The power spectrum estimated by the FFT using 32 768 (2^{15}) data points at an assumed sampling rate 1 Hz, where the power spectrum is obtained by averaging 100 different data sets, and (b) a segment of 8192 data points used for the power spectrum estimate in (a).

(AR) model generating $1/f^{\alpha}$ noise based on this understanding has been proposed by Kasdin [6]. The AR model is linear and composed of consecutive terms with a unit time delay,

$$x(t) = \sum_{i=1}^{w} -a_i x(t-i) + \varepsilon(t), \tag{1}$$

$$a_k = \left(k - 1 - \frac{\alpha}{2}\right)\frac{a_{k-1}}{k}, \tag{2}$$

where w is the model size (the largest time delay), $a_0 = 1.0$, and $\varepsilon(t)$ is assumed to be IID Gaussian random variables with mean zero and the standard deviation 1.0. Each time delay in the model becomes each underlying period in the data. Also, if time delays in the model are consecutive, the periods in the data are continuous. Hence, the data generated by Eqs. (1) and (2) have no characteristic timescale or (equivalently) reflect features across all timescales. We consider that this is the typical $1/f$ noise process many people expect. Using Eqs. (1) and (2) with the model size $w = 2500$ and $\alpha = 1.0$, we generate the data and estimate the power spectrum. The number of data points in one sample is 32 768 (2^{15}) and the power spectrum is estimated by averaging over 100 samples [27]. The data generated by Eqs. (1) and (2) exhibit a clear $1/f$ spectrum in the whole consecutive frequency range as shown in Fig. 1(a). Figure 1(b) shows an enlargement of the behavior of a sample. The behavior shows irregular fluctuations and a wave form similar to a slow undulation, which are characteristic of $1/f$ noise [5,6]. We have thus verified that the model composed of Eqs. (1) and (2) actually generates $1/f$ noise. From this result, it seems that the contribution of continuous periods is indispensable to generate $1/f$ noise, which appears to be the major reason to use autoregressive processes to generate $1/f$ noise [28,29].

IV. GENERATING $1/f$ NOISE WITH SEPARATE PERIODICITIES

In the previous section we showed that data generated by a system composed of consecutive time delays (periods) exhibit a clear $1/f$ spectrum. In this section we investigate whether data generated by a system composed of inconsecutive (separate) time delays can exhibit a $1/f$ spectrum.

A. The reduced autoregressive model

To build a linear model composed of separate time delays (linear model with mixed periodicities), we adopt the reduced autoregressive (RAR) model [3]. RAR models include only terms that contribute significantly to the model as assessed by an information criterion [24,30]. The terms included in the RAR model correspond to periodicities in the time series generated by the RAR model [3]. The RAR model has proven to be effective in modeling both linear and nonlinear dynamics [3,25].

The form of the RAR models is

$$\begin{aligned} x(t) &= a_0 + a_1 x(t-l_1) + a_2 x(t-l_2) + \cdots + a_w x(t-l_w) + \varepsilon(t) \\ &= a_0 + \sum_{i=1}^{w} a_i x(t-l_i) + \varepsilon(t), \end{aligned} \tag{3}$$

where $1 \leqslant l_1 < l_2 \cdots < l_w$, a_i are parameters to be determined, and $\varepsilon(t)$ is assumed to be IID Gaussian random variables, which are interpreted as fitting errors. The parameters a_i are chosen to minimize the sum of squares of the fitting errors. As Eq. (3) shows, RAR models can deal with both consecutive and separate time delays. Time delays in the RAR model correspond to the underlying periodicities in the data [3]. In this sense, RAR models explicitly indicate the periods in the data, and the data generated by RAR models have multiple *discrete* and *definite* timescales. For building RAR models from time series data, many candidate linear terms with different time delays are usually prepared in the form of a dictionary, and the linear terms that can extract the peculiarity of the time series as much as possible are selected [3,24,25]. We use the total error bottom-up method as an effective selection method, because the approach has proven to be effective in modeling nonlinear dynamics and can obtain better models in most cases than others with a reasonable computation time [31,32].

Selection algorithms usually employ some information criteria to find the optimal (best) model among many. It is often the case that the minimum of an information criterion corresponds to the optimal model size. For determining the optimal model we adopt the description length (DL) suitably

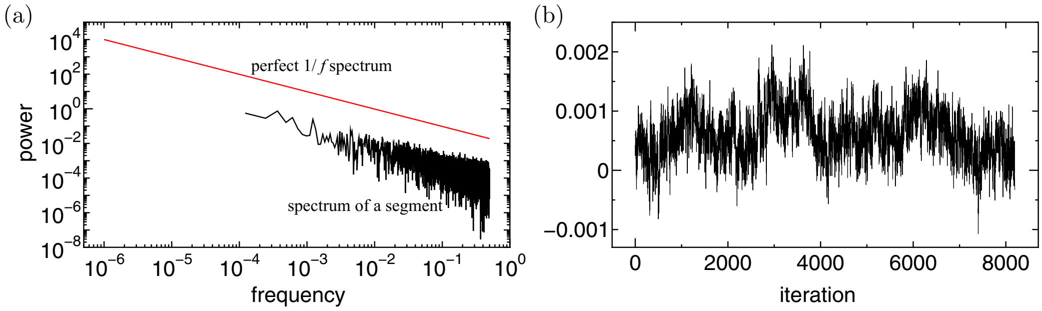

FIG. 2. A schematic diagram of a perfect $1/f$ spectrum and a segment from the time series exhibiting a perfect $1/f$ power spectrum. (a) The straight line is a perfect $1/f$ spectrum of 2^{20} (around 1×10^6) data points and the power spectrum of an arbitrary segment of 8192 data points of the signal in (a), where the assumed sampling rate is 1 Hz, the power spectrum is estimated by the FFT, and (b) a segment of 8192 data points used for the power spectrum estimate in (a).

modified by Judd and Mees [24], in the form

$$L(k) = \left(\frac{n}{2} - 1\right) \ln \frac{\mathbf{e}^T \mathbf{e}}{n} + (k+1)\left(\frac{1}{2} + \ln \gamma\right) - \sum_{i=1}^{k} \ln \delta_i, \tag{4}$$

where n is the length of the time series to be fitted, $\mathbf{e}$ stands for the vector composed from fitting errors, k is the number of parameters (or model size), γ is related to the scale of the data, and the variables δ can be interpreted as the relative precision to which the parameters are specified. The factor γ is a constant and typically fixed to be $\gamma = 32$. The concept of minimum description length (MDL) ensures that a RAR model built with an MDL modeling criterion will detect any periodicities present in the data, if the time series data are sufficiently long [3,33]. More thorough arguments for the details of the RAR model and the DL can be found in Refs. [3,24,25].

In the next section we apply the RAR modeling technique to $1/f$ noise data.

B. Generating $1/f$ noise using an optimal RAR model

As described above, since RAR models are linear models which can be composed of separate time delays, data generated by RAR models are inherently characterized by those separate periods [3]. In contrast, $1/f$ noise has been considered to have no characteristic timescale and no distinctive periodicity, since the $1/f$ power spectrum consists of significant contributions from a broad range of frequencies [6]. However, we demonstrate that a RAR model can generate data whose spectrum estimated by the DFT exhibits a clear $1/f$ spectrum in a wide frequency range.

To build a RAR model we start from an artificially generated genuine $1/f$ noise time series. It is possible to generate $1/f$ noise by various systems (models) [6,26,28,29]. These time series data are influenced by the systems. Whether data can be treated as $1/f$ noise is usually determined, not by the features of the data, but by the behavior of the power spectrum estimated by the DFT. Hence, we prefer, if possible, to generate the time series data purely conforming to the idea of the DFT without the influence of specific systems. To obtain such data we apply the technique of the linear surrogate data method [34]. We show the schematic diagram in Fig. 2. We first set a “perfect” $1/f$ spectrum as shown in Fig. 2(a). The amplitude at $f = 0$ corresponds to the mean of the time series and can be set arbitrary. We generate the randomized phases (while the power spectrum is completely preserved), and invert the transform using the randomized phases. Then we obtain time series data which have perfect $1/f$ power spectra [35]. As the inverse Fourier transform and the perfect $1/f$ spectrum are used in this approach, the time series is completely linear and stationary. Also, the time series definitely reflects all periodicities over the entire $1/f$ spectrum in terms of the DFT.

In practice, we can collect only a sample from the process over a finite time interval. Sampling a segment from a perfect $1/f$ time series produces a time series which is no longer perfect. Figure 2(b) shows a segment of the time series exhibiting the perfect $1/f$ spectrum. The behavior shows irregular fluctuations and a wave form similar to a slow undulation, and it is a characteristic of $1/f$ noise [5,6]. As Fig. 2(a) shows, the spectrum still exhibits a $1/f$ appearance. We use a segment from the perfect $1/f$ time series when building a RAR model.

Bassingthwaighte and Raymond suggested not to take a subsample of 1/2–1/8 the length of the total signal to avoid a spurious wraparound of the FFT generated signal [35,36]. Based on this consideration and to avoid unnecessary confounding factors, we use a segment much shorter than 1/8 the length of the perfect signal. We generate a perfect $1/f$ spectrum of 2^{30} (around 1×10^9) data points and use a segment of 300 000 data points for building RAR models (training data), where the data points are around 1/3600 the length of the perfect signal.

We note that the relationship between the maximum time delay l_W in the dictionary and the number of data points n is important. Even if the value of l_W is smaller than that of n, when l_W is large relative to n, the term with l_W is not often selected. In our experience we have found that around $l_W \times 100$ data points is large enough to select the term with l_W in most cases, although the number of necessary data points depends on the nature of the data. As we use 300 000 data points, we choose a largest time delay 3000, and the constant function gives 3001 candidate terms in the dictionary. Using the dictionary we build RAR models. As these training data are a segment of the perfect $1/f$ noise, we consider that these

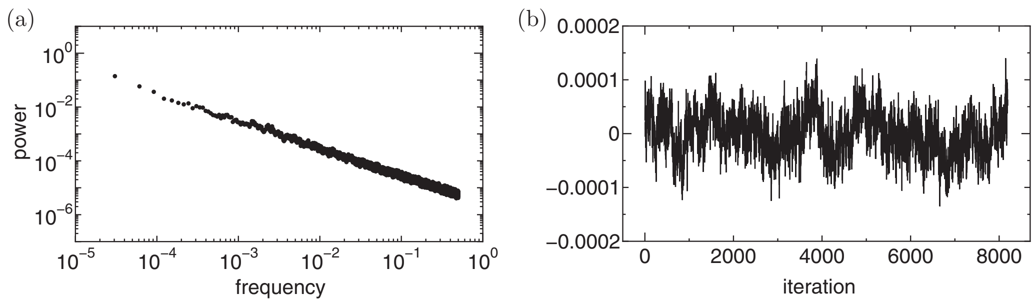

FIG. 3. Power spectrum with a double-logarithmic scale and a time series, where the data are generated by the optimal RAR model. (a) The power spectrum estimated by the FFT using 32 768 (2^{15}) data points at the assumed sampling rate 1 Hz, where the power spectrum is obtained by averaging 100 different data sets, and (b) a segment of 8192 data points used for the power spectrum estimate in (a).

data are composed of consecutive periodicities and do not have a characteristic timescale. Hence, we expect that the optimal RAR model might contain all terms with consecutive time delays up to the maximum delay l_W. However, only 25 terms are contained in the optimal RAR model, where the shortest and longest time delays are 1 and 2507, respectively [37].

As described above, each time delay in the RAR model corresponds to particular periodicities in the data [3]. Hence, when time delays in the model are not consecutive, the periods in the data should also be not continuous. It is commonly expected that the power spectrum of such data estimated by the DFT has peaks at frequencies corresponding to the periodicities. On the other hand, we can expect that as the optimal RAR model is built using the $1/f$ noise data, the RAR model might be able to generate data which have a $1/f$ spectrum, although the RAR model is only composed of mixed periodicities. To verify this expectation, we estimate the power spectrum of the data generated by the optimal RAR model using the FFT, assuming that the data are not contaminated by observational noise, where the number of data points is 32 768 (2^{15}). Hence, the frequency range of the power spectrum is between $1/32\,768$ (around $3.051\,758 \times 10^{-5}$) and 0.5. As the inverse number of the period is defined as the frequency, the inverse number of the time delay i of a time delay term $t - i$ corresponds to the frequency [3]. As the maximum time delay in the RAR model is 2507, the corresponding frequency is $1/2507$ (around $3.988\,831 \times 10^{-4}$). That is, the frequencies corresponding to all time delays except $t - 1$ in the RAR model fall within the frequency range of this power spectrum, and 32 768 data points should be large enough to detect these time delays in the RAR model. Contrary to a naive expectation, Fig. 3(a) shows that there is no clear peak of the individual power corresponding to the periodicities in the RAR model and an undoubted $1/f$ spectrum is observed in the whole frequency range (the slope of the least squares fit is $-1.001\,013$), where the power spectrum is an average of 100 time series data sets generated by the RAR model [38]. Figure 3(b) shows the time series generated by the RAR model. The behavior is similar to that of Fig. 2(b) because the time series has irregular fluctuations and a wave form similar to a slow waviness [5,6].

When we look at the power spectrum and the behavior of the data shown in Figs. 3(a) and 3(b) without the knowledge of the underlying dynamical system, we would conclude that the data are generated from a typical $1/f$ noise process as Eqs. (1) and (2). This result indicates that even if stationary linear models are composed of mixed periodicities, there are cases where the power spectrum estimated by the DFT can exhibit a clear $1/f$ spectrum in a wide frequency range.

V. CONCLUSION

The purpose of this paper is to consider a problem of source systems which generate time series data with long-range correlations, where the data and the systems are stationary and linear. We examine the correspondence between periodicities in the time series data and periodicities composing the source system when the data exhibit a $1/f$ spectrum or there are long-range correlations in the data.

An RAR model (stationary linear model with mixed periodicities) can generate data exhibiting a $1/f$ spectrum in a wide frequency range (data with long-range correlations). This fact indicates that continuous periodicities are not always necessary for linear models to generate data with long-range correlations. We note that while we demonstrate this with a relatively simple toy system (the linear RAR model driven by simple noise processes), data from real world sources are undoubtedly more complicated. Hence, we can consider that data with long-range correlations might be generated by more various systems than we expected in the real world. This result also leads us to reaffirm that the DFT basically provides information on periodicities of the time series data behavior and the DFT does not necessarily provide the information on periodicities of the source system that generates the data, even if the model and data are stationary linear, as it aggregates complex information.

ACKNOWLEDGMENTS

We would like to thank the anonymous referees for their very valuable remarks. We would like to acknowledge the partial support of JSPS KAKENHI Grant No. 25282094 and the Collaborative Research Program of Research Institute for Applied Mechanics, Kyushu University. T.T. also acknowledges the support of JSPS KAKENHI Grant No. 17K05590. M.S. is funded by Discovery Project (DP180100718).

[1] T. Ohira and T. Yamane, Phys. Rev. E **61**, 1247 (2000).
[2] M. B. Priestley, *Non-Linear and Non-Stationary Time Series Analysis* (Academic, London, 1988).
[3] M. Small and K. Judd, Phys. Rev. E **59**, 1379 (1999).
[4] J. B. Johnson, Phys. Rev. **26**, 71 (1925).
[5] M. B. Weissman, Rev. Mod. Phys. **60**, 537 (1988).
[6] N. J. Kasdin, Proc. IEEE **83**, 802 (1995).
[7] T. Musha, Int. J. Phys. Sci. **7**, 5717 (2012).
[8] V. Palenskis and K. Maknys, Sci. Rep. **5**, 18305 (2015).
[9] N. Leibovich, A. Dechant, E. Lutz, and E. Barkai, Phys. Rev. E **94**, 052130 (2016).
[10] We emphasize that in what follows the FFT algorithm itself is an innocent bystander.
[11] H. J. Jensen, K. Christensen, and H. C. Fogedby, Phys. Rev. B **40**, 7425 (1989).
[12] A. Paonita, Nonlinear Processes Geophys. **17**, 733 (2010).
[13] R. M. May, Nature (London) **261**, 459 (1976).
[14] M. A. Caloyannides, J. Appl. Phys. **45**, 307 (1974).
[15] R. F. Voss and J. Clarke, Nature (London) **258**, 317 (1975).
[16] T. Musha, Y. Kosugi, G. Matsumoto, and M. Suzuki, IEEE Trans. Biomed. Eng. **BME-28**, 616 (1981).
[17] M. Kobayashi and T. Musha, IEEE Trans. Biomed. Eng. **BME-29**, 456 (1982).
[18] T. Musha, H. Takeuchi, and T. Inoue, IEEE Trans. Biomed. Eng. **BME-30**, 194 (1983).
[19] D. L. Gilden, T. Thornton, and M. W. Mallon, Science **267**, 1837 (1995).
[20] E. Novikov, A. Novikov, D. Shannahoff-Khalsa, B. Schwartz, and J. Wright, Phys. Rev. E **56**, R2387 (1997).
[21] J. M. Hausdorff and C.-K. Peng, Phys. Rev. E **54**, 2154 (1996).
[22] V. N. Livina and T. M. Lenton, Geophys. Res. Lett. **34**, L03712 (2007).
[23] I. Eliazar and J. Klafter, Phys. Rev. E **82**, 021109 (2010).
[24] K. Judd and A. Mees, Physica D **82**, 426 (1995).
[25] K. Judd and A. Mees, Physica D **120**, 273 (1998).
[26] Y. Aizawa and T. Kohyama, Prog. Theor. Phys. **71**, 847 (1984).
[27] We initially investigated the smallest model size to generate data exhibiting a clear $1/f$ spectrum with 32 768 data points. We found that the model size 2500 is minimum. That is, data can exhibit a clear $1/f$ spectrum when the model size is larger than 2500.
[28] B. Kaulakys, Phys. Lett. A **257**, 37 (1999).
[29] R. Blender, X. Zhu, and K. Fraedrich, Adv. Sci. Res. **6**, 137 (2011).
[30] T. Nakamura, K. Judd, A. I. Mees, and M. Small, Int. J. Bifurcation Chaos **16**, 2153 (2006).
[31] T. Nakamura, D. Kilminster, K. Judd, and A. Mees, Int. J. Bifurcation Chaos **14**, 1129 (2004).
[32] For a strictly linear time delay combinatorial search, an exhaustive algorithm is feasible, although time consuming. We have found that an exhaustive search can find the correct models when there are correct models [30]. We have demonstrated elsewhere that the bottom-up approach is an expedient approach to an equivalent result in many cases [31].
[33] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, Series in Computer Science (World Scientific, Singapore, 1989).
[34] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).
[35] B. Pilgram and D. T. Kaplan, Physica D **114**, 108 (1998).
[36] J. B. Bassingthwaighte and G. M. Raymond, Ann. Biomed. Eng. **22**, 432 (1994).
[37] The pairs of time delays and parameters (l_i, a_i) in Eq. (3) are $(1, 0.586\,22)$, $(2, 0.054\,07)$, $(3, 0.079\,12)$, $(4, 0.024\,53)$, $(5, 0.033\,37)$, $(6, 0.012\,89)$, $(7, 0.018\,65)$, $(8, 0.012\,19)$, $(9, 0.014\,86)$, $(11, 0.019\,15)$, $(14, 0.017\,70)$, $(18, 0.012\,28)$, $(20, 0.008\,90)$, $(24, 0.011\,27)$, $(32, 0.010\,74)$, $(41, 0.012\,10)$, $(55, 0.010\,72)$, $(79, 0.009\,61)$, $(115, 0.008\,83)$, $(205, 0.007\,62)$, $(334, 0.006\,06)$, $(398, 0.006\,46)$, $(680, 0.005\,40)$, $(1202, 0.005\,98)$, and $(2507, 0.004\,92)$. The mean and the standard deviation of Gaussian dynamical noise $\varepsilon(t)$ is zero and $0.000\,021\,722$, respectively.
[38] An additional remark on the number of data points to generate $1/f$ noise by the RAR model is as follows. We confirmed that when the data points are smaller than 32 768, the power spectra estimated by the FFT exhibit a $1/f$ spectrum and the straight line as in Fig. 3(a). However, when the number of data points is larger than 32 768 (for example, 2^{16}), the straight line does not continue and the area considered as white noise appears in the lower-frequency area. We need a time delay larger than 2507 to obtain a $1/f$ spectrum in the wider frequency range.