Stationarization via surrogates

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Stationarization via surrogates

Pierre Borgnat and Patrick Flandrin

École Normale Supérieure de Lyon and CNRS, Laboratoire de Physique (UMR 5672 CNRS), Université de Lyon, 46, Allée d'Italie, 69364 Lyon Cedex 07, France
E-mail: pierre.borgnat@ens-lyon.fr and patrick.flandrin@ens-lyon.fr

Received 23 June 2008
Accepted 6 October 2008
Published 5 January 2009

Online at stacks.iop.org/JSTAT/2009/P01001
doi:10.1088/1742-5468/2009/01/P01001

Abstract. The surrogate data method, classically used for non-linearity tests, amounts to the use of some constrained noise providing a reference for statistical testing. It is revisited here as a method for stationarization and this feature is put forward in the context of non-stationarity testing. The stationarization property of surrogates is first explored in a time–frequency perspective and used for devising a test of stationarity relative to an observation time. Then, more general forms of surrogates are developed, directly in time–frequency or mixed domains of representation (ambiguity and time-lag domains included), and it is shown how they allow for other tests of non-stationary features: detection of the existence of a transient in some noise; assessment of non-stationary cross-correlations.

Keywords: stochastic processes (theory), new applications of statistical mechanics

J. Stat. Mech. (2009) P01001

 1742-5468/09/P01001+14$30.00

Contents

1. Introduction

When dealing with experimental data, making a distinction between stationary and non-stationary behaviors is often an important pre-processing step that may condition any subsequent analysis or modeling. Whereas the concept of stationarity (in short, independence of statistical properties with respect to some absolute time) seems to be unambiguous, its practical use turns out to be more subtle, with additional implicit assumptions regarding, e.g., observation scales and a need for statistical criteria of decision aimed at assessing the significance of observed fluctuations over time (or space) in a single observation. There have not been many works devoted to this question. In the statistical literature, stationarity tests have been proposed; see for instance [1, 2]. Some deal only with restrictive, parametric forms of non-stationarity (e.g., existence of trends, or variance evolution). Others use some specific assumption on the data (e.g., in [3]–[5]), or a parametric modeling [6], which is not generally adapted. Some studies have been focused on the testing of stationarity of a system using a reconstruction of the dynamics in an embedding state space where the lack of stationarity is associated with a change of recurrence times (or maps) [7]–[9]. This loss of recurrence is found practical as the basis of a stationarity test for signals that are outputs of dynamical systems, but operating the test is not straightforward and it is not easy to adapt it for large classes of signals. Another group of studies have put forward the concept of local stationarity of signals. The context is then more the detection of changes and segmentation of the signal in local stationary pieces [10]–[12]. The questions of testing more general forms of stationarity, for large classes of signals, have not been studied much. They have recently been revisited [13, 14], in a framework combining a time–frequency perspective [15] with a new use of the well-known technique of *surrogate data* [16, 17]. A main original feature of this work is to advocate the use of some empirical ingredient derived from the data to avoid having to resort to specific assumptions on the kind of stationarity that is tested, or on the nature of the signals. With the objective of testing stationarity, a basic ingredient is to introduce some 'controlled noise' in the problem so as to empirically characterize, in a data-driven way, the null hypothesis of stationarity. This will be the role given here to surrogates. Another original feature, which is the contribution in the last section of this article, is to propose new methods for directly designing surrogate data in the time–frequency framework—a possibility that has not been explored previously in the classical literature on surrogate data.

After recalling the classical method of surrogates, its new interpretation as a stationarization method is put forward in section 3. In section 4, a stationarity test framework using surrogates already put forward in [13] is outlined. An application to experimental data is proposed in section 5. Then, new variations are developed in section 6, introducing new versions of surrogates in time–frequency domains and devising methods using them for transient detection and non-stationarity cross-correlation assessment. A conclusion will then be drawn.

2. Revisiting surrogate data

Surrogates were first introduced by Theiler and co-authors [16] as a complement to statistical methods that test for non-linearity. This is a technique of resampling that creates new time series directly by manipulating the data. The leading idea is that, when facing experimental data and given a specific test of non-linearity of the system producing the data, one needs to assess statistically that some evidence for non-linearity is not a mere artifact of random statistical fluctuations. More precisely, one needs to find a way to obtain statistical knowledge about the null hypothesis of linearity, and derive from that a significant threshold for its rejection. Surrogate time series are obtained as new samples constrained to satisfy the null hypothesis, here linearity of the system, and keeping at the same time other relevant properties of the signals [17, 18].

A simple and elegant version of surrogates satisfying these properties was proposed in [16]. Second-order statistics of the original signal are kept but all other properties, specifically higher order statistics, are randomized. Given that correlations are the Fourier transforms of the spectrum (by the Wiener–Khintchin theorem), it turns out that keeping the spectrum of the signal (which is the squared amplitude of the direct Fourier transform) fulfils the constraint. In practice, given the original data $x(t)$, one first takes the Fourier transform $X(f) = (\mathbb{F}x)(f) = \int \mathrm{e}^{-\mathrm{i}2\pi tf} x(t)\,\mathrm{d}t$. The magnitude of $X(f)$ is kept unchanged while its phase is replaced by a random one φ_f, i.i.d. and uniformly distributed over $[-\pi, \pi]$. The inverse Fourier transform of this randomized distribution gives a surrogate time series:

$$s(t) = \int \mathrm{e}^{\mathrm{i}2\pi tf} |X(f)|\, \mathrm{e}^{\mathrm{i}\varphi_f}\, \mathrm{d}f. \qquad (1)$$

This simple procedure for surrogate data is efficient in many situations. This procedure is illustrated in figure 1. Many variations have been proposed since then. The review [17] gives a survey of the various surrogate techniques that have been proposed and improved along the years. For instance, one is aimed at preserving additional constraints such as

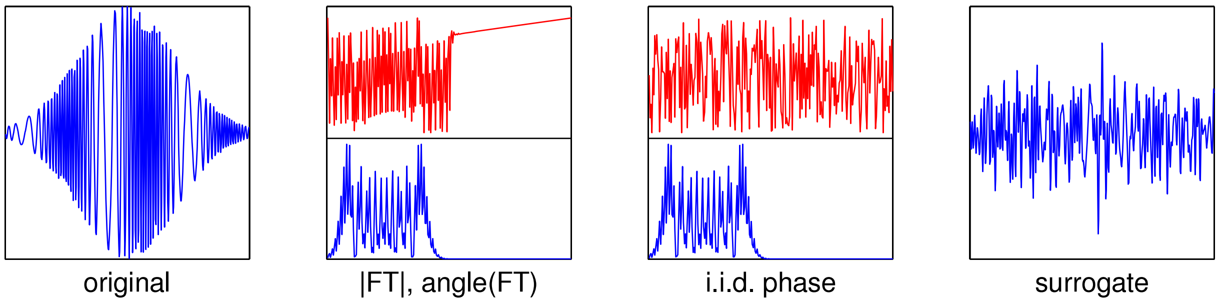

Figure 1. *Surrogate.* From left to right: original signal, its Fourier transform (top: phase, bottom: magnitude), the same after phase randomization, the surrogate.

the probability distribution [18]; another one was proposed for multivariate data, keeping the cross-spectrum of the data [19].

In this work, we propose to revisit surrogates in a less classical way, namely the use related to the property of stationarity.

3. Stationarization via surrogates

The theoretical definition of stationarity refers to a strict invariance of statistical properties under all and every time shift. In practical situations, this concept is loosely relaxed so as to encompass stationarity over some limited interval of observation, and deterministic stationarity (periodicity) as well as random stationarity. This agrees with the physical intuition associated with this notion. In order to test this property, the null hypothesis of stationarity is built directly from the data. More precisely, we want to construct a family of *stationarized* time series from the observation, each of them having a global frequency spectrum that exactly identifies with that of the data, while also being reproduced locally. Indeed, for the same spectrum density, 'non-stationary' signals differ from 'stationary' ones by temporal *structures* encoded in the spectrum phase. The simple surrogates of [16] are an adequate solution because one scrambles those temporal structures and keeps only the time-averaged spectrum as a constraint, hence stationarizing the signal.

Note that, to our knowledge, there seems to have been no or little consideration of the stationarizing property of surrogate data and, more precisely, of its possible use in the context of tests for stationarity. In a converse manner, Keylock studied in [20] surrogates constrained to retain the non-stationary of the signal, only evoking the possibility of using surrogates for stationarity testing.

A general framework for exhibiting this stationarizing property is the time–frequency perspective (see, e.g., [15]). Given a signal $x(t)$, an estimate of the time-varying spectrum at time t is given by the multitaper spectrogram [21]:

$$S_{x,K}(t,f) = \frac{1}{K}\sum_{k=1}^{K}\left|\int_{-\infty}^{+\infty} x(s)\, h_k(s-t)\, \mathrm{e}^{-\mathrm{i}2\pi f s}\, \mathrm{d}s\right|^2, \tag{2}$$

with K short-time windows $h_k(t)$ chosen as the K first Hermite functions. With $K=1$, one recovers the well-known spectrogram which is the squared magnitude of the short-term

J. Stat. Mech. (2009) P01001

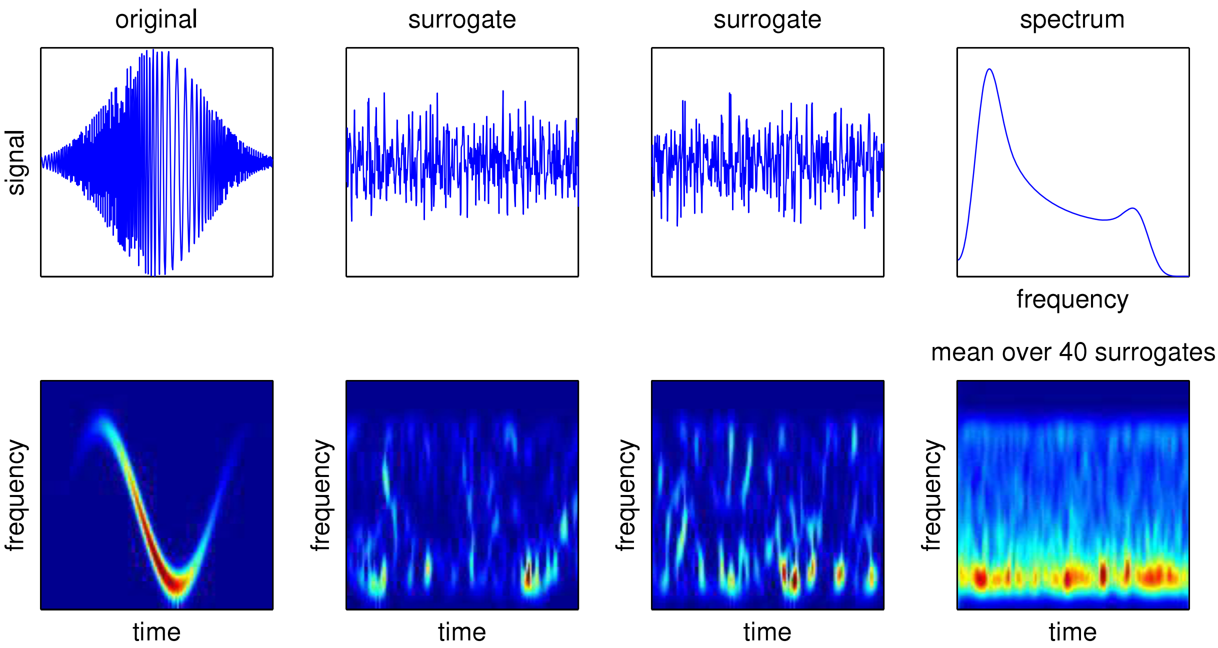

Figure 2. *Surrogates in time–frequency domains.* Top: a non-stationary signal and two representative surrogates, with their common time-averaged spectrum on the far right. Bottom: multitaper spectrogram of the signals. One sees that for surrogates, structures are destroyed and one recovers a time-varying spectrum that fluctuates around the mean stationary spectrum (displayed on the right). Spectrograms are color coded (red for large values, blue for zero). Parameters: $K = 5$; 512 points in time and frequency.

Fourier transform of x with window h_1. The advantage of the multitaper approach is that it provides a better *estimate* of the Wigner–Ville spectrum for stochastic processes, whereas it is a *reduced interference distribution* for deterministic signals. Indeed, the mean over K tapers results in reduced estimation variance without some extra time averaging that would be inappropriate in a non-stationary context. The choice of K was studied thoroughly in [22]. Here, averaged spectrograms are needed, so K should be larger than 1, but the value is not critical.

Figure 2 illustrates the result of this stationarization via surrogates. For a given (non-stationary) signal, one sees that its time–frequency distribution of power (estimated by a multitaper spectrogram) displays a clear organized structure and evolution along time, here a modulation both in amplitude and frequency. In contrast, a surrogate drawn from this signal reveals no specific structure in time: its spectrogram shows fluctuations, yet all seem to be around a mean stationary behavior as seen in the (ensemble-averaged) spectrogram displayed on the right. These are evidences of stationarity.

4. Stationarity testing with surrogates

Using surrogates to characterize the null hypothesis of stationarity, a test was proposed in [13] that amounts, in a time–frequency setting, to comparing *local* features versus *global* ones obtained by marginalization over time, relatively to a chosen observation scale. Using

J. Stat. Mech. (2009) P01001

a time–frequency framework with the objective of probing local ranges of stationarity is not new; it was used for instance for change detection [10, 11] and segmentation of stationary regions [12]. Contrasting global and local time–frequency features were also already present in [23, 24]; the novel feature of [13] is formalizing the idea for hypothesis testing by combining it with surrogates. First we recall this work and the next section will extend it and develop variations around this general methodology testing for some hypothesis related to (non-)stationary behaviors.

The principle of the test is contrasting instantaneous spectral features with global, time-averaged ones. In practice, multitaper spectrograms are evaluated at N time positions $\{t_n, n = 1, \ldots, N\}$, with a spacing $t_{n+1} - t_n$ which is a fraction of the width of the K windows $h_k(t)$. The number of tapers K is chosen between 5 and 10; the trade-off here is between smoothed estimates and the computational cost when K increases (the cost is linear in K). The method was validated in [13] with K as small as 5. The local contrast is computed as

$$c_n^{(x)} := \kappa\left(S_{x,K}(t_n, \cdot), \langle S_{x,K}(t_n, \cdot)\rangle_{n=1,\ldots,N}\right), \tag{3}$$

where κ is some suitable spectral distance. Studies in [13] have shown that a combination of a Kullback–Leibler distance with a log-spectral deviation offers a good measure of contrast in many situations. This distance reads as

$$\kappa\left(G, H\right) = \left(1 + \int \left|\log \frac{G(f)}{H(f)}\right| \mathrm{d}f\right) \cdot \int \left(\tilde{G}(f) - \tilde{H}(f)\right) \log \frac{\tilde{G}(f)}{\tilde{H}(f)} \, \mathrm{d}f, \tag{4}$$

where $\tilde{G}$ and $\tilde{H}$ are normalized versions of the spectrum. The fluctuations in time of these divergences $c_n^{(x)}$, computed as variances, give the test statistics:

$$\Theta_1 = L\left(c_n^{(x)}, \langle c_n^{(x)}\rangle_{n=1,\ldots,N}\right) := \frac{1}{N}\sum_{n=1}^{N}\left(c_n^{(x)} - \langle c_n^{(x)}\rangle\right)^2. \tag{5}$$

The distribution of the null hypothesis is provided by the same exact operation applied to a collection of the surrogates s_j:

$$\left\{\Theta_0(j) = L\left(c_n^{(s_j)}, \langle c_n^{(s_j)}\rangle_{n=1,\ldots,N}\right), j = 1, \ldots, J\right\}. \tag{6}$$

From this distribution, a one-sided test is derived where the threshold γ is obtained empirically from the null hypothesis built by the surrogates, after the specification of some false alarm percentage:

$$\begin{cases} \Theta_1 > \gamma & : \text{'non-stationarity'}; \\ \Theta_1 < \gamma & : \text{'stationarity'}. \end{cases} \tag{7}$$

Figure 3 illustrates the different steps of the test on some signal. One sees that it works well in differentiating non-stationary from stationary signals, the surrogates belonging to the class of stationary signals per construction. Note that a variant of the statistical test, using a one-class support vector machine instead of equations (3)–(7) to learn the stationary statistics, was proposed in [14]. More variations of the test are possible; our focus here is to show how one can adapt this framework to situations where other methods of surrogates are needed.

J. Stat. Mech. (2009) P01001

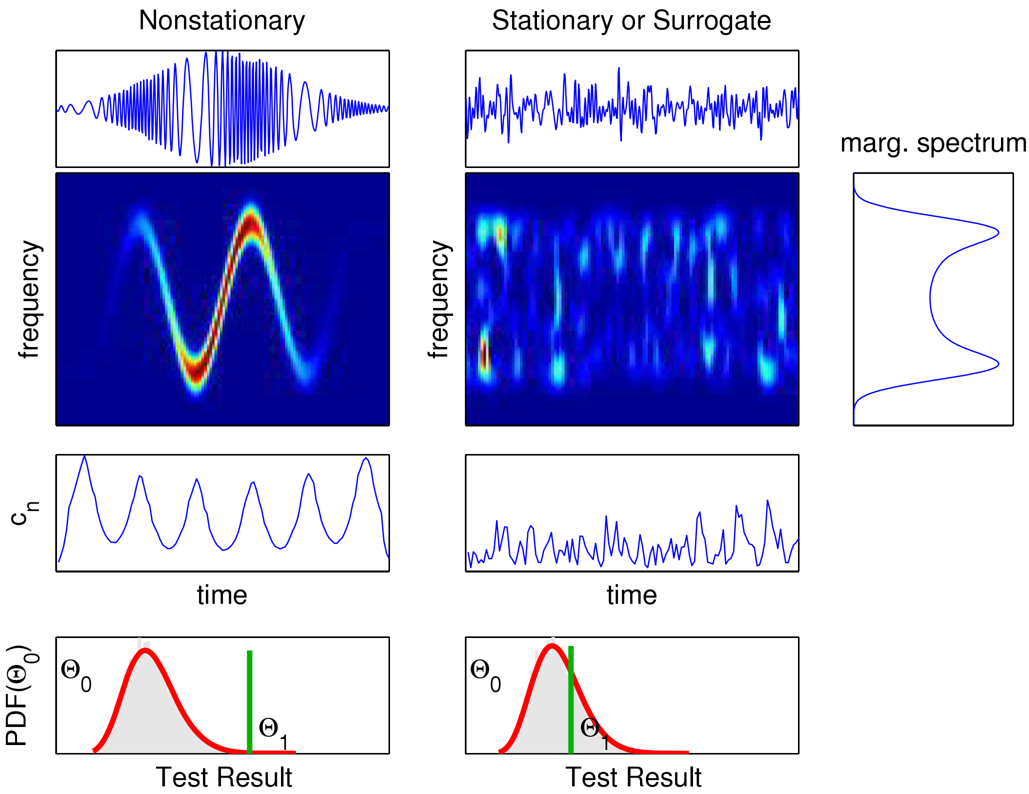

Figure 3. *Stationarity test.* From top to bottom: signal, multitaper spectrogram, local contrast c_n, and statistics of the test. Left: original (non-stationary) signal; right: stationary surrogate. Last row: the distribution of $\Theta_0(j)$ is displayed in gray (fitted with a Gamma law [13]). One sees that Θ_1 (the green line) is an outlier for the distribution on the left, and is inside the distribution for a stationary signal (on the right). Parameters: $K = 5$, $J = 50$.

5. Application of the stationarity test to experimental data

Let us show how the proposed method applies to experimental data. An example is taken here from experiments in dynamic light scattering used to investigate the internal dynamics of a living cell nucleus [26]. The reader is referred to [26] for the experimental set-up and the biological motivations. The measured raw signal is the scattered light intensity recorded as a function of time. A specific problem is that this raw signal display modulations at several timescales at once: slow modulations over several tens of seconds, short duration bursts around 7 s to 10 s, and fast fluctuations (see figure 4, top left). Hence, two aims are relevant: (1) automatically find the timescales of non-stationarity to extract them from the signal; (2) decide when the remaining fluctuations can be well-modeled by a stationary light scattering process, so that a well-defined auto-correlation function can be estimated. This auto-correlation function is indeed interesting because it was observed to be relevant to the internal dynamics of the nucleus in [26]. Those questions cannot be answered using stationarity tests from statistical literature (e.g., [1, 2]), due to the complexity

J. Stat. Mech. (2009) P01001

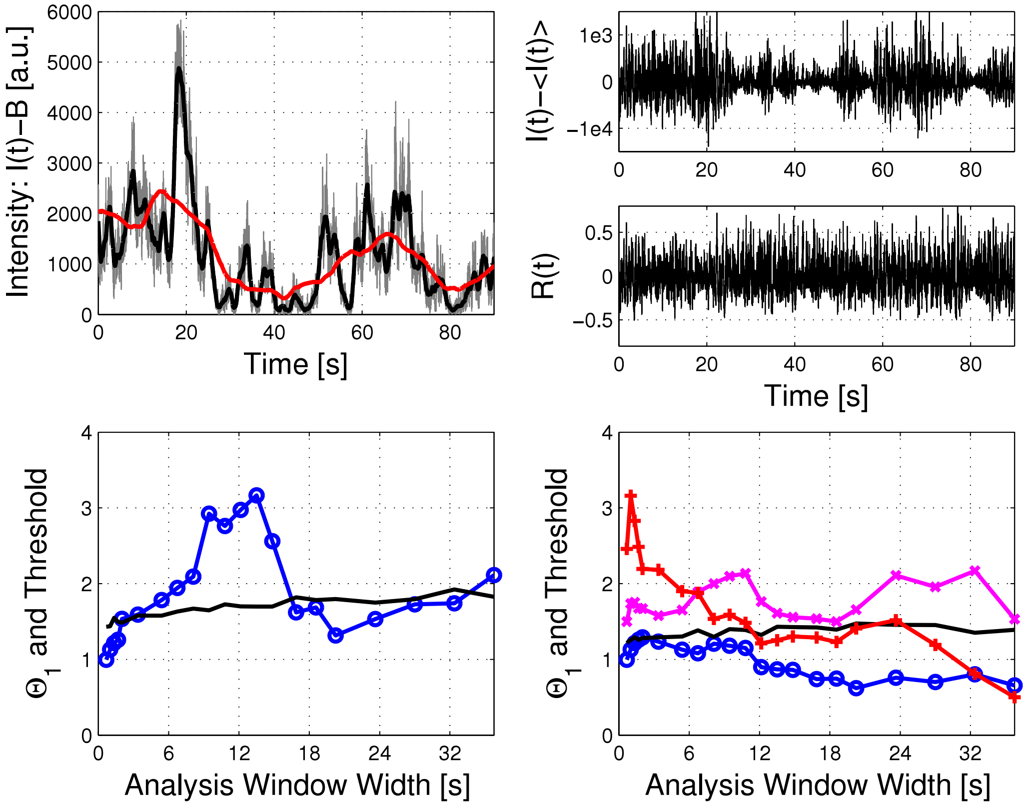

Figure 4. *Test of stationarity of dynamic light scattering.* Data from [26]: SHEP cell in the G1 phase; the acquisition frequency is 10^3 Hz. Top left: raw scattering intensity and moving average $\langle I(t)\rangle_T$ of the signal for $T = 1$ s (black) and 15 s (red; baseline B_0 removed). Bottom left: result of the stationarity test for $I(t)$ according to window width (threshold γ for stationarity in black; test statistics Θ_1 in blue). Top right: fluctuations $I(t) - \langle I(t)\rangle_T$, and (renormalized) fluctuations $R(t)$ (for B_0 here). Bottom right: result of the stationarity test for $R(t)$ (threshold in black), for different fixed baselines B (red: $B_0 - 600$, blue: B_0, magenta: $B_0 + 300$). Only with the correct one, B_0, does the test validate that the fluctuation is stationary for almost all analysis scales.

of the data that are not amenable to *a priori* modeling. Neither is it a question of segmenting the data in stationary pieces, but more of finding hierarchical scales of non-stationarity.

First, the stationarity test described in the previous section is applied to a signal of dynamic light scattering observed for 90 s (SHEP cell in the G1 phase; acquisition frequency is 10^3 Hz). The width of the analysis window in equation (2) is changed from 1 to 35 s and the stationarity test is repeated for each width. Hence this analysis takes the meaning of a stationarity test relative to the timescales of representation (and observation). The result is reported in figure 4, left column. There appear two regions of non-stationarity: one for scales larger than 32 s, and one from 6 to 15 s roughly. This validates an empirical conclusion of [26]: there exist slow modulations in the signal for scales around 10 s, and for scales larger than 30 s.

J. Stat. Mech. (2009) P01001

The second step is then to assess the stationarity of the fluctuating scattering signal, once the modulations are removed. For that, the signal is normalized as

$$R(t) = \frac{I(t) - \langle I(t)\rangle_T}{\langle I(t)\rangle_T - B}, \tag{8}$$

where $\langle I(t)\rangle_T$ is the moving average of $I(t)$ over T, and B is the baseline of the signal. The smoothing time T is fixed to 1 s here, so as to be certain that the non-stationarity evidenced by the first part of the analysis is accounted for in the slow modulation $\langle I(t)\rangle_T$. Unfortunately, the correct value of the baseline, B_0, is not known beforehand. The stationarity test gives the possibility of estimating it. Indeed, the property of stationarity is recovered for the fluctuations $I(t) - \langle I(t)\rangle_T$ only if normalized properly by the correct instantaneous mean intensity $\langle I(t)\rangle_T - B_0$. In figure 4, right column, we show how choosing the correct baseline B_0 provides a proper stationary signal for almost all widths of the analysis window (i.e., the representation timescale), whereas for an incorrect baseline, non-stationarity is still detected.

This example shows the potentiality of the proposed method as a tool for empirical data analysis. Here, the method was found to be efficient for ascertaining the existence of hierarchical scales or modulations, and disentangling their effects.

6. Surrogates in more general time–frequency contexts

Simple classical surrogates are often sufficient for probing the null hypothesis of stationarity. However, there exist situations where more refined schemes for surrogates are needed. In a sense, as seen in section 3, drawing surrogates amounts to keeping only the average spectrum of the original signal and imposing no further constraint. In other contexts, here for transient detection or cross-correlation analysis, this method is outperformed by ones acting directly in a time–frequency domain, so as to preserve features apparent in such a representation beside the averaged, marginal spectrum. The purpose of the current section is specifically aimed at defining constraints directly in the time–frequency plane and proposing surrogate methods that satisfy them.

6.1. Time–frequency surrogates

Time–frequency distributions are only one possible representation of the evolution in time of the energetic content of a signal, as it is well known from general lectures on the subject (e.g., the reader is referred to [15]). For instance, there is equivalence of the content displayed in three of the usual domains of representation:

(i) the time-lag domain, where the correlation is usually defined: $C_x(t,\tau) = \mathbb{E}\{x^*(t - \tau/2)x(t+\tau/2)\}$;

(ii) the time–frequency domain, which is the Fourier transform (over lag τ) of the correlation, providing the time-varying Wigner–Ville spectrum: $W_x(t,f) = \int \mathbb{E}\{x^*(t-\tau/2)x(t+\tau/2)\}\,\mathrm{e}^{-\mathrm{i}2\pi f t}\,\mathrm{d}\tau$;

(iii) the ambiguity domain, which is the Fourier transform of correlation over the time variable t: $A_x(\xi,\tau) = \int \mathbb{E}\{x^*(t-\tau/2)x(t+\tau/2)\}\,\mathrm{e}^{\mathrm{i}2\pi\xi t}\,\mathrm{d}t$.

Equivalent definitions exist for the cross-correlations, time-varying cross-spectrum and cross-ambiguity, by changing the first x in the other signal y. Classical surrogates manipulate the direct data in a linear way. We propose here to build surrogates in the 2D domains, via a direct phase randomization of the 2D representation.

For instance, if one needs to test for the significance of some feature of W_x in the time–frequency domain, the adaptation *mutatis mutandis* of the surrogate technique reads as follows: (i) do the (2D) Fourier transform of W_x (or its estimate through the multitaper spectrogram (2)) which is the ambiguity domain; (ii) keep its amplitude and replace the phase by an admissible phase, obtained as the phase of a realization of the ambiguity function of a white noise; (iii) come back into the time–frequency domain by inverting the Fourier transform. The constraint preserved in this method is the magnitude of the ambiguity, which is known to be associated with the correlations and the geometry of the time–frequency distribution [15].

A second constraint is meaningful for representations in the time–frequency domain: as it is of energetic, or quadratic, nature, one may expect it to satisfy a positivity constraint. The theoretical study of the positivity is beyond the scope here; however it is worth mentioning that the Wigner–Ville spectrum does not satisfy it [15]. This an additional motivation for the use of a smoothed estimate such as spectrograms which are always positive. For a fair comparison with spectrograms, the additional constraint for the surrogates is the positivity of the resulting representation. This is not ensured by the previous procedure. Following the lead of [17, 18], we propose an iterative method that was first presented in [25], which asymptotically corrects the representation toward positivity. Namely, the iterative algorithm from step n to step $n+1$ is sketched as follows:

(i) take the positive part $(S^{(n)}(t,f))_+$ of the time–frequency surrogate;
(ii) compute the ambiguity under positivity constraint: $\tilde{A}^{(n)}(\xi,\tau) = \mathbb{F}_t^{-1}\mathbb{F}_f(S^{(n)}(t,f))_+$;
(iii) form the new ambiguity as $A^{(n+1)} = |\tilde{A}^{(n)}|\,\mathrm{e}^{\mathrm{i}(\arg A^{(n)}+\delta\varphi^{(n)})}$, hence keeping the magnitude and adjusting the phase by some $\delta\varphi^{(n)}$ (more on this later on);
(iv) compute the new time–frequency surrogate: $S^{(n+1)}(t,f) = \mathbb{F}_\xi\mathbb{F}_\tau^{-1}(A^{(n+1)}(\xi,\tau))$,

until the positivity constraint is approximatively satisfied (given an *a priori* threshold on the negative amount in $S^{(n+1)}(t,f)$). The correction in phase may be operated in two different manners. The first one is a gradient descent method where $\delta\varphi^{(n)} = \lambda(\arg\tilde{A}^{(n)} - \arg A^{(n)})$ (with some $\lambda < 1$); this results in convergence of the iterations. A second method is to use a random correction of the phase by adding $\delta\varphi^{(n)} = \lambda^n \arg(A_{w^{(n)}})$ where $w^{(n)}$ is a newly synthesized white noise at step n and $\lambda < 1$. This method is similar to a simulated annealing convergence and is found to perform particularly well in practice.

6.2. Detection of transients

The aforementioned method for synthesizing time–frequency surrogates constrained in the ambiguity domain and positive is used with the objective of transient detection. The question is how to assess the level of significance of a ‘time–frequency’ patch in a fluctuating background. Testing against a null hypothesis of classical surrogates is inappropriate because, as is apparent in figure 5, the extent of time–frequency patches depends only on the marginal spectrum. This is more a characteristic of a specific geometry in the time–frequency domain, which is encoded in the ambiguity domain.

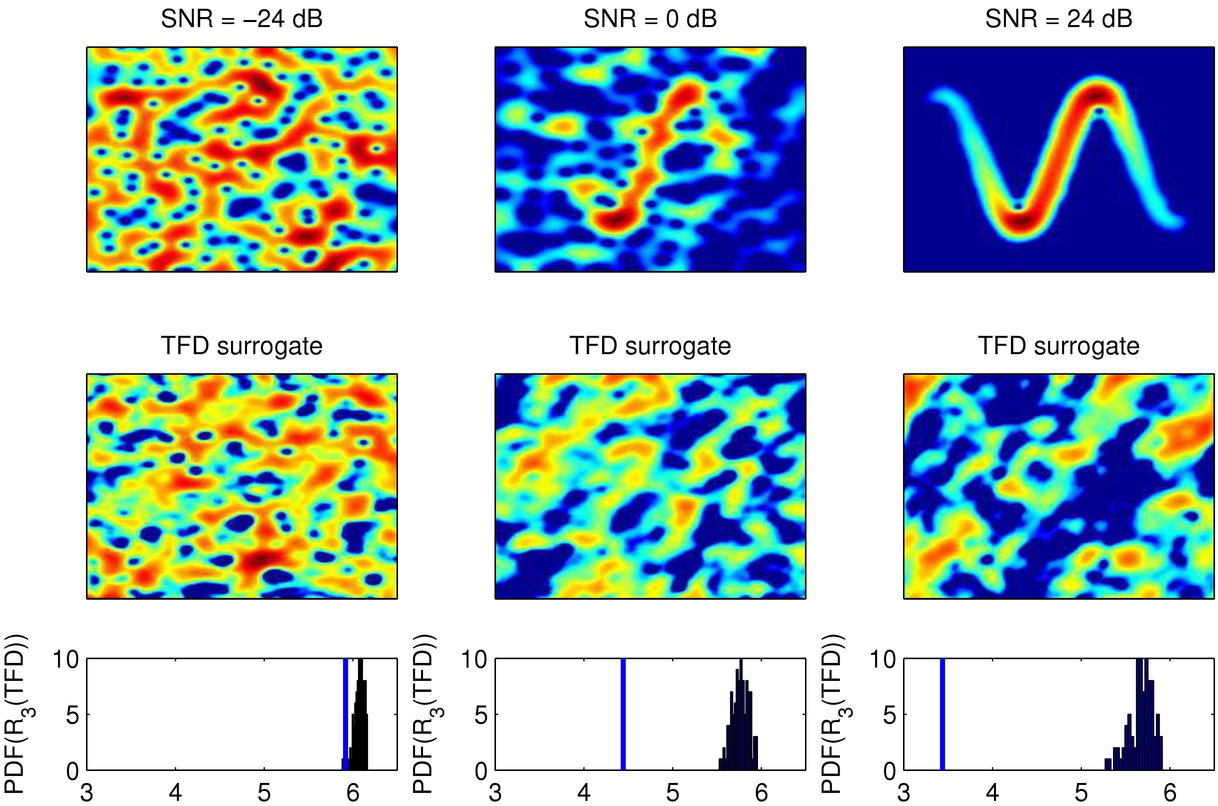

Figure 5. *Transient detection.* The detection test is applied to a chirp embedded in a white Gaussian noise with several SNR (−24, 0 and 24 dB). From top to bottom: TFD of the signal with noise, TFD surrogate with positivity constraint; statistics for the test (Rényi entropy of order 3): (blue) vertical bar for the signal, (black) histogram for the surrogates.

A test for the existence of a statistically significant transient is designed following the structure of the test of stationarity. The test statistics used here is the Rényi entropy of the time–frequency distribution [27]. Indeed, the entropy is relevant as a measure for estimating signal complexity and randomness in the time–frequency plane. It also roughly counts the number of components in a signal. Entropy of order $\alpha > 0$ of a time–frequency distribution S is defined as follows:

$$R_\alpha(S) = \frac{1}{1-\alpha} \int\int \left(\frac{S(t,f)}{\int\int S(t,f)\,\mathrm{d}t\,\mathrm{d}f} \right)^\alpha \mathrm{d}t\,\mathrm{d}f. \tag{9}$$

Note that Shannon entropy appears as $\alpha \to 1$. For a pure random signal, one expects a high entropy value; if a signal contains a transient, the resultant organization in the time–frequency domain should induce a smaller entropy.

An example of the result from the proposed detection test is shown in figure 5. A transient signal is embedded in white Gaussian noise for several SNR. The null hypothesis of the test is obtained using a collection of positive, time–frequency surrogates from the original noisy data. The statistics Θ is the Rényi entropy of order 3, which was found suitable (and numerically stable) in [27]. One sees that the chosen statistics is sensitive to the existence of a transient when it appears meaningfully above the fluctuation level

(SNR of 0 or 24 dB). For SNR = −24 dB, the transient is not significant in the noise and the test tells us so.

6.3. Detection of non-stationary cross-correlations

A second variation on non-stationary tests is the assessment of the existence and evolution of cross-correlations in multivariate data. Here the principle is mapped out from the test of stationarity and relies on a contrast between the local cross-correlations and the global (time-averaged) ones. We then compare the result obtained with the corresponding one using a collection of surrogates.

The difficulty here is selecting a relevant null hypothesis. Three candidates are compared in the following. First, individual and independent surrogates of the signals can be used, hence preserving only the marginal spectrum of each signal. Second, one may rely on the multivariate formulation of original surrogates for a signal, that keeps the marginal cross-spectrum [19], and hence the averaged cross-correlations. Despite this improvement, it does not keep the 'geometrical' structure of the quadratic distribution in the time-lag plane. The third proposition is to directly design time-lag surrogates for the cross-correlations, using the method of section 6.1 via a Fourier phase randomization of the 2D cross-correlations. Note that in this case, positivity is neither necessary nor a relevant constraint.

The chosen local measure of contrast is here the kurtosis of the divergence between local and averaged cross-correlations. Then, the variance of this contrast gives the statistics for the test. Figure 6 shows the result of the procedure for the correlated signals displayed at the top of the figure, depending on the type of surrogate that is used for the null hypothesis. One sees that independent or multivariate surrogates do not perform well for discerning the existence of significant cross-correlation patches. In contrast, 2D time-lag surrogates, preserving the mean geometric structure in the time-lag domain appear to provide a more relevant null hypothesis. This illustrates the potential interest of these 2D surrogates.

7. Conclusion

Time–frequency surrogates are introduced here as noises that are controlled and constrained to follow several properties of a given signal: the time-averaged spectrum for usual surrogates, or other constraints such as on the magnitude in the ambiguity domain and positivity are accessible by direct synthesis of 2D surrogates. The interesting property of stationarization that surrogates possess was stressed and its usefulness was illustrated using several variants of statistical testing for non-stationary features. This work is a step toward the use of these specific controlled noises in the context of non-stationary analysis.

A final word is that all stationarization techniques considered in this work involve some randomization, here in the Fourier domain of the signal or its representation in a suitable domain. An open issue in this use of controlled noise as an ingredient for statistical testing is the question of possible relationships with other resampling plans, such as bootstrap, jackknife, and cross-validation ones [28]. This question surfaces naturally and will be addressed in future works.

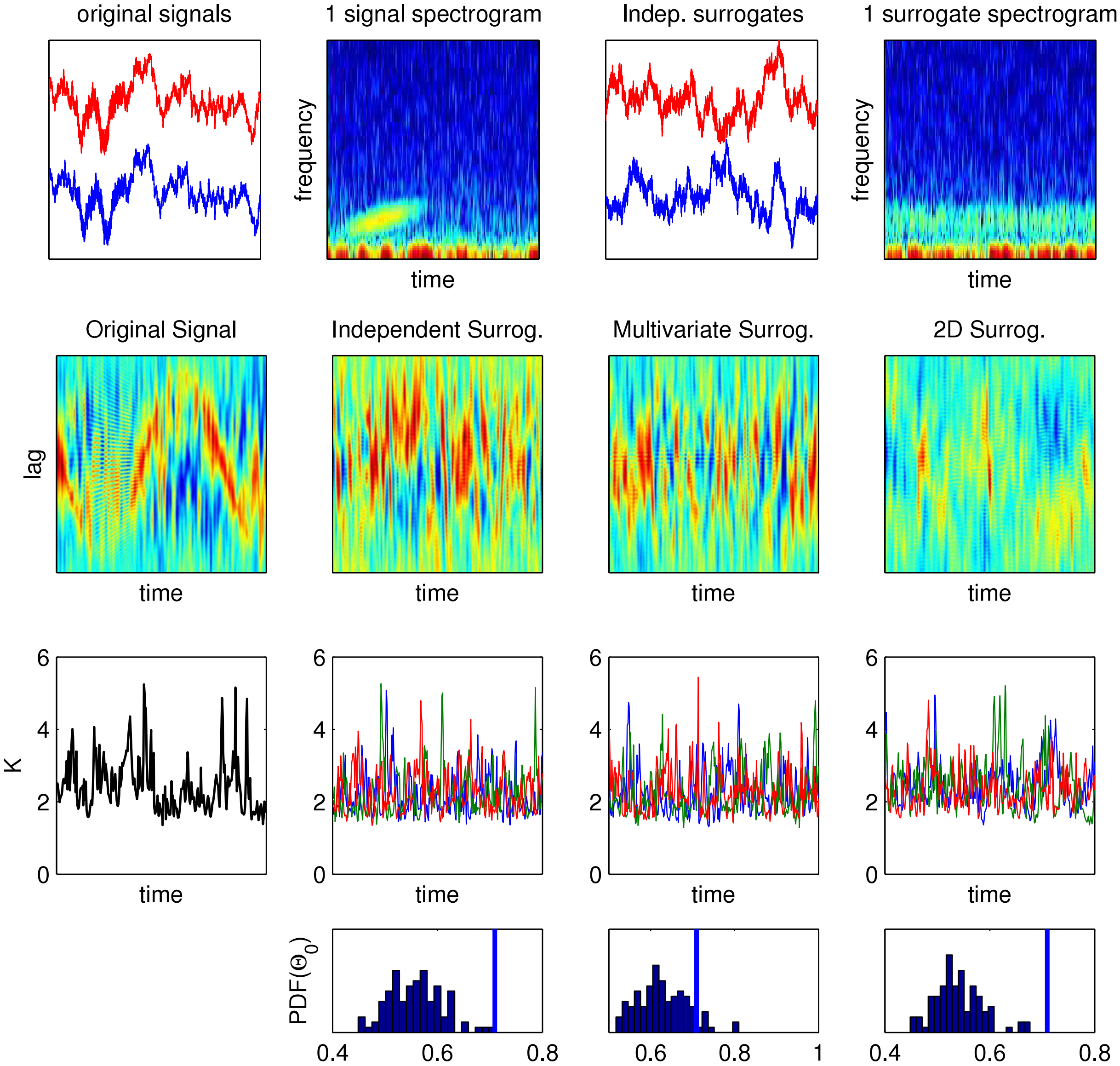

Figure 6. *Non-stationary cross-correlations.* Top (from left to right): original signals with non-zero cross-correlation (displayed underneath: mainly a time-varying sinusoidal delay), the spectrogram of one of them, two independent surrogates and the spectrogram of one of them. Second line: cross-correlations of original surrogate and of surrogates (independent, multivariate or in the time-lag domain from left to right). Third line: local contrast (measured here with kurtosis of the difference between the estimate of the cross-correlation and the time-averaged one). Bottom: PDF of the test statistics Θ_0 from the different versions of surrogates, and the statistics Θ_1 for the original signals (vertical bar).

Acknowledgments

This work is part of an undergoing project supported by ANR StaRAC and conducted with the collaboration of P-O Amblard (GIPSA-lab, Grenoble), C Richard (UTT, Troyes) and J Xiao (ENS Lyon & ECNU, Shanghai). The authors thank M Suissa, C Place, E Goillot and E Freyssingeas (ENS Lyon) for giving access to their data. The authors thank an anonymous referee for suggestions and comments, and for pointing out some references.

References

[1] Kwiatkowski D, Phillips P C B, Schmidt P and Shin Y, *Testing the null hypothesis of stationarity against the alternative of a unit root*, 1992 *J. Economet.* **54** 159

[2] Hobijn B, Franses P H and Ooms M, *Generalization of the KPSS-test for stationarity*, 2004 *Statis. Neerlandica* **58** 482

[3] Priestley M B and Rao T S, *A test for non-stationarity of time-series*, 1969 *J. R. Stat. Soc.* **31** 141

[4] Vaton S, *A new test of stationarity and its application to teletraffic data*, 1998 *Proc. IEEE ICASSP-98 (Seattle)* vol 6, pp 3449–52

[5] Brcich R F and Iskander D R, *Testing for stationarity in the frequency domain using a sphericity statistic*, 2006 *Proc. IEEE ICASSP-06 (Toulouse)* vol III, pp 464–7

[6] Kay S, *A new nonstationarity detector*, 2008 *IEEE Trans. Signal Process.* **56** 1440

[7] Rieke C, Sternickel K, Andrzejak R G, Elger C E, David P and Lehnertz K, *Measuring nonstationarity by analyzing the loss of recurrence in dynamical systems*, 2002 *Phys. Rev. Lett.* **88** 244102

[8] Rieke C, Andrzejak R G, Mormann F and Lehnertz K, *Improved statistical test for nonstationarity using recurrence time statistics*, 2004 *Phys. Rev.* E **69** 046111

[9] Verdes P F, Granitto P M and Ceccatto H A, *Overembedding method for modeling nonstationary systems*, 2006 *Phys. Rev. Lett.* **96** 118701

[10] Laurent H and Doncarli C, *Stationarity index for abrupt changes detection in the time–frequency plane*, 1998 *IEEE Signal Process. Lett.* **5** 43

[11] Mallat S, Papanicolaou G and Zhang Z, *Adaptive covariance estimation of locally stationary processes*, 1998 *Ann. Stat.* **24** 1

[12] Davy M and Godsill S, *Detection of abrupt signal changes using support vector machines: an application to audio signal segmentation*, 2002 *Proc. IEEE ICASSP-02 (Orlando, FL)*

[13] Xiao J, Borgnat P and Flandrin P, *Testing stationarity with time–frequency surrogates*, 2007 *Proc. EUSIPCO-07 (Poznan)*

[14] Xiao J, Borgnat P, Flandrin P and Richard C, *Testing stationarity with surrogates—a one-class SVM approach*, 2007 *Proc. IEEE Stat. Sig. Proc. Workshop SSP-07 (Madison, WI)*

[15] Flandrin P, 1999 *Time–Frequency/Time–Scale Analysis* (San Diego, CA: Academic)

[16] Theiler J, Eubank S, Longtin A, Galdrikian B and Farmer J D, *Testing for nonlinearity in time series: the method of surrogate data*, 1992 *Physica* D **58** 77

[17] Schreiber T and Schmitz A, *Surrogate time series*, 2000 *Physica* D **142** 346

[18] Schreiber T and Schmitz A, *Improved surrogate data for nonlinearity tests*, 1996 *Phys. Rev. Lett.* **77** 635

[19] Prichard D and Theiler J, *Generating surrogate data for time series with several simultaneously measured variables*, 1994 *Phys. Rev. Lett.* **73** 951

[20] Keylock C J, *Constrained surrogate time series with preservation of the mean and variance structure*, 2006 *Phys. Rev.* E **73** 030767.1

[21] Bayram M and Baraniuk R G, *Multiple window time-varying spectrum estimation*, 2000 *Nonlinear and Nonstationary Signal Processing* ed W J Fitzgerald *et al* (Cambridge: Cambridge University Press) pp 292–316

[22] Xiao J and Flandrin P, *Multitaper time–frequency reassignment for nonstationary spectrum estimation and chirp enhancement*, 2007 *IEEE Trans. Signal Process.* **55** 2851

[23] Martin W, *Measuring the degree of non-stationarity by using the Wigner–Ville spectrum*, 1984 *Proc. IEEE ICASSP-84 (San Diego, CA)* pp 41B.3.1–4

[24] Martin W and Flandrin P, *Detection of changes of signal structure by using the Wigner–Ville spectrum*, 1985 *Signal Process.* **8** 215

[25] Flandrin P, *Time–frequency surrogates*, 2008 *Invited talk at the AMS Joint Mathematics Mtg (San Diego, Jan. 2008)*

[26] Suissa M, Place C, Goillot E and Freyssingeas E, *Internal dynamics of a living cell nucleus investigated by dynamic light scattering*, 2008 *Eur. Phys. J.* E **26** 435

[27] Baraniuk R G, Flandrin P, Janssen A J E M and Michel O J J, *Measuring time–frequency information content using the Rényi entropies*, 2001 *IEEE Trans. Inf. Theory* **47** 1391

[28] Efron B, 1982 *The Jackknife, the Bootstrap and Other Resampling Plans* (Philadelphia, PA: SIAM)

J. Stat. Mech. (2009) P01001