**Astronomy
&
Astrophysics**

Unevenly-sampled signals: a general formalism for the Lomb-Scargle periodogram

R. Vio[1], P. Andreani[2,3], and A. Biggs[4]

[1] Chip Computers Consulting srl, Viale Don L. Sturzo 82, S. Liberale di Marcon, 30020 Venice, Italy
e-mail: robertovio@tin.it,
[2] ESO, Karl Schwarzschild strasse 2, 85748 Garching, Germany
[3] INAF - Osservatorio Astronomico di Trieste, via Tiepolo 11, 34143 Trieste, Italy
e-mail: pandrean@eso.org
[4] ESO, Karl Schwarzschild strasse 2, 85748 Garching, Germany
e-mail: abiggs@eso.org

ABSTRACT

The periodogram is a popular tool that tests whether a signal consists only of noise or if it also includes other components. The main issue of this method is to define a critical detection threshold that allows identification of a component other than noise, when a peak in the periodogram exceeds it. In the case of signals sampled on a regular time grid, determination of such a threshold is relatively simple. When the sampling is uneven, however, things are more complicated. The most popular solution in this case is to use the *Lomb-Scargle* periodogram, but this method can be used only when the noise is the realization of a zero-mean, white (i.e. flat-spectrum) random process. In this paper, we present a general formalism based on matrix algebra, which permits analysis of the statistical properties of a periodogram independently of the characteristics of noise (e.g. colored and/or non-stationary), as well as the characteristics of sampling.

Key words. methods: data analysis – methods: statistical

1. Introduction

Spectral analysis is a popular tool for testing whether a given experimental time series $\{x(t_0), x(t_1), \ldots, x(t_{M-1})\}$ contains only noise, i.e. $x(t_j) = n(t_j)$, or whether some other component $s(t)$ is present, i.e. $x(t) = s(t) + n(t)$. The classic approach is to fit the time series with the model function

$$x(t_j) = \sum_{k=0}^{N-1} a_k \cos(2\pi f_k t_j) + b_k \sin(2\pi f_k t_j) + n(t_j), \qquad (1)$$

$j = 0, 1, \ldots, M - 1$. If, for example, a periodic component $s(t)$ is present with a frequency f_l close to one in the set $\{f_k\}$, then the periodogram $\{\widehat{p_k}\}$,

$$\widehat{p_k} = a_k^2 + b_k^2, \qquad k = 0, 1, \ldots, N - 1, \qquad (2)$$

will show a prominent peak close to $k = l$. If $s(t)$ is semi-periodic or even non-periodic, the situation is more complicated since more peaks are expected. The main problem with the use of this technique is the definition of a detection threshold that fixes the contribution of noise in such a way that, when a peak exceeds it, the presence of a component $s(t)$ can be claimed. In the case of signals sampled on a regular time grid, the determination of such a threshold is a relatively simple procedure, but this is not the case when the condition of regularity does not apply. In this respect, several solutions have been proposed (see Lomb 1976; Ferraz-Mello 1981; Scargle 1982; Gilliland & Baliunas 1987; Reegen 2007; Zechmeister & Kürster 2009, and

references therein) that, however, work only under rather restrictive conditions (e.g. white and/or stationary noise) and are difficult to extend to more general situations.

In this paper, a general formalism is presented that allows analysis of the statistical properties of periodograms independently of the specific characteristics of the noise and the sampling of the signal. In Sect. 2 the formalism is presented for the case of even sampling and its extension to arbitrary sampling in Sect. 3. The usefulness of the proposed formalism is illustrated in Sect. 4, where the case of white noise with a mean different from zero and that of colored noise is calculated. In Sect. 5 the relationship between the periodogram and the least-squares method is considered. Finally, on the basis of simulated signals and a real time series, we discuss in Sect. 6 whether the use of algorithms specifically developed for computing the periodogram of unevenly-sampled time series is really advantageous.

2. Periodogram analysis in the case of even sampling

If a continuous signal $x(t)$ is sampled on a set of N equispaced time instants $t_0, t_1, \ldots, t_{N-1}$, a time series x_j, $j = 0, 1, \ldots, N - 1$ is obtained. Its discrete Fourier transform (DFT) is given by

$$\widehat{x_k} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-i2\pi k j/N}, \qquad k = 0, 1, \ldots, N - 1, \qquad (3)$$

with $i = \sqrt{-1}$. The sequence $\{x_j\}$ can be recovered from $\{\widehat{x}_k\}$ by means of

$$x_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \widehat{x}_k e^{2\pi k j/N}, \qquad j = 0, 1, \ldots, N-1. \quad (4)$$

The set $\{k/N\}_{k=0}^{N-1}$ provides the so-called *Fourier frequencies*. Implicit in the use of DFT is the assumption that $\{x_j\}$ is a periodic sequence with period $N\Delta t$ where $\Delta t = t_{j+1} - t_j$.

In matrix notation, Eqs. (3) and (4) can be written in the form

$$\widehat{x} = \mathbf{F}x \quad (5)$$

and

$$x = \mathbf{F}^*\widehat{x}. \quad (6)$$

Here, x and $\widehat{x}$ are column arrays that contain, respectively, the sequences $\{x_j\}$ and $\{\widehat{x}_k\}$, and $\mathbf{F}$ is the so-called "*Fourier matrix*", which is an $N \times N$ square, symmetric matrix whose (k, j)-entry[1] is given by

$$F_{kj} = \frac{1}{\sqrt{N}} e^{-i2\pi k j/N}. \quad (7)$$

The superscript "*" denotes the *complex conjugate transpose*. Matrix $\mathbf{F}$ is unitary, i.e.,

$$\mathbf{F}\mathbf{F}^* = \mathbf{F}^*\mathbf{F} = \mathbf{I}, \quad (8)$$

with $\mathbf{I}$ the identity matrix. Another useful property is that

$$\mathbf{F}\mathbf{F} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = \mathbf{F}\mathbf{F}^{\mathrm{T}} = \mathbf{F}^{\mathrm{T}}\mathbf{F}^{\mathrm{T}} = \mathbf{H} \quad (9)$$

with

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}. \quad (10)$$

By means of Eq. (9) it can be shown that

$$\mathbf{F}_{\mathfrak{R}}\mathbf{F}_{\mathcal{I}} = \mathbf{F}_{\mathfrak{R}}^{\mathrm{T}}\mathbf{F}_{\mathcal{I}} = \mathbf{F}_{\mathfrak{R}}\mathbf{F}_{\mathcal{I}}^{\mathrm{T}} = \mathbf{F}_{\mathfrak{R}}^{\mathrm{T}}\mathbf{F}_{\mathcal{I}}^{\mathrm{T}} = 0, \quad (11)$$

where $\mathbf{F}_{\mathfrak{R}} \equiv \mathfrak{R}[\mathbf{F}]$ and $\mathbf{F}_{\mathcal{I}} \equiv \mathcal{I}[\mathbf{F}]$ and where $\mathfrak{R}[.]$ and $\mathcal{I}[.]$ are the real and the imaginary parts of a complex quantity. Indeed,

$$\mathbf{F}\mathbf{F} = \mathbf{F}_{\mathfrak{R}}\mathbf{F}_{\mathfrak{R}} + \mathbf{F}_{\mathcal{I}}\mathbf{F}_{\mathcal{I}} + i2\mathbf{F}_{\mathfrak{R}}\mathbf{F}_{\mathcal{I}} = \mathbf{H}, \quad (12)$$

and $\mathbf{H}$ is a real matrix, hence Eq. (11) has to hold. From Eq. (3) it is also possible to see that $\widehat{x}_0$ is a real quantity. In the case that N is an even number, the same holds for $\widehat{x}_{N_\dagger+1}$ where $N_\dagger = \lceil N/2 \rceil$ and $\lceil z \rceil$ represents the smallest integer greater than z. Finally, dealing with $\widehat{x}$, only half of this array can be considered, since $\widehat{x}_{N-k} = \widehat{x}_k^*$, $k = 1, 2, \ldots, N_\dagger - \delta_{2N_\dagger}^N$, where $\delta_i^j = 1$ if $i = j$ and zero otherwise. With this notation, the periodogram of x is defined as

$$\widehat{p}_k = 2(\mathfrak{R}[\widehat{x}_k]^2 + \mathcal{I}[\widehat{x}_k]^2) = 2|\widehat{x}_k|^2, \quad k = 0, 1, \ldots, N_\dagger - 1, \quad (13)$$

where $\widehat{x}_k$ is the kth entry of the array $\widehat{x}^{\mathrm{T}} = [\widehat{x}_0, \widehat{x}_1, \ldots, \widehat{x}_{N_\dagger-1}]$[2] and $|.|$ denotes the *Euclidean* norm. By means of

$$\widehat{z} = \begin{pmatrix} \widehat{\underline{x}}_{\mathfrak{R}} \\ \widehat{\underline{x}}_{\mathcal{I}} \end{pmatrix}, \quad (14)$$

a column array obtained by the column concatenation of $\widehat{x}_{\mathfrak{R}} \equiv \mathfrak{R}[\widehat{\underline{x}}]$ and $\widehat{x}_{\mathcal{I}} \equiv \mathcal{I}[\widehat{\underline{x}}]$, Eq. (13) can be rewritten in the form

$$\widehat{p}_k = 2\left(\widehat{z}_k^2 + \widehat{z}_{N_\dagger+k}^2\right), \qquad k = 0, 1, \ldots, N_\dagger - 1. \quad (15)$$

In Sect. 5 it is shown that this periodogram is equivalent to that obtainable by means the least-squares fit of model (1) with $M = N$, $t_j = j$ and $f_k = k/N$.

An important point to stress is that, if x is the realization of a (not necessarily Gaussian) random process, then each $\widehat{x}_k$ is given by the sum of N random variables. This is because of the linearity of the *Fourier operator* $\mathbf{F}$. Thanks to the *central limit theorem*, therefore, the entries of $\widehat{x}$ can be expected to be Gaussian random quantities. As a consequence, the entries of $\widehat{z}$ can also be expected to be Gaussian random quantities with covariance matrix $\mathbf{C}_{\widehat{z}} = \mathrm{E}[\widehat{z}\widehat{z}^{\mathrm{T}}]$ given by

$$\mathbf{C}_{\widehat{z}} = \begin{pmatrix} \underline{\mathbf{F}}_{\mathfrak{R}}\mathbf{C}_x\underline{\mathbf{F}}_{\mathfrak{R}}^{\mathrm{T}} & \underline{\mathbf{F}}_{\mathfrak{R}}\mathbf{C}_x\underline{\mathbf{F}}_{\mathcal{I}}^{\mathrm{T}} \\ \underline{\mathbf{F}}_{\mathcal{I}}\mathbf{C}_x\underline{\mathbf{F}}_{\mathfrak{R}}^{\mathrm{T}} & \underline{\mathbf{F}}_{\mathcal{I}}\mathbf{C}_x\underline{\mathbf{F}}_{\mathcal{I}}^{\mathrm{T}} \end{pmatrix}. \quad (16)$$

Here, E[.] denotes the *expectation operator*, $\mathbf{C}_x = \mathrm{E}[xx^{\mathrm{T}}]$ is the covariance matrix of x, and $\underline{\mathbf{F}}$ the matrix obtained by the first $N_\dagger$ rows of the *Fourier matrix* $\mathbf{F}$. From Eqs. (16) and (11), it is easy to deduce that, if x is the realization of a standard white-noise process, i.e., $\mathbf{C}_x = \mathbf{I}$, then $\mathbf{C}_{\widehat{z}}$ is a diagonal matrix with $(\mathbf{C}_{\widehat{z}})_{11} = 1$, $(\mathbf{C}_{\widehat{z}})_{N_\dagger N_\dagger} = 0$ and $(\mathbf{C}_{\widehat{z}})_{kk} = 0.5$. In other words, the entries of $\widehat{z}$ are mutually uncorrelated. In turn, this means that $\mathfrak{R}[\widehat{x}_k]$ is uncorrelated with $\mathcal{I}[\widehat{x}_k]$. If x is a colored (not necessarily stationary) noise process, i.e. $\mathbf{C}_x$ is not a diagonal matrix, then this holds also for $\mathbf{C}_{\widehat{z}}$. However, from x it is possible to obtain an array y containing mutually uncorrelated entries by means of the transformation

$$y = \mathbf{C}_x^{-1/2}x. \quad (17)$$

The matrix $\mathbf{C}_x^{-1/2}$ can be computed via

$$\mathbf{C}_x^{-1/2} = \mathbf{U}^{\mathrm{T}}\mathbf{\Sigma}^{-1/2}\mathbf{U} \quad (18)$$

with $\mathbf{U}$ the orthogonal matrix whose columns contain the eigenvectors of $\mathbf{C}_y$ and $\mathbf{\Sigma}$ a diagonal matrix containing the corresponding eigenvalues $\{\lambda_l\}_{l=0}^{N-1}$. This decomposition is particularly simple and computationally efficient if $\mathbf{C}_x$ is a *circulant* matrix since it can be diagonalized according to

$$\mathbf{C}_x = \mathbf{F}^*\mathrm{diag}[\mathbf{F}c]\mathbf{F}, \quad (19)$$

where "diag[q]" denotes a diagonal matrix whose diagonal entries are given by the array q and c is the first column of $\mathbf{C}_x$. Because of this, the decomposition (18) can be directly computed with $\mathbf{U} = \mathbf{F}$ and $\mathbf{\Sigma} = \mathrm{diag}[c]$; hence, $y = \mathbf{F}^*\mathbf{\Sigma}^{-1/2}\mathbf{F}$. This means that the whitening operation can be performed in the harmonic domain through the following procedure: a) computation of $\widehat{x}$, i.e. the DFT of x; b) computation of $\{\widehat{y}_k\} = \{\widehat{x}_k/\lambda_k^{1/2}\}$; and c) the inverse DFT of $\widehat{y}$. A potential difficulty in using transformation (17) is that $\mathbf{C}_x$ is required to be of full rank. Such a

---

[1] In the following, the element in the nth row and mth column of an $N \times M$ matrix $\mathbf{A}$ will be indicated with A_{mn} or alternatively with $(\mathbf{A})_{mn}$, $n = 0, 1, \ldots N - 1$, $m = 0, 1, \ldots, M - 1$.

[2] From now on, if r is an $N \times 1$ column array, then $\underline{r}$ is a column array that contains the first $N_\dagger = \lceil N/2 \rceil$ entries of r, i.e. $\underline{r} = [r_0, r_1, \ldots, r_{N_\dagger-1}]^{\mathrm{T}}$. Similarly, if $\mathbf{A}$ is an $N \times M$ matrix, then $\underline{\mathbf{A}}$ is a matrix that contains the first $N_\dagger$ rows of $\mathbf{A}$.

condition can be expected to be satisfied in most practical applications. Some problems can arise if the time step of the sampling is much shorter than the decorrelation time of n[3]. Indeed, some columns (and therefore some rows) of $\mathbf{C}$ could almost be identical i.e., this matrix could become numerically *ill-conditioned*. In this case, the most *natural* solution consists of averaging the data that are close in time.

When the periodogram is used to test whether $x = n$ vs. $x = s + n$, a threshold $L_{\widehat{p}_k}$ has to be defined such that, with a pre-fixed probability, a peak in $\widehat{p}_k$ that exceeds $L_{\widehat{p}_k}$ can be expected to not arise because of the noise. This requires knowledge of the statistical properties of $\widehat{p}$ under the hypothesis that $x = n$. The simplest situation is when n is the realization of a standard white-noise process. In fact, since the entries of $\widehat{z}$ are uncorrelated random Gaussian quantities, from Eq. (15) it can be derived that the entries of $\widehat{p}$ are (asymptotically) independent quantities distributed according to a χ_2^2 distribution[4]. As a consequence, independent of the frequency k, a threshold L_{Fa} can be determined that corresponds to the level that a peak due to the noise would exceed with a prefixed probability α when a number N_f of (*statistically independent*) frequencies are inspected. More specifically, L_{Fa} is the highest value for which $1-[1-\exp(-L_{\widehat{p}_k})]^{N_f} \le \alpha$ (Scargle 1982), in formula

$$L_{\mathrm{Fa}} = \sup_{L_{\widehat{p}_k}} \left\{ 1 - [1 - \exp(-L_{\widehat{p}_k})]^{N_f} \le \alpha \right\}. \tag{20}$$

If the entire periodogram is inspected, then $N_f = N_\dagger$. Commonly, L_{Fa} is called the *level of false alarm*.

Threshold (20) is not applicable when the noise is colored. However, the requirement to fix a different level for each frequency can be avoided if the original signal x is transformed into

$$y = \mathbf{C}_n^{-1/2} x. \tag{21}$$

Indeed, under the hypothesis that $x = n$, the entries of y are uncorrelated and unit-variance random quantities. As seen above, this operation should not be difficult. Some problems emerge if the decomposition (18) is carried out by means of the efficient DFT approach described above (a necessary approach in the case of very long sequences of data). This is because, in forming $\mathbf{C}_n$, it is necessary to take the periodicity of the sequence x forced by the DFT into account. In other words, it is necessary to impose a spurious correlation among the first and the last entries in n. For instance, for a stationary noise process, $\mathbf{C}_n$ is a *Toeplitz* matrix, but it has to be approximated with a *circulant* one. When dealing with sampled signals, this is an unavoidable problem that no technique can completely solve. A classical solution to relieving this situation is the *windowing method*, i.e., the substitution of x_j in Eq. (3) with $\eta_j x_j$, where $\{\eta_j\}$ is some prefixed discrete function (window) that makes the signal gently reduce to zero at the extreme of the sampling interval (e.g. see Oppenheim & Shafer 1989). This method can be expected to work satisfactorily only when the decorrelation time of noise is shorter than the length of the signal.

3. Periodogram analysis in the case of uneven sampling

If a signal $x(t)$ is sampled on an uneven set of time instants, some problems emerge: it is no longer possible to define a set of

"*natural*" frequencies such as those obtained by the *Fourier* transform. In turn, this implies some ambiguities in the definition of the Nyquist frequency that, loosely speaking, corresponds to the highest frequencies that contain information on the signal of interest (e.g. see Vio et al. 2000; Koen 2006). As a consequence, Eq. (3) has to be modified. In the following, with no loss of generality, it is assumed that $x(t)$ is sampled at M arbitrary time instants $t_0, t_1, \ldots, t_{M-1}$ with $t_0 = 0$, $t_{M-1} = M-1$ and the remaining t_j arbitrarily distributed within this interval. Moreover, a set of N frequencies $k = 0, 1, \ldots, N - 1$ is considered with $N \gtreqless M$. Such a set corresponds to the frequencies that are typically inspected when looking for a periodicity. However, others can be chosen. With these conditions, a transformation corresponding to the one given by Eq. (5) is

$$\widehat{x} = \mathcal{F} x, \tag{22}$$

where

$$\mathcal{F}_{kj} = \frac{1}{\sqrt{M}} e^{-i2\pi k \tilde{t}_j / N}, \tag{23}$$

$\tilde{t}_j = t_j/\Delta_m t$, $\Delta_m t = \gamma \min[\{t_{j+1} - t_j\}]$ and γ is a real positive number. Because of the normalization by $\Delta_m t$, time $\tilde{t}_j$ is expressed in units of the shortest sampling time interval. Apart from the substitution of the *Fourier* matrix $\mathbf{F}$ with $\mathcal{F}$, the uneven sampling of signals does not modify the formalism introduced in the previous section. In particular, the covariance matrix $\mathbf{C}_{\widehat{z}}$ defined in Eq. (16) becomes

$$\mathbf{C}_{\widehat{z}} = \begin{pmatrix} \mathcal{F}_{\Re} \mathbf{C}_x \mathcal{F}_{\Re}^{\mathrm{T}} & \mathcal{F}_{\Re} \mathbf{C}_x \mathcal{F}_{I}^{\mathrm{T}} \\ \mathcal{F}_{I} \mathbf{C}_x \mathcal{F}_{\Re}^{\mathrm{T}} & \mathcal{F}_{I} \mathbf{C}_x \mathcal{F}_{I}^{\mathrm{T}} \end{pmatrix}. \tag{24}$$

The $N \times M$ *Fourier matrix* $\mathcal{F}$ does not have the properties (8)−(12). As consequence, and also in the case that x is the realization of a standard white-noise process i.e. $\mathbf{C}_x = \mathbf{I}$, matrix $\mathbf{C}_{\widehat{z}}$,

$$\mathbf{C}_{\widehat{z}} = \begin{pmatrix} \mathcal{F}_{\Re} \mathcal{F}_{\Re}^{\mathrm{T}} & \mathcal{F}_{\Re} \mathcal{F}_{I}^{\mathrm{T}} \\ \mathcal{F}_{I} \mathcal{F}_{\Re}^{\mathrm{T}} & \mathcal{F}_{I} \mathcal{F}_{I}^{\mathrm{T}} \end{pmatrix}, \tag{25}$$

is not diagonal. In general, $\mathbf{C}_{\widehat{z}}$ is not even diagonalizable. For example, this happens when the periodogram of a time series containing M data is computed on N frequencies with $N > M$ (a typical situation in practical applications). This implies that the entries of $\widehat{z}$ cannot be made mutually uncorrelated. Obviously, the same holds for the entries of $\widehat{p}$ as given by Eq. (15). As a consequence, although a number N of frequencies are considered in $\widehat{p}$, at most only $M/2$ of them are statistically independent[5]. Particularly troublesome is that, for a given frequency k, $\Re[\widehat{x}_k]$, and $I[\widehat{x}_k]$ (i.e. $\widehat{z}_k$ and $\widehat{z}_{N_\dagger+k}$) are also correlated. This makes it difficult to fix the statistical characteristics of $\widehat{p}_k$. In this respect, two choices are possible. The first consists in the determination, for each frequency, of the PDF of $\widehat{p}_k$. Actually, this is a rather involved approach, because $\widehat{z}_k$ and $\widehat{z}_{N_\dagger+k}$ have variance $(C_{\widehat{z}})_{kk}$ and $(C_{\widehat{z}})_{N_\dagger+k,N_\dagger+k}$, respectively, and covariance $(C_{\widehat{z}})_{k,N_\dagger+k}$. Therefore, once $\widehat{z}_k$ and $\widehat{z}_{N_\dagger+k}$ are normalized to unit variance, each $\widehat{p}_k$ is given by the sum of two correlated χ_1^2 random quantities. Although available in analytical form (Simon 2006), the resulting PDF is rather complex and hence difficult to handle

---

[3] The decorrelation time of a random signal $n(t)$ is the time interval Δt such that two values $n(t_1)$ and $n(t_2)$, with $t_2 - t_1 = \Delta t$, can be considered as uncorrelated.

[4] χ_2^2 denotes the chi-square distribution with two degrees of freedom.

[5] This is because, if $N > M$, the rank of the $N \times N$ matrix $\mathbf{C}_{\widehat{z}}$ is smaller than or equal to M. This implies that the array $\widehat{z}$ has at most M degrees of freedom. Since each entry of $\widehat{p}$ is given by the sum of two entries of $\widehat{z}$, then a periodogram has at most $M/2$ degrees of freedom.

(for an alternative approach, see Reegen 2007). Moreover, there is the additional problem that $L_{\widehat{p}_k}$ changes with k. A simpler alternative is the use of two uncorrelated and unit-variance random quantities, $\widehat{v}_k$ and $\widehat{v}_{N_\dagger + k}$, obtained through the transformation

$$\widehat{v} = \Sigma_\star^{-1/2} \mathbf{U}_\star \widehat{z}_\star. \tag{26}$$

Here, $\widehat{v}^{\mathrm{T}} = [\widehat{v}_k, \widehat{v}_{N_\dagger + k}]$, $\widehat{z}_\star^{\mathrm{T}} = [\widehat{z}_k, \widehat{z}_{N_\dagger + k}]$, $\Sigma_\star^{-1/2}$ is a diagonal matrix whose entries are given by the reciprocal of the square root of the non-zero eigenvalues of the covariance matrix

$$\mathbf{C}_\star = \begin{pmatrix} (C_{\widehat{z}})_{kk} & (C_{\widehat{z}})_{k,N_\dagger + k} \\ (C_{\widehat{z}})_{k,N_\dagger + k} & (C_{\widehat{z}})_{N_\dagger + k, N_\dagger + k} \end{pmatrix}, \tag{27}$$

zero otherwise, and $\mathbf{U}_\star$ is an orthogonal matrix that contains the corresponding eigenvectors[6]. Indeed, if the periodogram is defined as

$$\widehat{p}_k = \widehat{v}_k^2 + \widehat{v}_{N_\dagger + k}^2, \qquad k = 0, 1, \ldots, N_\dagger - 1, \tag{28}$$

then each $\widehat{p}_k$ is given by the sum of two independent, unit-variance, Gaussian random quantities. As a consequence, the corresponding PDF is, independently of k, a χ_2^2 whose *cumulative distribution function* (CDF) is the exponential function. This permits determining the statistical significance of $\widehat{p}_k$ for a *specified* frequency k. Things become more complex if N_f frequencies are inspected when looking for a peak. Indeed, also after the operation (26), it happens that $\mathrm{E}[\widehat{p}_k \widehat{p}_l] \neq 0$ for $k \neq l$, i.e., the frequencies of the periodogram remain mutually correlated. This is an unavoidable problem. Because of it, N_f does not correspond to the number of *independent frequencies*, so the *level of false alarm* (20) cannot be computed. However, since $L_{\widehat{p}_k}$ is the same for all the frequencies, an upper limit can be fixed for L_{Fa} by setting $N_f = \lceil M/2 \rceil$. The periodogram obtained by means of Eq. (28) corresponds to the original *Lomb-Scargle* periodogram.

Since the transformation (17) does not depend on the characteristics of the signal sampling, the strategy of following in the case that x is the realization of (not necessarily stationary) colored noise is simply the one in Sect. 2 i.e. transformation of x to an array y with uncorrelated entries. After that, the *Lomb-Scargle* periodogram can be computed. It is worth noticing that this simple result has been possible thanks to a formulation of the problem in the time domain and the use of the matrix notation. The same results could have been obtained by following the popular approach of working in the harmonic domain but at the price of a much more difficult derivation.

4. Two examples

To illustrate the usefulness and the simplicity of the proposed formalism in handling different situations from the classical ones, we show two examples in this section.

The first consists of a periodogram of a mean-subtracted time series. The evaluation of the reliability of a peak in the periodogram of a signal x requires that (under the *null* hypothesis $x = n$) n be the realization of a zero-mean noise process. In most experimental situations, this condition is not fulfilled and one works with a centered (i.e. mean-subtracted) version χ of x. This, however, introduces some (often neglected) problems. The

---

[6] The diagonal elements λ_1 and λ_2 of $\Sigma_\star$ can be trivially computed through the solution of the quadratic equation $\lambda^2 - \mathrm{tr}[\mathbf{C}_\star] + \det[\mathbf{C}_\star] = 0$, with tr[.] and det[.] denoting the *trace* and *determinant* operators. The arrays u_1 and u_2, which constitute the columns of $\mathbf{U}_*$, can be obtained by solving the equations $(\mathbf{C}_\star - \lambda_l \mathbf{I}) u_l = 0$, $l = 1, 2$.

case where x is the realization of a discrete white noise process with variance σ_x^2 has been considered several times in the literature. An example is the paper by Zechmeister & Kürster (2009) where a rather elaborate solution is presented. With the approach proposed here, a simpler solution can be obtained if one takes into consideration that the subtraction of the mean from x forces a spurious correlation among the entries of χ in such a way that the covariance matrix $\mathbf{C}_\chi = E[\chi \chi^{\mathrm{T}}]$ is given by

$$\mathbf{C}_\chi = \sigma_x^2 \left(\mathbf{I} - \frac{1}{M} \right), \tag{29}$$

where M is number of entries of x and 1 an $M \times M$ matrix with every entry equal to unity. Since this matrix is singular, it cannot be diagonalized and therefore χ cannot be whitened. In any case, if in Eq. (24) matrix C_x is substituted for $\mathbf{C}_\chi$ and one sets

$$\widehat{z} = \begin{pmatrix} \widehat{\chi}_{\mathfrak{R}} \\ \widehat{\chi}_{\mathcal{I}} \end{pmatrix}, \tag{30}$$

then it is a trivial matter to decorrelate $\widehat{z}_k$ and $\widehat{z}_{N_\dagger + k}$ by means of Eqs. (26), (27) and to compute the periodogram through Eq. (28). This result can be easily extend to the case where, because of measurement errors, each entry of x has its own variance $\sigma_{x_j}^2$ and a weighted mean is subtracted from the data sequence i.e. $\chi_j = x_j - \sum_l \eta_j x_l / \sum_l \eta_l$, with $\eta_l = 1/\sigma_{x_l}^2$. Indeed, it is sufficient to substitute $\mathbf{C}_\chi$ as given by Eq. (29) with

$$\mathbf{C}_\chi = \mathrm{diag}[\sigma^2] - \frac{1}{\sum_l \eta_l}, \tag{31}$$

where $\sigma^2 = [\sigma_{x_0}^2, \sigma_{x_1}^2, \ldots, \sigma_{x_{N-1}}^2]^{\mathrm{T}}$. The rest of the procedure remains unmodified.

The second example consists of zero-mean colored noise. The improvement in the quality of the results obtainable with the approach presented in the previous section is visible in Fig. 1. The top left panel shows a discrete signal $x_j = 0.5 \sin(2\pi f j) + n_j$, $f = 0.127$, simulated on a regular grid of 120 time instants but with missing data in the ranges [31 70] and [76 115]. Here, n is the realization of a discrete, zero-mean, colored noise process whose autocovariance function is given in the top right panel. From the bottom left panel, it is evident that *Lomb-Scargle* periodogram of the original sequence x provides rather ambiguous results concerning the presence of a sinusoidal component. On the other hand, such component is well visible in the bottom right panel that shows the *Lomb-Scargle* periodogram of the sequence $y = \mathbf{C}_n^{-1/2} x$.

5. Periodogram and least-squares fit of sinusoids

The formalism proposed here is also useful in the context of more theoretical questions (but with important practical implications). For example, a point often overlooked in the astronomical literature is the relationship between the periodogram and the least-squares fit of sine functions. Often these two methods are believed to be equivalent. Actually, this is true only when the sampling is regular and the frequencies of the sinusoids are given by the *Fourier* ones. Indeed, if $t_j = \tilde{t}_j$ and $f_k = k/N$, $k = 0, 1, \ldots, N - 1$, then Eq. (1) can be written in the form

$$x - \mathfrak{F}^{\mathrm{T}} a = n, \tag{32}$$

with

$$\mathfrak{F} = \frac{2}{\sqrt{N}} \begin{pmatrix} \mathcal{F}_{\mathfrak{R}} \\ \mathcal{F}_{\mathcal{I}} \end{pmatrix}, \tag{33}$$
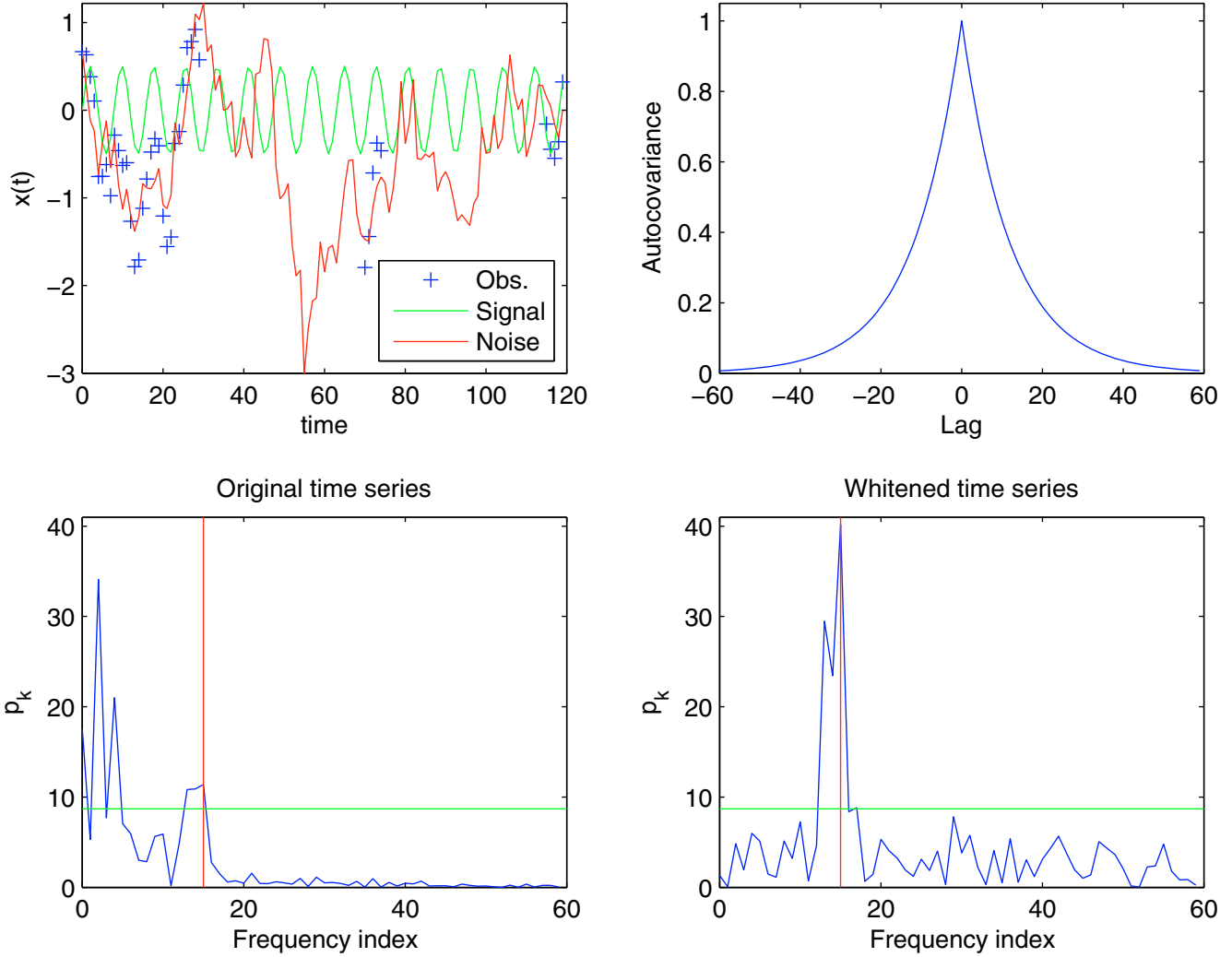
Fig. 1. Results concerning the numerical experiment, presented in Sect. 4, on the detection of a sinusoidal component in colored noise. *The top left panel* shows an observed time series (blue crosses) obtained through the simulation of signal $\boldsymbol{x} = \{x_j\}_{j=0}^{119}$ with $x_j = 0.5 \sin(2\pi f j) + n_j$, $f = 0.127$, on a regular grid of 120 time instants but with the data in the ranges [31 70] and [76 115]. Here, $\boldsymbol{n}$ (red line) is the realization of a discrete, zero-mean, colored noise process whose autocovariance function is given in *the top right panel*. For comparison, the sinusoidal component is also plotted (green line). *The bottom left panel* shows the *Lomb-Scargle* periodogram of the original sequence $\boldsymbol{x}$ computed on 120 frequencies $k = 0, 1/120, \ldots, 119/120$, whereas *the bottom right panel* shows the *Lomb-Scargle* periodogram corresponding to its whitened version. In both cases, only the first 60 frequencies are shown and the vertical red line corresponds to the frequency of the sinusoidal component. The horizontal green line in *the bottom panels* provides the threshold corresponding to a 0.01 *level of false alarm* (number of independent frequencies $N_f = 60$), i.e. the probability that the periodogram of a pure noise signal exceeds such a threshold by chance is 1%.

and $\boldsymbol{a} = [a_0, a_1, \ldots, a_{N-1}, b_0, b_1, \ldots, b_{N-1}]^T$. The least-squares solution $\bar{\boldsymbol{a}}$ of system (32) is given by

$$\bar{\boldsymbol{a}} = (\mathfrak{F}\mathfrak{F}^T)^+ \mathfrak{F} \boldsymbol{x}, \tag{34}$$

where "+" denotes *Moore-Penrose pseudo-inverse* (Björck 1996). In the case of even sampling, i.e. when $\mathcal{F}_\mathfrak{R} = \mathbf{F}_\mathfrak{R}$ and $\mathcal{F}_\mathcal{I} = \mathbf{F}_\mathcal{I}$, it happens that

$$\bar{\boldsymbol{a}} = \mathfrak{F} \boldsymbol{x}. \tag{35}$$

In other words, coefficients $\{a_k\}$ and $\{b_k\}$, as given by the least-squares approach, can be obtained through the DFT of $\boldsymbol{x}$, because, as shown by means of Eqs. (11), $(\mathfrak{F}\mathfrak{F}^T)^+ \mathfrak{F} = \mathfrak{F}$. In the case of uneven sampling, this identity is not fulfilled. Any kind

of periodogram computed through Eq. (22) and the least-squares fit of sine functions has to be expected to give different results. Moreover, as only under the two above-mentioned conditions do the sine functions constitute an orthonormal basis, the least-squares fit of a single sine function per time does not in general provide the same result as the simultaneous fit of all the sinusoids as in Eq. (34) (e.g. see Hamming 1973, p. 450). In particular, if an unevenly-sampled signal is given by the contribution of two or more sinusoids, the one-at-a-time fit of a single sine function provides biased results. This also holds for the *Lomb-Scargle* periodogram, which is equivalent to the least-squares fit of a single sinusoid with a specified frequency, with the constraint that the corresponding coefficients "a" and "b" are uncorrelated (Scargle 1982; Zechmeister & Kürster 2009).
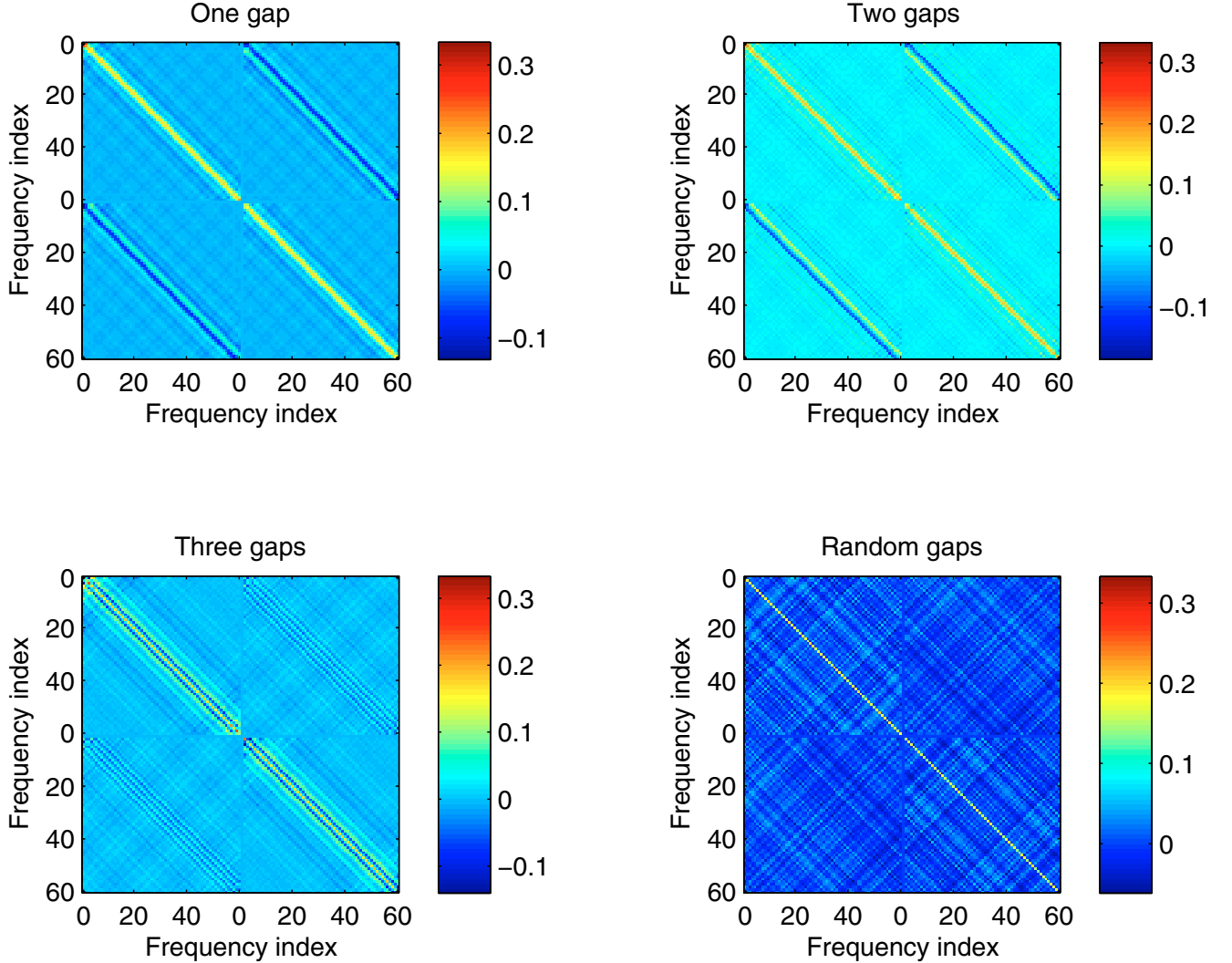
Fig. 2. Covariance matrix of the real and the imaginary parts, say $\Re[\widehat{x}_k]$ and $\mathcal{I}[\widehat{x}_k]$ of the Fourier transform $\{\widehat{x}_k\}$ of an unevenly sampled white-noise signal $\boldsymbol{x}$ (i.e. matrix $\mathbf{C}_{\widehat{z}}$ as given by Eq. (25)). In particular, $\boldsymbol{x}$ is assumed to be the realization of a zero-mean, unit-variance, white-noise process sampled at 120 time instants regularly spaced but with 80 missing data. Three different cases have been considered with missing data in the following ranges: **a)** [31 110] (*top left panel*), **b)** [31 70], and [76 115] (*top right panel*), **c)** [6 25], [36 75] and [96 115] (*bottom left panel*). In a fourth case, the missing data are randomly distributed (*bottom right panel*). Because of the small size of the figure, it is necessary to stress that for each panel, none of the prominent diagonal structures visible in *the top right*, as well in the bottom left quadrant, correspond to the main diagonal of the quadrant itself (i.e. none of them provide the covariance between $\Re[\widehat{x}_k]$ and $\mathcal{I}[\widehat{x}_k]$ that is seen in Fig. 3).

6. Discussion

As demonstrated in Sect. 3, when noise has arbitrary statistical characteristics, the computation of the periodogram of an unevenly-sampled signal requires two steps:

- whitening and standardization of the noise component (in this way a signal $\boldsymbol{y}$ is obtained);
- computation of the *Lomb-Scargle* periodogram of $\boldsymbol{y}$.

The first step, unavoidable even in the case of regular sampling, is a computationally expensive operation. Therefore, for time series containing more than a few thousand data points, dedicated algorithms exploiting the specific structure of $\mathbf{C}_n$ (e.g. often this matrix is of banded type) have to be developed for implementing Eq. (21). However, this problem is beyond the aim of the present paper. The second step is much less time consuming. Indeed, in the case of time series containing some thousands of points and when the periodogram has to be computed on a similar number of frequencies, the direct implementation of Eqs. (22)−(28) results in a few seconds of computation time only. In other words, in many practical situations, no dedicated algorithm is really necessary. However, fast algorithms have been proposed for very long time series (Press et al. 1992).

The last issue that has to be considered is to which extent the use of the *Lomb-Scargle* periodogram is really advantageous. Indeed, the action of the algorithms dealing with uneven sampling is essentially directed, for each frequency k, to force $\widehat{p}_k$ to be the sum of two independent Gaussian random quantities. However, although not clearly emphasized, it has already been pointed out that this operation is not critical (e.g. see Scargle 1982). It can be expected that a periodogram computed simply through

$$\widehat{p}_k = 2|\underline{\widehat{y}}_k|^2, \tag{36}$$
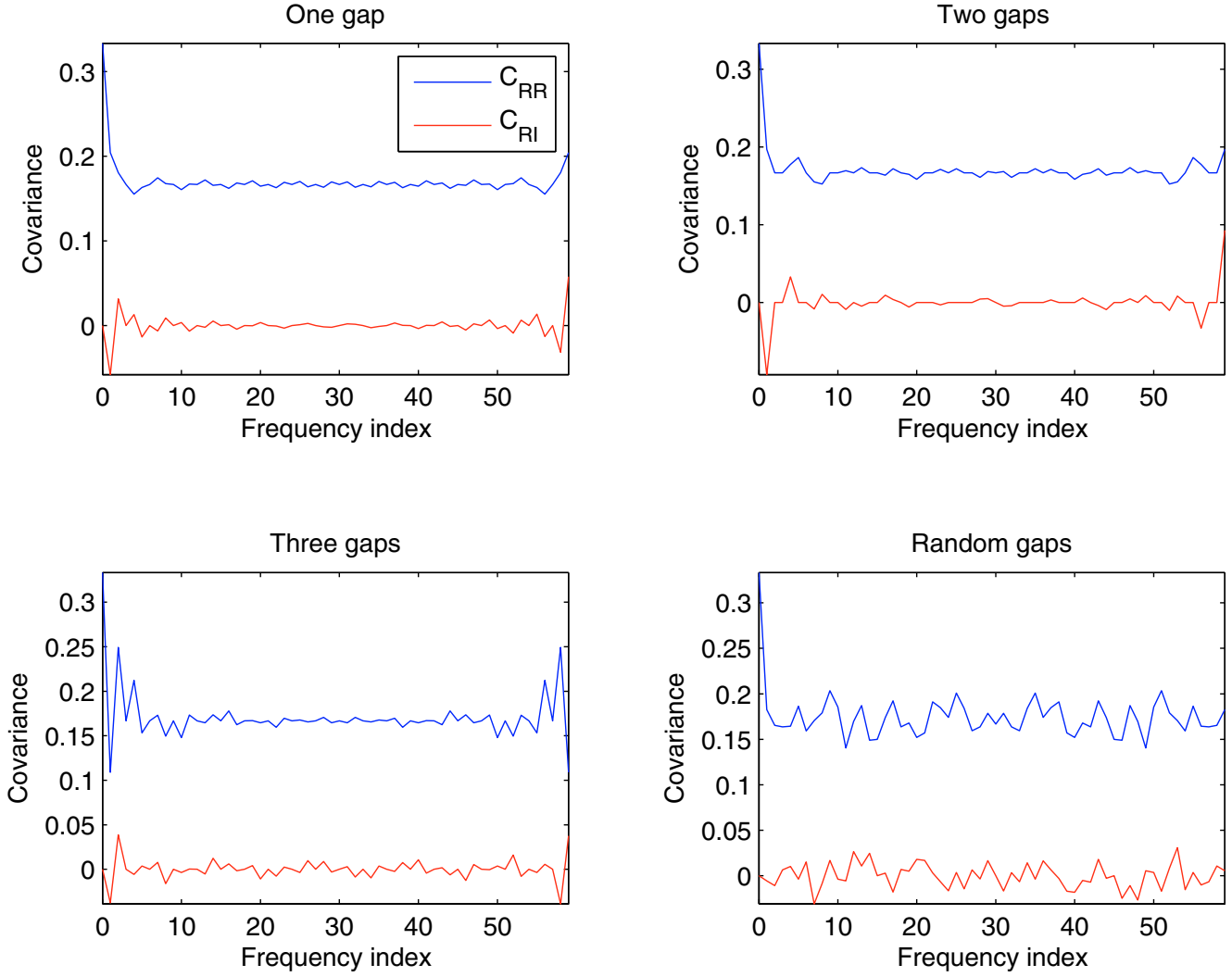
Fig. 3. Variance of $\Re[\widehat{x}_k]$ and covariance of $\Re[\widehat{x}_k]$ with $\mathcal{I}[\widehat{x}_k]$ as a function of the frequency k for the experiments in Fig. 2. The blue line corresponds to the main diagonal of the top left quadrant in each panel of Fig. 2, whereas the red one corresponds to the main diagonal in the top right, as well in the bottom left quadrant.

with $\widehat{y} = \mathcal{F}y$, is often very close to the one given by Eq. (28). A rigorous demonstration of this fact is difficult because of its strict dependence on the specific sampling. However, with the help of some numerical experiments and of the formalism introduced here, some insights are possible. In particular, we consider the covariance matrix $\mathbf{C}_{\widehat{z}}$ of a white noise signal when sampled on different uneven time grids.

6.1. Numerical simulations

In our simulations, we take the realization of a zero-mean, unit-variance, Gaussian white-noise process n sampled on $M_s = 120$ time regularly-spaced instants, but with 80 missing data (i.e. $M = 40$). The available signal x can be written in the form

$$x = \mathbf{W}n, \tag{37}$$
$$\mathbf{W} = \mathrm{diag}[w], \tag{38}$$

where $\mathbf{W} = \mathrm{diag}[w]$ and w is an array whose entries are equal to one in correspondence to a value of n that is available and zero

otherwise. In this case, Eq. (3) can be written as

$$\widehat{x} = \mathbf{FW}n \tag{39}$$
$$= \mathcal{F}n, \tag{40}$$

with $\mathcal{F} = \mathbf{FW}$. Three different cases have been considered where the missing data have time indices in the ranges a) [31 110], b) [31 70] and [76 115], c) [6 25], [36 75] and [96 115], whereas they are randomly distributed in a fourth case. The related covariance matrices $\mathbf{C}_{\widehat{z}}$, computed through Eq. (25) with $N = M_s$, are shown in Fig. 2, whereas Fig. 3 displays the corresponding main diagonal of the blocks $\mathcal{F}_\Re \mathcal{F}_\Re^\mathrm{T}$ and $\mathcal{F}_\Re \mathcal{F}_\mathcal{I}^\mathrm{T}$. Especially from Fig. 3 it is evident that, for an arbitrary frequency k, the covariance between $\Re[\widehat{x}_k]$ and $\mathcal{I}[\widehat{x}_k]$ is quite close to zero. This means that each entry of $\widehat{p}$, as given by Eq. (36), can be assumed to be distributed according to a χ_2^2. From Fig. 2 it is also evident that, in the case of gaps present in the sampling pattern, there are nearby frequencies k and l for which not only $\Re[\widehat{x}_k]$ and $\Re[\widehat{x}_l]$ are mutually correlated, but also $\Re[\widehat{x}_k]$ with $\mathcal{I}[\widehat{x}_l]$ and $\Re[\widehat{x}_l]$ with $\mathcal{I}[\widehat{x}_k]$. These correlations, especially those between the real and the imaginary components, disappear in the case of random sampling.
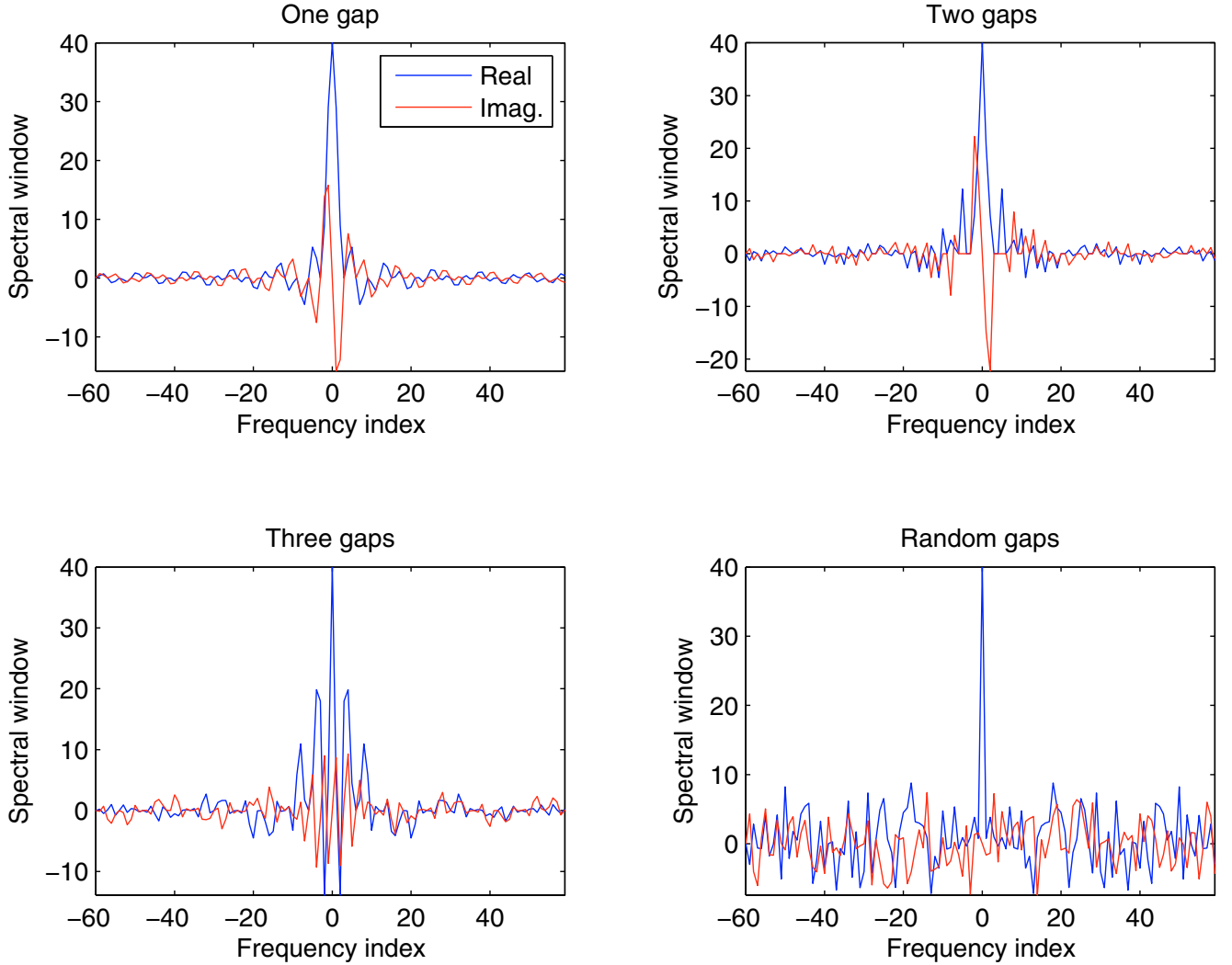
Fig. 4. Real (blue line) and imaginary (red line) parts of the spectral windows $\widehat{\boldsymbol{w}}$ that have been used to simulate the irregular sampling of the signal in the experiments corresponding to Fig. 2 (i.e., $\widehat{\boldsymbol{w}}$ is the Fourier transform of the sampling pattern $\boldsymbol{w}$ of $\boldsymbol{x}$, see Eqs. (37), (38)).

6.2. Interpretation of the results of the simulations

To understand these results, it is necessary to take into account that Eq. (39) can be rewritten in the form

$$\widehat{\boldsymbol{x}} = \mathbf{F}\mathbf{W}\mathbf{F}^*\mathbf{F}\boldsymbol{n}, \tag{41}$$
$$= \widehat{\mathbf{W}}\widehat{\boldsymbol{n}}. \tag{42}$$

As $\mathbf{W}$ is a (singular) diagonal matrix, $\widehat{\mathbf{W}}$ is a (singular) circulant matrix. This implies that $\widehat{\boldsymbol{x}}$ is given by the *circular convolution* of $\widehat{\boldsymbol{n}}$ (the DFT of the original signal, inclusive of the missing data) with the *spectral window* $\widehat{\boldsymbol{w}}$ (the DFT of the sampling pattern). Since for two arbitrary frequencies, say k and l, $\mathfrak{R}[\widehat{n_k}]$, $\mathcal{I}[\widehat{n_k}]$, $\mathfrak{R}[\widehat{n_l}]$, and $\mathcal{I}[\widehat{n_l}]$ are mutually independent, any correlation existing between $\mathfrak{R}[\widehat{x_k}]$, $\mathcal{I}[\widehat{x_k}]$, $\mathfrak{R}[\widehat{x_l}]$, and $\mathcal{I}[\widehat{x_l}]$ is induced by the correlation between the entries of $\mathfrak{R}[\widehat{w}]$ and $\mathcal{I}[\widehat{w}]$. Now, as shown in Fig. 4, in the presence of gaps the real and the imaginary parts of $\widehat{\boldsymbol{w}}$ are both only significantly different from zero in a narrow interval of frequencies surrounding the origin, i.e., corresponding to the lowest frequencies ($\boldsymbol{w}$ can be interpreted as low-pass filter). This produces the correlations observed among $\mathfrak{R}[\widehat{x_k}]$, $\mathfrak{R}[\widehat{x_l}]$, $\mathfrak{R}[\widehat{x_l}]$ and $\mathcal{I}[\widehat{x_k}]$ for nearby frequencies k and l. On the other hand, in the case of random sampling $\widehat{\boldsymbol{w}}$ mimics the behaviour of white noise, making the correlations less important.

Moreover, Fig. 5 shows that, in the case of sampling with gaps, the cross-correlation between $\mathfrak{R}[\widehat{w}]$ with $\mathcal{I}[\widehat{w}]$ is significant but not centered at zero lag. This explains why the quantities $\mathfrak{R}[\widehat{x_k}]$ and $\mathcal{I}[\widehat{x_l}]$, $k \neq l$ can be strongly correlated in spite of the small correlation between $\mathfrak{R}[\widehat{x_k}]$ and $\mathcal{I}[\widehat{x_k}]$. Similar arguments also hold for the case of sampling with periodic gaps (rather common in astronomical experiments). Indeed, in practical applications the period of these gaps is somewhat smaller than the mean sampling time step of an uninterrupted sequence of data. As a consequence, both $\mathfrak{R}[\widehat{w}]$ and $\mathcal{I}[\widehat{w}]$ present a sharp and narrow peak in correspondence to a frequency close to the origin. Actually, the *spectral window* of a sampling with periodic gaps is also characterized by the presence of *aliases*. However, these aliases too are sharp and narrow, and their importance decreases for increasing frequencies. The combination of these facts leads again to the quantities $\mathfrak{R}[\widehat{x_k}]$, $\mathfrak{R}[\widehat{x_l}]$, $\mathfrak{R}[\widehat{x_l}]$, and $\mathcal{I}[\widehat{x_k}]$ almost being uncorrelated if the frequencies k and l are not sufficiently close enough. Moreover, since with periodic gaps the cross-correlation between $\mathfrak{R}[\widehat{w}]$ with $\mathcal{I}[\widehat{w}]$ can also be significant, but not centered on zero lag, then for each frequency k the correlation between $\mathfrak{R}[\widehat{x_k}]$ and $\mathcal{I}[\widehat{x_k}]$ is negligible. These considerations are confirmed by Figs. 6 and 7. Figure 6 displays the covariance matrix $\mathbf{C}_{\widetilde{z}}$, computed through Eq. (25), of a discrete zero-mean,
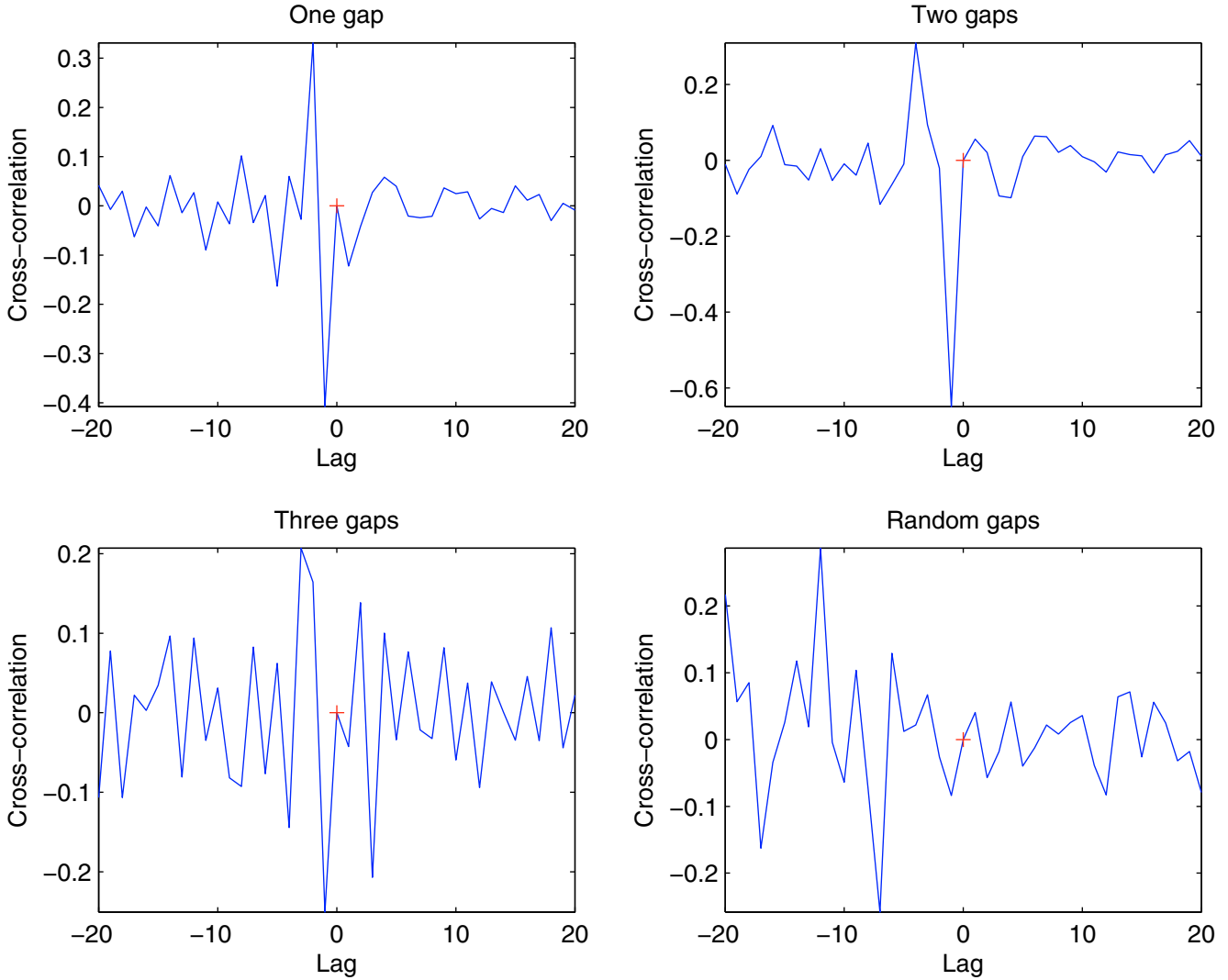
Fig. 5. Cross-correlation between the real and the imaginary parts of the spectral windows $\widehat{w}$ that are shown in Fig. 4. The red cross in each panel corresponds to the point $(0, 0)$.

unit-variance, and white-noise process sampled on 600 time instants randomly distributed (i.e. not rebinned) along six cycles (i.e. 100 points per cycle) of the sampling pattern shown in the bottom panel of Fig. 7; 1200 frequencies have been considered. The top panel of Fig. 7 displays the main diagonal of the blocks $\underline{\mathcal{F}}_{\mathfrak{R}}\underline{\mathcal{F}}_{\mathfrak{R}}^{\mathrm{T}}$ and $\underline{\mathcal{F}}_{\mathfrak{R}}\underline{\mathcal{F}}_{I}^{\mathrm{T}}$.

The numerical simulations indicate that the consequences of unevenly-sampled data seem to concern the number of independent frequencies in $\widehat{p}$ rather than the correlation between $\mathfrak{R}[\widehat{x}_k]$ and its imaginary counterpart $I[\widehat{x}_k]$. At present, no general method has been developed to deal with this problem. However, it has been pointed out in the literature that the number of independent frequencies is not a critical parameter to test the significance level of a peak in $\widehat{p}$. In particular, empirical arguments indicate that this number can be safely set to $M/2$ (e.g. see Press et al. 1992). The conclusion is that forcing each entry of $\widehat{p}$ to be the sum of two independent Gaussian quantities only has minor effects. This is shown by Fig. 8 where the top panel shows that the *Lomb-Scargle* periodogram of the time series with periodic gaps considered above is quite similar to that provided simply by $\widehat{p}_k = 2|\widehat{x}_k|^2$ with $\underline{\widehat{x}} = \underline{\mathcal{F}}\underline{x}$. This is evident in the bottom panel

of the same figure where the similarity of the two periodiograms is demonstrated by their absolute difference.

A final point to underline, which has important practical implications, is that for long time series the small differences visible in Fig. 8 should decrease. Indeed, as seen above, $\widehat{p}_k$ will be significantly correlated with $\widehat{p}_l$ only if the two frequencies k and l are close enough. The only exception is represented by the frequencies at the extremes of the frequency domain where the assumption of periodic signal intrinsic to DFT forces a spurious correlation. For longer and longer time series, this spurious correlation will affect a smaller and smaller fraction of frequencies and, as consequence, a larger and larger fraction of them will be mutually independent. This is shown in Fig. 9 where, in the context of the previous experiment, the mean absolute difference between the two periodograms is plotted as a function of N_{s}, the number of cyclic sampling patterns ($N_{\mathrm{s}} = 6$ in Fig. 8). This argument also explains why in many practical situations the number of independent frequencies can be safely fixed to $M/2$. In conclusion, only in the case of signals that contain a small number of data, the *Lomb-Scargle* periodogram can be expected to exhibit noticeable differences from the periodogram given by
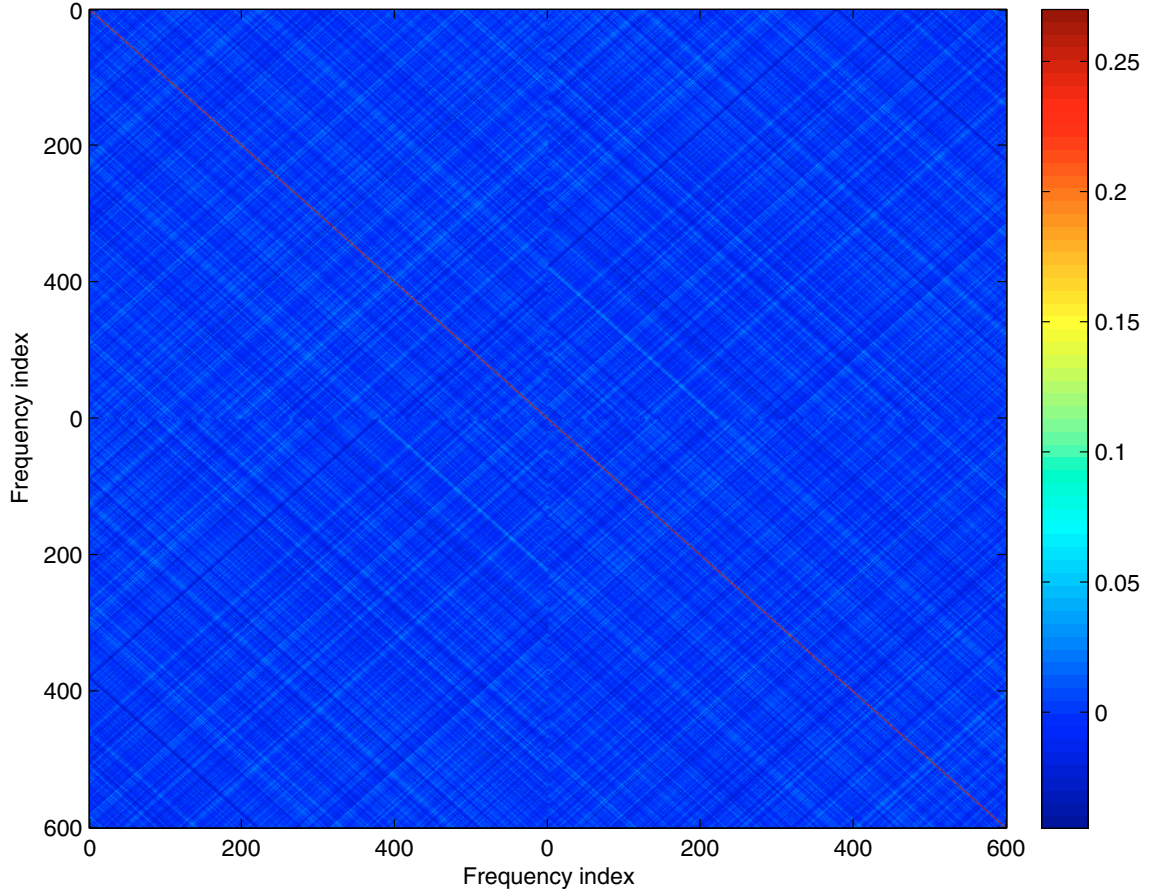
Fig. 6. Covariance matrix of the real and the imaginary parts, say $\Re[\widehat{x}_k]$ and $\mathcal{I}[\widehat{x}_k]$ of the Fourier transform $\{\widehat{x}_k\}$ of an unevenly sampled white-noise signal x with periodic gaps (i.e. matrix $\mathbf{C}_{\widehat{z}}$ as given by Eq. (25)). In particular, x is the realization of a discrete zero-mean, unit-variance, white-noise process sampled on a grid of 600 time instants randomly distributed (i.e. not rebinned) according to the cyclic sampling pattern shown in *the bottom panel* of Fig. 7. In total, 1200 frequencies have been considered.
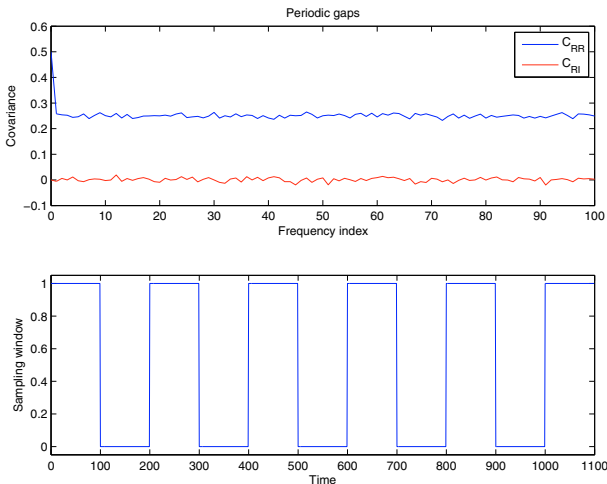

Fig. 7. *Top panel*: variance of $\Re[\widehat{x}_k]$ and covariance of $\Re[\widehat{x}_k]$ with $\mathcal{I}[\widehat{x}_k]$ (only the first 100 frequencies are shown) for the experiment concerning the realization of a discrete zero-mean, unit-variance, white-noise process on 600 time instants randomly distributed (i.e. not rebinned) according to a cyclic sampling. In total, 1200 frequencies have been considered (see text in Sect. 6). *Bottom panel*: cyclic sampling used in the experiment.

Eq. (36). Often, using it does not change anything. Comparable results can be expected with less sophisticated approaches.

### 6.3. Application to an astronomical time series

As an example of an unsophisticated method able to produce results similar to those obtainable with the *Lomb-Scargle* periodogram, we consider the rebinning of the original time series on an arbitrarily dense regular time grid (in this way a signal with a regular sampling is obtained but some data are missing) followed by applying any of the *fast Fourier transform* (FFT) algorithms available nowadays. Figure 10 shows an experimental (mean-subtracted) time series versus its rebinned version. This time series, which is characterized by rather irregular sampling, was obtained with the VLA array (Biggs et al. 1999) and consists of polarisation position angle measurements at an observing frequency of 15 GHz for one of the images of the double gravitational lens system B0218+357. The original sequence contains only $M = 45$ data and it is rebinned on a regular grid of $M_r = 92$ time instants. In spite of this, as visible in Fig. 11, the corresponding periodograms, computed on $N = M_r$ equispaced frequencies by means of the *Lomb-Scargle* and the FFT approach, are remarkably similar. Here, the highest frequency approximately corresponds to the Nyquist frequency that is related to the shortest sampling time step. The main conclusion of this example is to point out that, although in the previous section we stated that use of the *Lomb-Scargle* periodogram can be expected to be effective only for time series that contain a small number of data, this is not a sufficient condition to guarantee that the method is truly useful.
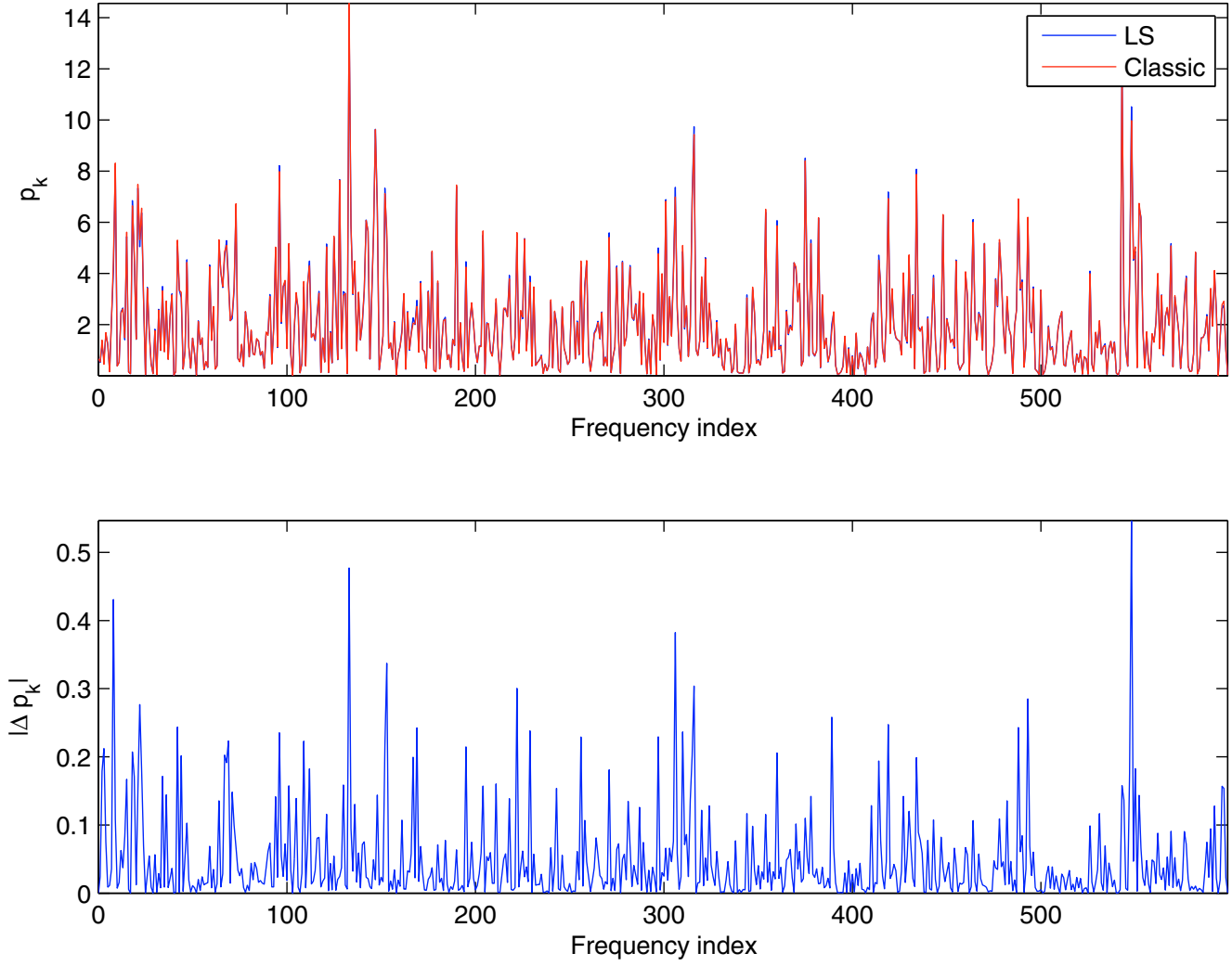
Fig. 8. *The top panel* shows the *Lomb-Scargle* periodogram and the classic periodogram given by $\widehat{p}_k = 2|\widehat{x}_k|^2$ for the experiment in Fig. 7. *The bottom panel* shows the smallness of their absolute difference.
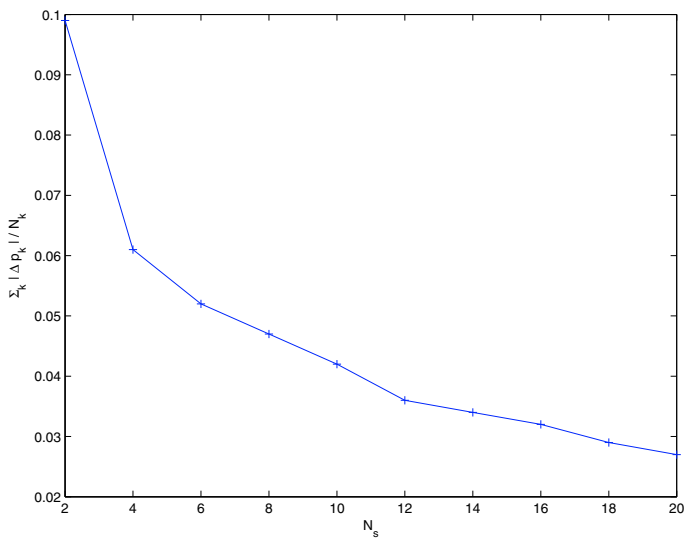
Fig. 9. Mean absolute difference $\sum_k |\Delta \widehat{p}_k|/N_k$ (N_k = number of frequencies) between the *Lomb-Scargle* periodogram and the classic periodogram given by $\widehat{p}_k = 2|\widehat{x}_k|^2$ as a function of the number N_s of cyclic sampling patterns (Fig. 8 shows the case with $N_s = 6$).
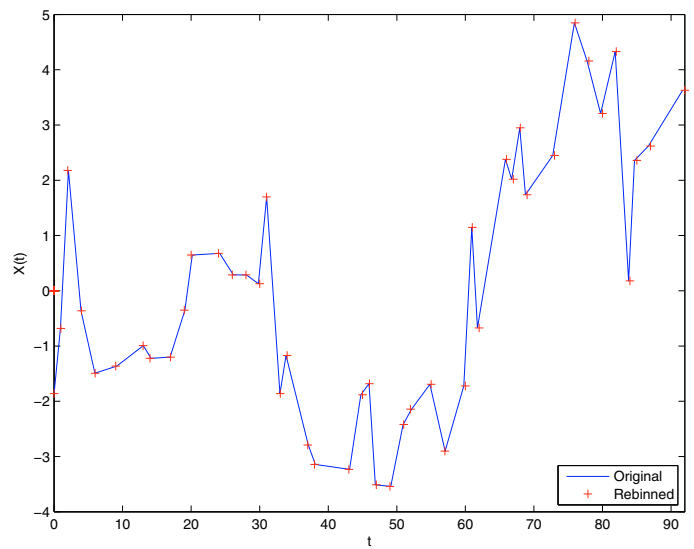
Fig. 10. Experimental (mean-subtracted) time series containing 45 unevenly-spaced data versus a rebinned version computed on a regular grid of 92 time instants (data taken from Biggs et al. 1999).
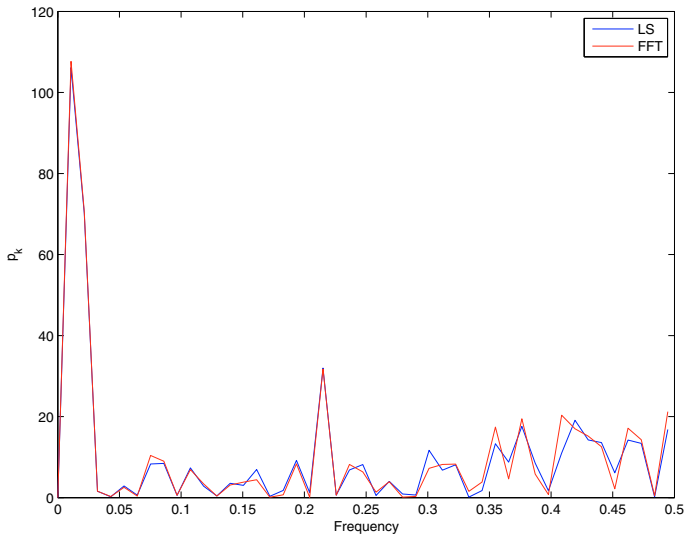
Fig. 11. Periodograms of the time series shown in Fig. 10. Frequency is in given in units of the Nyquist frequency corresponding to the shortest sampling time step. The periodogram of the original time series has been obtained by means of the *Lomb-Scargle* method and that of the rebinned version by means of a classic FFT algorithm.

7. Conclusions

In this paper we worked out a general formalism, based on the matrix algebra, that is tailored to analysis of the statistical properties of the *Lomb-Scargle* periodogram independently of the characteristics of the noise and the sampling. With this formalism it has become possible to develop a test for the presence of components of interest in a signal in more general situations than those considered in the current literature (e.g. when noise is colored and/or non-stationary). Moreover, we were able to clarify the relationship between the *Lomb-Scargle* periodogram and other techniques (e.g. the least-squares fit of sinusoids) and to fix the conditions under which the use of such method can be expected to be effective.

References

Biggs, A. D., Browne, I. W. A., Helbig, P., et al. 1999, MNRAS, 304, 349
Björck, A. 1996, Numerical Methods for Least Squares Problems (Philadelphia: SIAM)
Ferraz-Mello, S. 1981, AJ, 86, 619
Gilliland, R. L., & Baliunas, S. L. 1987, ApJ, 314, 766
Koen, C. 2006, MNRAS, 37, 1390
Hamming, R. W. 1973, Numerical Methods For Scientists and Engineering (New York: Dover)
Lomb, N. R. 1976, Ap&SS, 39, 447
Oppenhaim, A. V., & Shafer, R. W. 1989, Discrete-time Signal Processing (London: Prentice Hall)
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes (Cambridge: Cambridge University Press)
Reegen, P. 2007, A&A, 467, 1353
Scargle, J. D. 1982, ApJ, 263, 835
Simon, M. K. 2006, Probability Distributions Involving Gaussian Random Variables (Heidelberg: Springer)
Vio, R., Strohmer, T., & Wamsteker, W. 2000, PASP, 112, 74
Zechmeister, M., & Kürster, M. 2009, A&A, 496, 577