A spectral estimation approach to contrast transfer function detection in electron microscopy

José-Jesús Fernández, José R. Sanjurjo, José-Maria Carazo

Abstract

In this work we approach the task of estimating the contrast transfer function (CTF) of a transmission electron microscope by applying mathematical tools extracted from the field of spectral estimation in multidimensional signal processing, such as periodogram averaging and autoregressive modelling. We prove that the clarity and precision by which the CTF can be detected using these approaches is far better than by any conventional method based on the Fourier transform amplitude alone. We also present a unified signal processing-based framework in which recent development in CTF detection approach can be studied, helping to understand their respective benefits and problems.

PACS: 82.80.Ch; 07.05.Pj; 07.05.Kf; 42.30.Lr

Keywords: Data processing/image processing; Contrast transfer theory

1. Introduction

Central to the task of extracting useful information from an electron microscopy picture rests the ability to relate that image to properties of the object being imaged. In general, transmission electron microscopy images are affected by many forms of aberrations. Among them, the most important ones are those introduced by the objective lenses. The image degradation produced by the lens aberrations and defocusing is known in the field as the “contrast transfer function (CTF)”. It is critical for many applications to have a good knowledge of the actual shape of the CTF that is affecting the image.

In general, quite in-depth studies on the theoretical and experimental shape of the CTF have been carried out [1–9], leading to an approximate parametric model in which, for a given microscope setting, the most
important variables are the defocus, astigmatism as well as the percentage of phase versus amplitude contrast (ignoring some of the high-resolution attenuation effects described in some of the references given above).

A quite common way of processing an electron micrograph in order to estimate the CTF of the microscope is to calculate the Fourier transform of the image and, from a display of its amplitude, perform a fitting of the parametric model of the CTF [10, 11]. Since the CTF is a function with many zeroes, most of the attention is given to the location of these zeros as landmarks of the whole parametric fitting.

In the past, the procedure described above has been the single way to estimate the CTF of the microscope. Recently, however, Avila-Sakar et al. [12], Zhu and Frank [13], and Zhou et al. [14, 15] have presented a modification of this approach leading to much clearer images in which the zeroes of the CTF were substantially easier to identify.

In this work we have studied the mathematical reasons behind the success of those previous works, approaching the whole issue of CTF detection within the framework of spectral estimation in multidimensional signal processing. In this way we have been able to formulate the above referred methods as a special case of spectral estimation (a modification of the previously proposed procedures had to be introduced in order to gain mathematical consistency), studying their relative merits compared to the simple traditional approach described before, as well as providing guidelines on its use. Additionally, we have investigated a totally different method for spectral estimation that, in same cases, can produce even better results than those improved methods previously mentioned.

2. Contrast transfer function theory in transmission electron microscopy

Studies over many years in bright field electron microscopy have led to the proposal of a general model of image formation in transmission electron microscopy (TEM), expressing the Fourier relationship between the image and the object projection [13] as:

\[ G(\rho) = H(\rho)F(\rho) + N(\rho), \]

where \( G(\rho) \) is the Fourier transform of the recorded image, \( F(\rho) \) is the Fourier transform of the unknown object projection and \( N(\rho) \) is the Fourier transform of the noise.

The term \( H(\rho) \) is the contrast transfer function (CTF) of the transmission electron microscope that represents the distortions introduced by the microscope imaging system at each spatial frequency \( \rho \).

In the wave theory of image formation, the effects of lens aberrations and defocusing are attributed to a phase shift \( \chi(\rho) \) given as a function of the spatial frequency by

\[ \chi(\rho) = \pi \lambda (\Delta f \rho^2 - \frac{1}{2} C_s \rho^4 \lambda^2), \]

where \( C_s \) is the coefficient of spherical aberration, \( \Delta f \) is the defocusing and \( \lambda \) is the electron wavelength.

The CTF term \( H(\rho) \), reflecting the contributions of both phase and amplitude contrast transfer functions, is given by Erickson and Klug [1] as

\[ H(\rho) = - (\sin \chi(\rho) + Q(\rho) \cos(\rho)). \]

Eq. (2) and (3) constitute the mathematical model of the CTF of the microscope. These formulas depend upon the coefficient of spherical aberration \( C_s \) of the microscope (given by the manufacturer), the electron wavelength \( \lambda \) (derived from the used electron accelerating voltage) and finally, the defocusing. As the spherical aberration coefficient and the electron wavelength are given parameters, the CTF model of the microscope, according to Eq. (2) and (3), is completely specified by the defocusing \( \Delta f \) and the amplitude contrast proportion \( Q(\rho) \).

The simple formula (1) is valid only if completely coherent illumination is assumed. However, in practice, the effects due to partially coherent illumination [4] and chromatic defocus spread [5] have to be taken into
account too. A good model that accommodates all these issues can be achieved by representing the influence of those effects as a set of envelope functions that damp the CTF down, attenuating the high resolution region. The modulation transfer function of the electron micrographs [7] and other resolution limiting effects [8] have been also characterized by means of some kind of envelope functions.

Zhu and Frank [13] have carried out a good compendium of all the approximate envelope representations that have just been mentioned in the previous paragraphs. Also, the chapter devoted to Principle of Image Formation in the Electron Microscope in the book of Frank [16] presents an excellent review of many aspects in this subject.

In the electron microscopy field, the biological specimens to be studied are often supported on amorphous carbon films. To a good approximation, we can consider the final image (and therefore its spectrum) to be the addition of the specimen and the carbon film projections. We certainly do not know a priori the spectrum of the specimen, but the fact that the spectrum of the amorphous carbon film is approximately 'white' [9] allow us to start studying the CTF directly from the spectrum of the recorded image, \( G(\rho) \) in Eq. (1).

3. CTF determination

The effect of the CTF introduces typical modifications on the recorded images that are easily detected in the Fourier transform of a TEM image. In particular, the CTF possesses zones with alternating positive and negative signs, visible as bright circular (or elliptical, if there is some astigmatism) bands in the Fourier transform of the image. The zero transitions between the areas with opposite sign are visible as dark rings [3]. Those rings constitute the most obvious features that the CTF imprints on the Fourier transform of the recorded image.

As mentioned in the previous section, the CTF model of the microscope is specified by the defocusing and the proportion of amplitude contrast. The main effect of defocusing on the CTF is to change the position at which the first peak and subsequent oscillations occur. Thus, the way to estimate the defocus degree is from the positions of the rings that the CTF stamps on the Fourier transform of the TEM image. The method used to derive the defocus value from the position of a CTF zero or peak is the one implemented in the ICE program [15, 17], a well-known image processing package within the context of electron crystallography. Algorithms I and II, presented in Appendix A, thoroughly show this method. In order to compute the most reliable defocusing estimate, both Algorithms I and II may be applied and the resulting defocus degrees are then used together in a subsequent linear least-squares fitting. The reliability of the determined defocus value is judged from the consistency of the defocus values derived from different orders of CTF zeroes or peaks with each other.

With regard to the amplitude contrast fraction, its effect is more important at low spatial frequencies and consists of shifting the low-resolution CTF zeroes and peaks towards the origin of the Fourier space. For negative staining, amplitude and phase contrast usually complement each other, giving a roughly constant transfer function up to a resolution of about 2 nm [1], and to this resolution, little correction for the shape of CTFs is needed. For cryo-electron crystallography, usually only a small proportion of amplitude contrast is considered. It is for that reason that the method to derive defocus value from positions of CTF zeroes or peaks does not take into account the amplitude contrast.

As a consequence of the reasoning presented in the second paragraph of this section, the key to the estimation of the CTF is a proper location of its zeroes. However, in practice it is not so easy to see the CTF rings on the Fourier transform of the recorded image. The visibility of the CTF rings depends upon a great number of factors, including the damping factors mentioned above and others. The work of Zhou et al. [15] presents a summary of all these factors, encompassing electron statistics of the image [18], defocus value [3], specimen drift, partial coherence [4-6], instrumental instabilities, film MTF [7], structure factors of specimen, and ice thickness. As a consequence of all these factors, the Fourier transform of the images rarely
presents clear and well-defined CTF rings. This especially occurs with images coming from electron cryo-microscopy, due to their poor electron statistics and weak scattering factors.

Several automated computer-fitting approaches have been used to determine the CTF affecting an image from an inspection of the amplitude of its Fourier transform. They consist of a process of matching a theoretical CTF amplitude with the amplitude of the experimental Fourier transform by using an iterative non-linear least-squares fitting method [10].

Henderson et al. [11] used a simple least-squares refinement of defocusing and astigmatism parameters by minimizing an error function. This error function represented the global difference between the Fourier transform amplitudes and the amplitudes measured from electron diffraction pattern modulated by the theoretical CTF.

In practice, the error functions that are used have many local minima, so the only way to guarantee the global minimum to be reached is by providing a good estimation of the initial defocusing and astigmatism parameters to be used in the subsequent steps of the refinement process.

Also, there are other works in which, from the initial approximate values of defocusing and astigmatism estimated from the CTF rings, a refinement procedure is performed by searching for the best phase agreement with merged phases coming from other crystalline areas [19–21].

Therefore, it can be concluded that it is necessary to have available good initial estimates of defocusing and astigmatism parameters to be used as input in any of the CTF refinement methods. This involves to have a good Fourier transform on which, at least, one or a few CTF rings can be visualized. Within this context, our work precisely focuses on the subject of emphasizing the Thon rings of an experimental Fourier transform.

Historically, there have been several approaches to make the CTF zeroes be more visible. The works of Unwin and Henderson [22] and Amos et al. [23] described a method consisting of recording a high dose image immediately after the low dose exposure image. The low dose image was the one used to analyse the structure. The high dose image was the one which revealed the Thon rings and, therefore, it was used to estimate the CTF affecting the low dose image.

Lately, a new approach has arisen in the context of electron microscopy, consisting of dividing the original image into regions, and computing the average of the Fourier transform amplitude of those regions [12, 14, 15]. The result is a smoother Fourier transform amplitude making the detection of the Thon rings much easier. Zhu and Frank [13] apply a similar method, but computing the power spectrum of the regions into which the original image has been divided. In the forthcoming sections, the mathematical reasons behind the success of those works will be studied.

Finally, we have to comment another method used in this field to improve the location of the CTF zeroes. It computes the radial averaged Fourier transform amplitude of the experimental image [6, 13, 15], resulting in a smooth 1D graph of the CTF that is affecting the image. Obviously, this method is valid only for non-astigmatic or slightly astigmatic images, or in cases where the interest is centred on increasing the accuracy in other parameters, such as the various envelope functions, more than defocusing and astigmatism.

4. Stochastic processes

Let \( x \) denote a two-dimensional stationary real discrete stochastic (or random) process [24] formed by the ensemble of random variables \( \{x_{n_1,n_2}\} \). Each random variable is associated with a point of discrete space and has its own probability distribution function [25]. Stochastic processes, in contrast to deterministic ones, can only be described by statistical laws or models. In practice, the process is characterized by its autocorrelation function (ACF), commonly denoted by \( r_{xx} \), and it is defined, for a stationary real random process, by [25, 26]

\[
r_{xx}(m_1, m_2) = \mathbb{E}(x_{n_1+m_1,n_2+m_2} x_{n_1,n_2})
\]

independent of \( (n_1, n_2) \), where \( \mathbb{E} \) denotes mathematical expectation.
In order to avoid dealing with an infinite ensemble of realizations of the same stochastic process (from now on, called 2D sequences or, simply, sequences), the process is considered to be ergodic [24–27]. This means that the process can be studied from only one sequence \( \{x[n_1, n_2]\} \) by obtaining the ACF in the following way:

\[
\hat{r}_{xx}[m_1, m_2] = \lim_{L \to \infty} \frac{1}{(2L + 1)^2} \sum_{n_1 = -L}^{L} \sum_{n_2 = -L}^{L} x[n_1 + m_1, n_2 + m_2]x[n_1, n_2].
\] (4)

Obviously, the value at each point of the ACF function defined by Eq. (4) cannot be obtained in practice on a digital computer since they are only defined at the limit of \( L \to \infty \). It is for this reason that an estimate of it, obtained only from a finite sequence of the process \( \{x[n_1, n_2]\} \) for \( 0 \leq n_1 < N_1 \) and \( 0 \leq n_2 < N_2 \) has to be used. An estimate of the ACF, denoted by \( \hat{r}_{xx} \), is given by [26–28]

\[
\hat{r}_{xx}[m_1, m_2] = \begin{cases} 
\frac{1}{N_1 N_2} \sum_{n_1 = 0}^{N_1 - 1} \sum_{n_2 = 0}^{N_2 - 1} x[n_1 + m_1, n_2 + m_2]x[n_1, n_2] & \text{for } m_1 \geq 0, m_2 \geq 0, \\
\frac{1}{N_1 N_2} \sum_{n_1 = 0}^{N_1 - 1} \sum_{n_2 = -m_2}^{N_2 - 1} x[n_1 + m_1, n_2 + m_2]x[n_1, n_2] & \text{for } m_1 > 0, m_2 < 0.
\end{cases}
\] (5)

This quantity will be referred to as the biased ACF. As the matrix \( \hat{r}_{xx} \) is Hermitian, the remaining ACF estimates for \( m_1 < 0, m_2 \geq 0 \) and \( m_1 \leq 0, m_2 < 0 \) can be obtained in the following way:

\[
\hat{r}_{xx}[-m_1, -m_2] = \hat{r}_{xx}[m_1, m_2].
\]

4.1. Power spectrum

In the signal processing field, power spectra are associated with Fourier transforms. For deterministic signals, they are used to represent a function as a superposition of complex exponentials. For random signals, the notion of power spectrum has two interpretations. The first one involves transforms of averages, in concrete the Fourier transform of the ACF; it is thus essentially deterministic [24]. The second one leads to the representation of the process under consideration as superposition of complex exponentials with random coefficients. In this work we are going to use the first interpretation. Within this framework, the Fourier transform of the ACF of the stochastic (or random) process, denoted by \( P_{xx} \), is given by

\[
P_{xx}(\omega_1, \omega_2) = \sum_{m_1 = -\infty}^{\infty} \sum_{m_2 = -\infty}^{\infty} \hat{r}_{xx}[m_1, m_2]e^{-i\omega_1 m_1}e^{-i\omega_2 m_2},
\] (6)

defined over the range \(-\pi \leq \omega_1, \omega_2 \leq \pi \).

The equivalent description of a stochastic process by means of the Fourier transform of the ACF, \( P_{xx}(\omega_1, \omega_2) \), provides the distribution of the average power of the process as a function of the frequency. In fact, the integral of \( P_{xx}(\omega_1, \omega_2) \) over a band of frequencies is proportional to the power in the signal in that band. For this reason, the function \( P_{xx}(\omega_1, \omega_2) \) is called the power spectral density (PSD), or simply the power spectrum [27]. The relationship between the PSD and the Fourier transform of the ACF, as expressed by Eq. (6), is sometimes referred to as the Wiener–Khinchin theorem [28–30].

Although this interpretation may not be evident from Eq. (6), an alternative but equivalent definition of the PSD illustrates in a more clear way this property. The PSD may also be defined as [29]

\[
P_{xx}(\omega_1, \omega_2) = \lim_{L \to \infty} \mathbb{E} \left( \frac{1}{(2L + 1)^2} \sum_{n_1 = -L}^{L} \sum_{n_2 = -L}^{L} x[n_1, n_2]e^{-i\omega_1 n_1}e^{-i\omega_2 n_2} \right),
\] (7)

over the range \(-\pi \leq \omega_1, \omega_2 \leq \pi \).
Eq. (7) says that the PSD at frequency \((\omega_1, \omega_2)\) is found by first taking the amplitude squared of the Fourier transform of the sample sequence \(\{x[n_1, n_2]\}\) and then dividing by the sequence length to yield power. The expectation operator is required since the ergodicity of the process does not couple through the Fourier transform [30]. Finally, since the process is in general of infinite length, a limit operation is required.

In the study of a finite length sample sequence, the PSD, just like the ACF, cannot be computed from Eqs. (6) and (7) due to the limits. Then, an estimate is necessary to be used. Our study precisely focuses on the problem of spectral estimation, and will be thoroughly dealt with in next sections.

In the analysis of finite length sample sequences, the PSD is often preferable to the ACF because estimates of the PSD at neighbouring frequencies are approximately independent, and hence the interpretation is usually easier than that of the biased ACF.

5. **Spectral estimation**

As it has been enunciated in Section 4.1, the problem of spectral estimation is that of determining the distribution in frequency of the power of a stochastic process. Estimation of the PSD is usually complicated by the lack of a sufficiently long data record on which to base the spectral estimate (it is obvious that in order to compute reliable spectral estimates we would like to have large amounts of data).

Tradeoffs can be expected in spectral estimation; the paramount one is bias versus variance. As it will be seen, if the spectral estimator yields good estimates on the average (low bias), then we can expect much variability from one data realization to the next (high variance). On the other hand, if we choose a spectral estimator with low variability, then on the average the spectral estimate may be poor. The only way out of this dilemma is to increase the data record length.

We can divide spectral estimation techniques into parametric or non-parametric methods, but this classification is not straightforward. The philosophy that we will adopt in this work is that a parametric technique is one that assumes more a priori knowledge about the PSD other than the validity of the Wiener–Khinchin theorem (Eq. (6)). With this definition the Fourier-based methods (in the following, they will be termed as classical spectral estimation methods) are non-parametric spectral estimators, as is the minimum variance spectral estimator, while the autoregressive, moving average, and autoregressive moving average approaches are parametric. Non-parametric spectral estimators require no assumptions about the data other than wide-sense stationarity. The advantage of a parametric spectral estimator is that, when applicable, it yields a more accurate spectral estimate. In this way the bias and the variance can simultaneously be reduced over the non-parametric spectral estimator without having to increase the data record length.

5.1. **Classical spectral estimation**

Two families of spectral estimation techniques based on Fourier transform operations have evolved. This dichotomy is the result of the two equivalent definitions of PSD given by Eqs. (6) and (7). The set of techniques making use of Wiener–Khinchin theorem, Eq. (6), is known as the Blackman–Tukey approach, whereas within the ones relying on Eq. (7), the periodogram and the periodogram-averaging estimators are found.

5.1.1. **Periodogram spectral estimator**

The periodogram spectral estimator relies on Eq. (7). Let \(x\) be the two-dimensional ergodic discrete stochastic process under consideration. Then, by neglecting the expectation operator in Eq. (7), and using the
available sequence of the process \( \{x[n_1, n_2]\} \) for \( 0 \leq n_1 < N_1 \) and \( 0 \leq n_2 < N_2 \), the periodogram is defined as

\[
\hat{P}_{xx}(\omega_1, \omega_2) = \frac{1}{N_1 N_2} \left| \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x[n_1, n_2] e^{-j\omega_1 n_1} e^{-j\omega_2 n_2} \right|^2
\]

over the range \( -\pi \leq \omega_1, \omega_2 \leq \pi \).

Let \( X(\omega_1, \omega_2) \) be the Fourier transform of the sequence \( \{x[n_1, n_2]\} \), then

\[
\hat{P}_{xx}(\omega_1, \omega_2) = \frac{1}{N_1 N_2} |X(\omega_1, \omega_2)|^2.
\]  

Many studies [25, 27, 29] have shown that the periodogram is an asymptotically unbiased estimator of the PSD, but it is not consistent. This means that the bias, i.e. the difference between the expected value of the periodogram and the true PSD, approaches zero as the length of the sequence increases, but the variance does not. The standard deviation, which is the square root of the variance, is as large as the mean of the PSD that we are estimating, independently of the sequence length:

\[
\text{VAR}[\hat{P}_{xx}(\omega_1, \omega_2)] \approx P^2_{xx}(\omega_1, \omega_2).
\]  

It is for that reason that the periodogram is an unreliable estimator.

5.1.2. Periodogram-averaging spectral estimator

The reason whereby the periodogram estimates are statistically inconsistent (unstable) is that no statistical averaging is performed, i.e. the expectation operation in Eq. (7) has been neglected.

The way out to overcome this limitation is to use an averaged periodogram estimator in an attempt to approximate the expectation operation in Eq. (7). In practice, only one data record on which to base the spectral estimator is available. It is for that reason that the ergodicity property of the process is always assumed. Then, the approach consists of segmenting the available sequence \( \{x[n_1, n_2]\} \) into \( K_1 \times K_2 \) non-overlapping sections of length \( L_1 \times L_2 \), where \( N_1 = K_1 \times L_1 \) and \( N_2 = K_2 \times L_2 \):

\[
x_{k_1,k_2}[l_1, l_2] = x[l_1 + k_1 L_1, l_2 + k_2 L_2]
\]

with \( 0 \leq l_1 < L_1, 0 \leq l_2 < L_2, 0 \leq k_1 < K_1, 0 \leq k_2 < K_2 \).

Then, the periodogram estimator of each section is computed independently:

\[
\hat{P}_{xx}^{(\text{PER})}(\omega_1, \omega_2) = \frac{1}{L_1 L_2} |X_{k_1,k_2}(\omega_1, \omega_2)|^2.
\]

where \( X_{k_1,k_2}(\omega_1, \omega_2) \) denotes the Fourier transform of the section \( \{x_{k_1,k_2}[l_1, l_2]\} \).

Finally, the averaged periodogram estimator is calculated according to

\[
\hat{P}_{xx}^{(\text{AVPER})}(\omega_1, \omega_2) = \frac{1}{K_1 K_2} \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \hat{P}_{xx}^{(\text{PER})}(\omega_1, \omega_2).
\]

With regard to the spectral resolution (minimum frequency separation at which samples of \( \hat{P}_{xx} \) are independent) there has been a reduction factor of \( K_1 \) in \( \omega_1 \) and \( K_2 \) in \( \omega_2 \) as a result of computing the periodogram for each section independently. On the other hand, if the \( K_1 \times K_2 \) sections were uncorrelated and hence independent, the individual periodograms would be independent and hence uncorrelated. It
follows from Eq. (10) that the variance of the averaged periodogram estimator would be reduced by a factor of \( K_1 \times K_2 \):

\[
\text{VAR}[\hat{p}_{xx}^{\text{APER}}(\omega_1, \omega_2)] = \frac{1}{K_1^2 K_2^2} \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \text{VAR}[\hat{P}_{xx}^{\text{APER}}(\xi_{k_1,k_2}, \xi_{k_2,k_1}, \omega_1, \omega_2)]
\]

\[
\approx \frac{1}{K_1^2 K_2^2} \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} P_{xx}^2(\omega_1, \omega_2) = \frac{1}{K_1 K_2} P_{xx}^2(\omega_1, \omega_2).
\]  

(12)

In fact, as the sections are not overlapped, the single correlation that may exist among them should only be due to white noise. Therefore, the variance reduction factor will in general be less than \( K_1 \times K_2 \), although it represents a good approximation for processes not exhibiting sharp resonances [29].

In summary, the variance of the averaged periodogram is reduced at the expense of resolution. Here is the tradeoff that is found in spectral estimation approaches: variance versus resolution, or equivalently, variance versus bias.

For good variance reduction, a large number of sections should be chosen. However, for good resolution the sections should be as long as possible. For non-overlapping sections, both goals cannot be met simultaneously, and thus we are forced to trade off resolution (or, equivalently, bias) for variance by adjusting the number of sections \( K_1 \times K_2 \) or the length of them \( L_1 \times L_2 \).

However, as the sequence length \( N_1 \times N_2 \) increases, both the number of sections \( K_1 \times K_2 \) and section length \( L_1 \times L_2 \) can be allowed to increase so that, as \( N_1 \) and \( N_2 \) approach \( \infty \), the bias and the variance of the averaged periodogram spectral estimator can approach zero [27]. Consequently, periodogram averaging provides an asymptotically unbiased, consistent estimate of the true PSD \( P_{xx}(\omega_1, \omega_2) \).

Until this point, only non-overlapping sections have been assumed. Welch [31] studied the case of overlapping sections in the field of one-dimensional signal processing, and showed that if the overlap is one-half the section length, the variance is further reduced by almost a factor of 2 due to the doubling of the number of sections. Greater overlap does not continue to reduce the variance because the sections become less and less independent as the overlap increases. The extension to the two-dimensional case is not straightforward. We have thoroughly studied this case, considering the possibility of overlapping sections in both directions. Finally, we have come to the following formula, which relates the variance of the periodogram estimator and that of the periodogram averaging estimator for the case in which the maximum overlap is one-half the section length in both directions:

\[
\frac{\text{VAR}[\hat{p}_{xx}^{\text{APER}}(\omega_1, \omega_2)]}{\text{VAR}[\hat{P}_{xx}^{\text{APER}}(\omega_1, \omega_2)]} \approx \frac{1}{K_1 K_2} \left[ 1 + 2 \frac{K_1 - 1}{K_1} \rho(1, 0) + 2 \frac{K_2 - 1}{K_2} \rho(0, 1) + 4 \frac{K_1 - 1}{K_1} \frac{K_2 - 1}{K_2} \rho(1, 1) \right],
\]

(13)

where \( \rho(i, j) \) is the correlation coefficient between adjacent sections. The procedure followed to derive this formula is inspired by the one used by Welch [31], but in the two-dimensional case (see Appendix B).

5.1.3. Blackman–Tukey approach

The periodogram estimator studied in Section 5.1.1 can also be written [25, 27, 29] as

\[
\hat{p}_{xx}^{\text{APER}}(\omega_1, \omega_2) = \sum_{m_1 = -(N_1-1)}^{N_1-1} \sum_{m_2 = -(N_2-1)}^{N_2-1} \hat{r}_{xx}[m_1, m_2] e^{-i\omega_1 m_1} e^{-i\omega_2 m_2},
\]

where \( \hat{r}_{xx}[m_1, m_2] \) denotes the biased estimator of the ACF given by Eq. (5). In this formula, the periodogram is seen to be an estimator based on the Wiener–Khinchin theorem. This means that, although we view the periodogram method in Section 5.1.1 as a direct computation of the spectral estimate, the underlying correlation estimate of the ACF is in effect Fourier transformed to obtain the periodogram.
So the poor performance of the periodogram may be attributed to the poor performance of the ACF estimator, i.e. the biased ACF. In fact, from Eq. (5), we can see that \((N_1 - |m_1|) \times (N_2 - |m_2|)\) values of the sequence coming from the stochastic process are involved in computing a particular ACF value (also known as lag) \(\hat{r}_{xx}[m_1, m_2]\), no matter how large \(N_1\) and \(N_2\) are. Thus, as \(|m_1|\) and \(|m_2|\) are close to \(N_1\) and \(N_2\), respectively, only a few values of the sequence \(\{x[n_1, n_2]\}\) are involved in the computation. Therefore, we expect that the estimate of the ACF would be considerably more inaccurate for these values of \(m_1\) and \(m_2\) and, consequently, will also show considerable variation between adjacent values of \(m_1\) and \(m_2\).

On the other hand, when \(|m_1|\) and \(|m_2|\) are small, many more samples of the sequence are involved, and the variability at large values of \(|m_1|\) and \(|m_2|\) manifests itself in the Fourier transform as fluctuations at all frequencies and thus, for large \(N_1\) and \(N_2\), the periodogram estimate tends to vary rapidly with frequency.

One way to avoid this problem is to weight the higher ACF values (or lags) less, i.e. to compute the PSD estimate as:

\[
\hat{P}_{xx}^{(BT)}(\omega_1, \omega_2) = \frac{1}{N_1 - 1} \sum_{m_1 = -(N_1 - 1)}^{N_1 - 1} \frac{1}{N_2 - 1} \sum_{m_2 = -(N_2 - 1)}^{N_2 - 1} w[m_1, m_2] \hat{r}_{xx}[m_1, m_2] e^{-i\omega_1 m_1} e^{-i\omega_2 m_2},
\]

where \(w[m_1, m_2]\) is known as the lag window and has the following properties:

(i) \(0 \leq w[m_1, m_2] \leq w[0, 0] = 1,
(ii) \(w[ - m_1, - m_2] = w[ - m_1, m_2] = w[m_1, - m_2] = w[m_1, m_2],
(iii) \(w[m_1, m_2] = 0, \text{ for } |m_1| > M_1 \text{ or } |m_2| > M_2,\)

where \(M_1 < N_1\) and \(M_2 < N_2\).

Eq. (14) is called the Blackman–Tukey estimator. It is equivalent to the periodogram if \(w[m_1, m_2] = 1\) for \(|m_1| \leq M_1 = N_1 - 1\) and \(|m_2| \leq M_2 = N_2 - 1\). Also, the appropriate window can be found to be equivalent to periodogram averaging [27, 29].

As a consequence of applying the weighting to the ACF estimator, we can expect an increase in the bias. Therefore, the tradeoff between variance and bias appears again.

Many 2D windows are available. A 2D window is typically designed [25, 29] from the available 1D windows: Rectangular, Bartlett, Hanning, Hamming, Parzen, Tukey, etc. The work of Harris [32] constitutes a concise review of data windows and their spectral effects.

According to the frequency domain convolution, an alternate form of Eq. (14) consists of convolving the periodogram \(\hat{P}_{xx}^{\text{PER}}(\omega_1, \omega_2)\) with the Fourier transform of the lag window, commonly called spectral window:

\[
\hat{P}_{xx}^{(BT)}(\omega_1, \omega_2) = \hat{P}_{xx}^{\text{PER}}(\omega_1, \omega_2) \ast W(\omega_1, \omega_2),
\]

where \(\ast\) represents the convolution operation and \(W(\omega_1, \omega_2)\) denotes the spectral window.

All the lag windows used in this field have low-pass characteristics [25, 29, 32]. The result of the convolution is a smoothed version of the periodogram. It is for that reason that the Blackman–Tukey approach is also known as periodogram smoothing.

With regard to the different lag/spectral windows that we can use in this method, it has been proved [25] that the window shape, except for the case of rectangular lag window (i.e. the periodogram), does not have a substantial effect on the smoothing. Nevertheless, it is the lag window size the feature whose effect is critical: The larger the lag window, the shorter the spectral window is, tending to an impulse function, and as a result, the bias gets small. On the other hand, for a small variance, the lag window should be as short as possible. So, again, a bias (or, equivalently, spectral resolution) versus variance tradeoff is evident.

5.1.4. Summary

In summary, classical spectral estimation based on a very limited data set involves trading off bias/resolution for variance. A basic problem is that the PSD depends on an infinite number of
autocorrelation function lags, all of which need to be estimated to obtain a good spectral estimate. If a large amount of data is available, so that many lags can be estimated reliably, Fourier-based methods perform very well in both variance and spectral resolution/bias. In addition, they are quite simple from a computational point of view, since the major computations are Fourier transforms that can exploit the computational efficiency of FFT algorithms. But if the available data are very limited, one may reasonably ask whether it might be better to assume a model for the PSD or ACF that depends on only a finite set of parameters. If the number of data points is large relative to the number of PSD parameters, then good estimates of the parameters, and hence of the PSD, would be expected. This approach is termed parametric spectral estimation. In the next section, we will study one method belonging to this alternative spectral estimation: autoregressive modelling.

5.2. Parametric modelling: autoregressive (AR) model

Many discrete random processes encountered in practice are well approximated by a rational transfer function model. In this model, in a two-dimensional real process an input driving sequence \( u[m, n] \) and the output sequence \( x[m, n] \) that is to model the data are related by the linear difference equation

\[
x[m, n] = - \sum_{k \neq 0} \sum_{l \neq 0} a[k, l] x[m - k, n - l] + u[m, n] \tag{15}
\]

This linear model is termed an 2D autoregressive (AR) model. It is important to distinguish between the driving noise of the model \( u[m, n] \) and any observation noise. The AR model noise is not an additive or observation noise which is typically encountered in signal processing applications. \( u[m, n] \) is an innate part of the model and gives rise to the random nature of the observed process \( x[m, n] \). Any observation noise then needs to be modelled within the AR process by modification of its parameters.

Often the driving process is assumed to be a white noise sequence of zero mean and variance \( \sigma^2 \). The PSD of the noise is then \( \sigma^2 \). The PSD of the AR output process becomes

\[
P_{xx}(\omega_1, \omega_2) = \frac{\sigma^2}{|A(\omega_1, \omega_2)|^2} \tag{16}
\]

where

\[
A(\omega_1, \omega_2) = \sum_{k} \sum_{l} a[k, l] e^{-j[\omega_1 k + \omega_2 l]}, \text{ with } a[0, 0] = 1.
\]

This model is sometimes termed an all-pole model. The range of summation has purposely been left unspecified. There are many possible choices of the region of support and, in general, the spectral estimates will be different [28]. It may be shown that for a stable and causal system the region of support for the AR parameters must be in the nonsymmetric half plane (NSHP) (Fig. 1). From the viewpoint of spectral estimation, the constraint of a stable and causal system is questionable [33], but we will assume causality. A strong motivation for doing so is that any PSD may be modeled by an AR process under the following conditions [34]:

(i) The region of support for the AR parameters is the NSHP.

(ii) The number of AR parameters allowed in the model is infinite. The explicit relationship between the parameters of the AR model and the ACF can be derived directly from (15). Let the region of support for the AR parameters (including \( a[0, 0] = 1 \)) be denoted by \( S_{\text{NSHP}} \), where

\[
S_{\text{NSHP}} = \{ [k, l] / k = -p_1, \ldots, 0, \ldots, p_1; l = 1, 2, \ldots, p_2 \} \cup \{ [k, l] / k = 0, 1, \ldots, p_1; l = 0 \}.
\]
and TNSHP denotes the truncated NSHP (Fig. 1). Note that the NSHP, denoted by $S_{\text{NSHP}}$, is obtained by letting $p_1 \to \infty$, $p_2 \to \infty$. Multiplying Eq. (15) by $x[m-i, n-j]$ and taking the expected value yields

$$
\sum_{[k,l] \in S_{\text{temp}}} \sum a[k,l] r_{xx}[i-k, j-l] = \mathcal{E}(x[m-i, n-j]u[m,n])
$$

since all fields are assumed to be stationary. If the 2D system is linear shift invariant, the output $x[-i, -j]$ results in the convolution sum

$$
x[-i, -j] = \sum_{[k,l] \in S_{\text{temp}}} h[k,l]u[-i-k, -j-l].
$$

Note that the region of support of the impulse response is the entire NSHP, which follows from the causal and stable filter assumption. Hence,

$$
\mathcal{E}(x[-i, -j]u[0,0]) = \sum_{[k,l] \in S_{\text{temp}}} h[k,l] \mathcal{E}(u[-i-k, -j-l]u[0,0])
$$

because $h[i,j]$ is the impulse response of a causal filter

$$
\mathcal{E}(x[-i, -j]u[0,0]) = \begin{cases} 
0 & \text{for } [i,j] \in S_{\text{NSHP}}, \\
|h[0,0]|^2 & \text{for } [i,j] = [0,0],
\end{cases}
$$

where $S_{\text{NSHP}}$ denotes the NSHP with the $[0,0]$ point deleted or $S_{\text{NSHP}} = S_{\text{NSHP}} \cup [0,0]$. Since $h[0,0]$ may be shown to be unity [28], it follows that

$$
\sum_{[k,l] \in S_{\text{temp}}} a[k,l] r_{xx}[i-k, j-l] = \begin{cases} 
0 & \text{for } [i,j] \in S_{\text{NSHP}}, \\
\sigma^2 & \text{for } [i,j] = [0,0].
\end{cases}
$$

Eq. (17) have been termed the 2D Yule–Walker equations. They define a nonlinear relationship between the parameters of an AR process and the ACF. However, given the ACF we may determine the AR parameters
by solving a set of linear equations. In practice, the ACF samples \( r_{xy}(i,j) \) will be estimated from the data, by using Eq. (5), which is recognized as the biased ACF estimator. The AR parameter estimates are found by substituting the estimated ACF samples into 2D Yule–Walker equations and solving them. Finally, the estimated PSD can be derived from (16).

We can see that AR spectral estimation requires more ACF samples than AR parameters to be estimated in order to solve the Yule–Walker equations, so there is no autocorrelation matching property of this method, in contrast to the 1D case or the 2D MESE (Maximum Entropy Spectral Estimation). Therefore, it is immediately obvious that AR spectral estimation and MESE are not equivalent in 2D, although they are in 1D [28].

We must take into account several considerations in model order selection. Because the best choice of the AR model order is usually not known a priori, it is necessary in practice to postulate several model orders. If an AR spectral estimator is applied to a process that is not AR, then the true AR model would be one of infinite order. Any finite order AR model will introduce bias errors due to modelling inaccuracies. On the other hand, if the model order is chosen too large, the estimation errors increase and spurious peaks may appear in the spectral estimate. A trade-off must be effected between the desires to choose large model order to reduce the bias and to choose small model order to reduce the estimation errors.

6. Results

The importance of the task of improving the direct observation of the CTF rings in the Fourier transform of a TEM image has been discussed in the previous sections. Also, we have introduced the different factors that make the Fourier transform of these images seldom presents clear CTF rings. They can be summarized stating that taking directly the square amplitude of a discrete Fourier transform produces a magnitude that is a very poor estimator of the PSD, and that it is the reason why the CTF rings are usually so hard to be visualized.

Trying to fit this problem into a more general signal processing framework, it is immediately clear that, from Eq. (1), the formation of a TEM image can be considered as a stochastic process. We can further make the assumption that the process is stationary, assuming that no changes in the image formation mechanism occur while a micrograph is taken. Also, we can consider the process as ergodic in order to study the process from only one micrograph. Therefore, the type of spectral estimation analyses presented in the previous sections can be applied to this case, aiming at reducing the variance of the PSD in order to improve the visibility of the CTF rings.

We have divided this section into two subsections. In Section 6.1, we have applied the periodogram, periodogram averaging (PA) and autoregressive modelling (AR) to an artificial image simulating TEM images that could be obtained from amorphous carbon film [9]. The goal was to objectively compare among the different methods, studying properties, advantages and disadvantages of each of them. In Section 6.2 we have applied these methods to a real TEM image. The goal in this case was to study how well the different approaches adapt themselves to real problems.

6.1. Simulated data

6.1.1. Generation of the artificial image

The objective in this section was to create an image which simulated somehow an image coming from a TEM. To this end, we generated an artificial image of \( 1024 \times 1024 \) pixels by assigning to each pixel random values obtained from a white gaussian noise distribution of mean zero and standard deviation one, denoted as \( \mathcal{N}(0,1) \). This random image intended to simulate an amorphous carbon film [9]. The Fourier transform of this image is represented in Eq. (1) by the term \( F \).
Then, a CTF with the following parameters was applied to that random image:

- Defocusing: 800 nm (No astigmatism was considered)
- Spherical aberration $c_s$: 2.0 mm
- Electron voltage: 120 kv
- Amplitude contrast: 0.2
- Sampling rate: 0.25 nm/pixel.

No damping factors were taken into account. The CTF is represented by the term $H$ in Eq. (1).

Finally, further noise had to be added up to the image after the CTF application according to the model of Eq. (1). We have considered white gaussian noise again, but this time distributed as $\mathcal{N}(0,2)$ in order to still make more difficult the observation of the CTF rings. Its Fourier transform is denoted by the term $N$ in Eq. (1). The final image obtained in this way has been the one used in the different spectral estimation approaches that we want to compare. It essentially simulates a quite noisy TEM image of a carbon film.

We can compare the performance of the different spectral estimators because we can compute the theoretical PSD of our artificial image: The PSD of white noise for an infinite sequence is a constant term equal to its variance [25]. On the other hand, the PSD of an infinite deterministic signal is equal to the squared Fourier transform amplitude [25], and, therefore, the PSD of the CTF is its squared amplitude. Finally, we can conclude that the theoretical PSD of our artificial image is

$$P_{xx}(\omega_1, \omega_2) = |CTF(\omega_1, \omega_2)|^2 \sigma_{r1}^2 + \sigma_{r2}^2 = |CTF(\omega_1, \omega_2)|^2 + 4,$$

where $\sigma_{r1}^2$ and $\sigma_{r2}^2$ represent the variances of the two random distributions used in the generation of the artificial image, distributed as $\mathcal{N}(0,1)$ and $\mathcal{N}(0,2)$, respectively.

### 6.1.2. Spectral estimation by Fourier-based methods

In this section, the estimation of the PSD of our artificial image will be accomplished by means of the periodogram and a series of periodogram averaging experiments. The periodogram smoothing approach will also be used, although in a form tailored to the radial symmetry of the CTF instead of the standard Blackman–Tukey approach.

The main goals of applying Fourier methods to the simulated image were:

(i) To study the tradeoff between bias/resolution and variance. This analysis was performed by means of a set of experiments, varying the number/size of the sections into which the image was divided.

(ii) To analyse the influence of the overlapping sections. Here we studied how overlapping sections in both directions affect the PSD estimate and variance reduction factor.

With regard to the first goal, Table 1 summarizes a set of experiments carried out to analyse the tradeoff resolution versus variance. The theoretical variance was calculated according to Eqs. (12) and (13).

The experiment labelled with a 0 in Table 1 corresponds to just taking the periodogram of the image. This result is shown in Fig. 2a. As the image is real, the PSD is uniquely specified [26] for $0 \leq \omega_1 \leq \pi$, $-\pi \leq \omega_2 \leq \pi$, and this is what is shown in the figure. One can notice the huge variability in the periodogram, precluding the separation of the CTF rings. It is clear from that the periodogram is indeed an unreliable estimator of the PSD of our image.

We then carried out the remaining experiments in an attempt to reduce the strong variability in the periodogram by computing different averaged periodogram estimates. We can appreciate how the variability is reduced as the number of sections increases at expense of resolution. Fig. 2b–Fig. 2d show the experiments labelled with 3, 6 and 9, respectively. The more sections used, the clearer the rings of the CTF are visualized. However, the spectral resolution is also decreased. The latter point is clearly observed in Fig. 2d, where the CTF rings corresponding to high resolution, i.e. those that should be the finest, are impossible to be distinguished.
Table 1
The columns are labelled indicating the number of the experiment that is going to be used later on in the main text to refer to it, number of sections into which the artificial image was divided, size of the sections (in pixels), overlapping between sections (in pixels), measured variance of the PSD estimate, theoretical variance and, finally, spectral resolution (in cycles/pixel)

<table>
<thead>
<tr>
<th>Exp.</th>
<th>Sections</th>
<th>Size</th>
<th>Overlap</th>
<th>Variance</th>
<th>Th. variance</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1024 × 1024</td>
<td></td>
<td>20.79</td>
<td>20.79</td>
<td>1/1024</td>
</tr>
<tr>
<td>1</td>
<td>2 × 2</td>
<td>512 × 512</td>
<td></td>
<td>5.27</td>
<td>5.20</td>
<td>1/512</td>
</tr>
<tr>
<td>2</td>
<td>3 × 3</td>
<td>256 × 256</td>
<td>256 × 256</td>
<td>4.22</td>
<td>4.11</td>
<td>1/512</td>
</tr>
<tr>
<td>3</td>
<td>4 × 4</td>
<td>128 × 128</td>
<td>128 × 128</td>
<td>1.37</td>
<td>1.30</td>
<td>1/256</td>
</tr>
<tr>
<td>4</td>
<td>7 × 7</td>
<td>128 × 128</td>
<td>64 × 64</td>
<td>0.95</td>
<td>0.87</td>
<td>1/256</td>
</tr>
<tr>
<td>5</td>
<td>8 × 8</td>
<td>128 × 128</td>
<td></td>
<td>0.39</td>
<td>0.33</td>
<td>1/128</td>
</tr>
<tr>
<td>6</td>
<td>10 × 10</td>
<td>128 × 128</td>
<td>30 × 30</td>
<td>0.33</td>
<td>0.25</td>
<td>1/128</td>
</tr>
<tr>
<td>7</td>
<td>12 × 12</td>
<td>128 × 128</td>
<td>50 × 50</td>
<td>0.31</td>
<td>0.24</td>
<td>1/128</td>
</tr>
<tr>
<td>8</td>
<td>15 × 15</td>
<td>128 × 128</td>
<td>64 × 64</td>
<td>0.27</td>
<td>0.20</td>
<td>1/128</td>
</tr>
<tr>
<td>9</td>
<td>16 × 16</td>
<td>64 × 64</td>
<td></td>
<td>0.12</td>
<td>0.08</td>
<td>1/64</td>
</tr>
<tr>
<td>10</td>
<td>31 × 31</td>
<td>64 × 64</td>
<td>32 × 32</td>
<td>0.09</td>
<td>0.05</td>
<td>1/64</td>
</tr>
</tbody>
</table>

Fig. 2. Fourier-based spectral estimates. (a) Periodogram, (b) PA resulting from the experiment labelled 3 in Table 1, (c) PA resulting from the experiment labelled 6 in Table 1, (d) PA resulting from the experiment labelled 9 in Table 1.

Another issue noteworthy from the experiments is the validity of the formulas of variance reduction factor. It can be appreciated how the real measured variances of the different PSD estimates are close to the corresponding theoretical variances computed from the Eqs. (12) and (13), in spite of the simplifications that had to be made to derive these equations.

Finally, the last column in Table 1 shows the spectral resolution of each PSD estimate in cycles/pixel. By multiplying this value by the maximum frequency, i.e. \( \frac{1}{\text{Sampling Rate}} \), the spectral resolution in \( \text{nm}^{-1} \) can be obtained.

In order to observe the reduction of the variance of the PSD, the graphs in Fig. 3a and Fig. 3b are shown. Fig. 3a corresponds to a scan of the periodogram at \( Y = \omega_2 = 0 \). Its high variability clearly can be
Fig. 3. Graphs of the Fourier-based spectral estimates. (a) Scan at \( Y = \omega_1 = 0 \) of the Periodogram. (b) Scan at \( Y = \omega_2 = 0 \) of the PA spectral estimate from the experiment labelled 6 in Table 1. (c) Radial average of the Periodogram. (d) Radial average of the PA spectral estimate from the experiment labelled 6 in Table 1. Here dotted line is the spectral estimate and solid line is the theoretical PSD for clarity, but in the following graphs dotted line will be the theoretical PSD to see better the details of the estimated curves.

appreciated. In Fig. 3b the scan at \( Y = \omega_2 = 0 \) of the averaged periodogram estimate presented in Fig. 2c is shown. This estimate corresponds to the experiment labelled 6 in Table 1, i.e. the averaging of the periodograms of \( 10 \times 10 \) sections of \( 128 \times 128 \) pixels with an overlap of \( 30 \times 30 \) pixels. The reduction in variability in both cases can be easily observed from Fig. 3a and Fig. 3b. However, despite this variability reduction, it is still not easy to determine the CTF rings from a direct observation of the graph in Fig. 3b, and a further processing of the signal is needed.

At this point we decided to introduce a variant of another PSD estimator approach introduced before that allowed us for specific information on the general shape of the CTF to be introduced in the process. This additional technique is known as periodogram smoothing, and essentially consists of smoothing the periodogram directly in order to reduce its variance. The standard approach amounts to defining square convolution kernels that are then used to smooth the periodogram, and it is usually referred to as the
Blackman–Tukey approach. In principle, this standard approach was of no interest for our application, since it can be proved that it is equivalent to a periodogram averaging without overlapping [27, 29]. However, with the prior knowledge that the CTF has a rotational (or ellipsoidal, if astigmatism is present) symmetry, we can now define a new application-specific type of periodogram smoothing operator that averages the periodogram – or improved estimator of it – along shells of constant radius, producing in this way smoother radial averages of PSD estimates.

Applying the shell-smoothing periodogram approach on the results obtained either when only the periodogram is calculated or when a periodogram averaging is performed, we obtained results that have a much lower variance. Fig. 3c corresponds to the radial average of the periodogram, whereas Fig. 3d corresponds to the radial average of the PSD estimate in Fig. 2c, i.e. the experiment labelled 6. In these graphs, we can still see the effect of the variance reduction, but the most important issue to be emphasized is the fact that all the CTF rings, or more correctly, all the CTF zeroes, exactly coincide with the CTF zeroes of the theoretical PSD given by \(|\text{CTF}|^2 + \sigma^2\). Certainly, this has been possible due to the non-astigmatic character of our artificial image. If the image were strongly astigmatic, this averaging could not be applied.

In relation to the second goal, i.e. the influence of overlapping the sections, we can conclude from the results in Table 1 that it allows the variance to be further decreased with no spectral resolution decrease. A maximum overlap of one half the section length in both directions has been allowed, as mentioned in Section 5.1.2. For instance, the experiments marked as 5–8 in Table 1 use an increasing overlapping of the sections: non-overlap, 30 x 30, 50 x 50, 64 x 64, respectively. As shown in the Table 1, a decrease in variance is achieved as the overlap increases: 0.39, 0.33, 0.31 and 0.27, respectively, with the advantage that no reduction in spectral resolution takes place. Nevertheless, if the sections overlapped more than one-half the section length, no further gain in variance would be obtained, because the sections would be less and less independent [31]. So, if we need to decrease the variance even further, the only solution would consist of dividing the image into a greater number of sections, but smaller, which involves a reduction of spectral resolution. For instance, experiment 9 versus experiment 8.

So, the strategy is clear: In order to decrease the periodogram variance, we must do periodogram averaging with sections of \(N \times N\), for instance, where \(N\) fits to our resolution requests. If we want to decrease the variance of the averaged periodogram even more, we can allow section overlapping, \(N/2 \times N/2\) at most. A shell-smoothing periodogram approach can now be used to obtain smooth radial profiles. However, if this is not yet enough, then \(N\) has to be decreased and the price to be paid for this variance reduction is spectral resolution.

So far, we have considered overlapping sections in both directions at the same time. We have done experiments with overlapping sections in only one direction. The results obtained (not shown here) tell us that the variance is decreased by a factor lower than that obtained with twofold overlapping. Therefore, twofold overlapping allows a further variance reduction with no decrease of resolution.

6.1.3. Autoregressive spectral estimation

By applying the AR model introduced in Section 5.2, the estimation of the PSD of our artificial image was obtained. The AR parameters were estimated by using Eq. (17) with the ACF estimator of Eq. (5), and the PSD estimate was then derived from Eq. (16).

Since the best choice of order is not generally known a priori, in practice it is usually necessary to postulate several model orders. We performed a set of experiments changing the model order, with \(p_1 = p_2\), obtaining the AR spectral estimates with model order ranging from \(p_1 = p_2 = 5\) to \(p_1 = p_2 = 40\) at steps of 5.

In Fig. 4 the PSD estimates from a 20 x 20 order AR model and from a 30 x 30 order AR model are shown. The improvement in the separation of the CTF rings can be seen. The most important point to be emphasized is the resolution increase of the PSD estimate as the model order is raised. This is clearly seen in Fig. 4.
In order to observe the coincidence of the PSD estimates with the theoretical PSD for several model orders, the graphs in Fig. 5 are shown. From Fig. 5 the improvement of the PSD estimate as the model order is increased can be appreciated. In every graph of Fig. 5, the scan of the PSD estimate at \( Y = \omega_2 = 0 \) and a radial average of the same PSD estimate are shown with dotted and dashed lines, respectively.

Only the PSD estimates coming from AR models with order 10 \( \times \) 10, 20 \( \times \) 20, 30 \( \times \) 30 and 40 \( \times \) 40 were analysed in that way, and the result of these analyses are shown in Fig. 5a, Fig. 5b, Fig. 5c and Fig. 5d, respectively. It can be clearly observed how the PSD estimate is improved as the model order is increased: the greater the model order, the more zeroes of the radial average of the PSD estimate coincide with those of the theoretical PSD. In Fig. 5a the first valley of the radial average coming from the 10 \( \times \) 10 AR model PSD estimate hardly matches the first zero of the theoretical PSD. In Fig. 5b, only a few valleys of the radial average of the PSD estimate match the zeroes of the theoretical PSD. In Fig. 5c, it can be clearly appreciated how all the CTF zeroes of the radial average of the PSD estimate match those of the theoretical PSD, although the height of the oscillations does not coincide. Finally, the radial average of the 40 \( \times \) 40 AR model PSD estimate almost perfectly matches the theoretical PSD, regarding both the zeroes as well as the height of the oscillations. Something similar occurs with the scan at \( Y = \omega_2 = 0 \) of the PSD estimates as the model order is increased.

Another issue worth notice is the fact that there is an increasing variability at low frequencies around the spectral origin, as far as the radial averages are concerned. This point may be explained by the fact that spurious peaks may appear in the AR PSD estimates as the model order increases [30]. At low frequencies the radial averaging involves a few terms, while at higher frequencies the radial averaging involves more terms. Therefore, it is logical to think that spurious peaks are overcome at higher frequencies and, as a consequence, a smoother shape is obtained. The shape of the scans at \( Y = \omega_2 = 0 \) of PSD estimates as model order increases would be justified by a similar reasoning: the spurious peaks that AR models may introduce in the PSD estimates as model order increases.

As our process is not pure AR, the true AR model would be one of infinite order. In practice, we calculate several PSD estimates by increasing the AR model order until a satisfactory PSD estimate is obtained. Here we must bear in mind that a too low order results in a highly smoothed PSD estimate and, on the other hand, a very high order may introduce spurious peaks into the spectral estimate.
6.1.4. Comparison

In order to carry out an objective comparison between the periodogram averaging and the autoregressive modelling spectral estimation approaches, the best results coming from each of the two methods were selected. Regarding the PA approach, the experiment labelled 8 in Table 1 was chosen in terms of the tradeoff variance versus spectral resolution. That particular PSD estimate resulted from averaging 225 (15 x 15) segments of 128 x 128 pixels each with one-half the segment length overlapping. This PSD estimate is shown in Fig. 6a.

As far as AR the approach is concerned, the PSD estimate resulting from a 40 x 40 order AR model was chosen. It is shown in Fig. 6b. Both pictures in Fig. 6 present the whole spectral space $-\pi \leq \omega_1, \omega_2 \leq \pi$.

The best AR result presented in Fig. 6b can be compared to those presented in Fig. 4. The improvement in the separation of the CTF rings at high frequencies can clearly be appreciated. This effect is due to the better spectral resolution of the result using a 40 x 40 order compared to those using lower order AR models.
Fig. 6. Best spectral estimates with Periodogram Averaging and Autoregressive Modelling. (a) PA resulting from the experiment labelled 8 in Table 1. (b) 40 x 40 order AR model.

Also, one can compare the best result of the PA approach, shown in Fig. 6a, to the other different results presented in Fig. 2. The selected best result is between the results shown in Fig. 2c and Fig. 2d, in terms of the tradeoff variance versus resolution.

From Fig. 6, the improvement in terms of spectral resolution of the AR spectral estimate with respect to that of PA spectral estimate is clearly appreciated. The CTF rings corresponding to high frequencies, which are the finest rings, are very well visualized in the AR spectral estimate, whereas in the PA spectral estimate those rings are very difficult to be distinguished.

The graph in Fig. 7 presents the theoretical PSD together with the radial averaging coming from the PA and the AR spectral estimates. The very good matching of the theoretical rings with those obtained from both spectral estimates can be clearly observed. Then, we can conclude that, after radial averaging, both methods detect very well the CTF zeroes. However, as far as the height of the oscillations is concerned, the AR spectral estimate is better than the PA estimate.

The disadvantage of the AR approach arises from a computational point of view. The PA spectral estimation approach is considerably cheaper than the AR method. The AR approach involves solving a set of linear equations where the number of equations depends on the square of the model order [30]. On the other hand, the PA approach consists of a set of fast Fourier transformations, where the number of transformations is given by the number of segments.

6.2. Experimental data

6.2.1. Imaging

The real data consisted of a micrograph containing a large crystal of bacteriophage φ29 P10 connector [35] imaged in the frozen hydrated state on a carbon support film, recorded using low dose microscopy. The electron micrograph was obtained using a Philips EM420 (operated at 120 kV) and recorded on Kodak SO-163 film. The micrograph was obtained with a nominal magnification of 36 000 x and an underfocus of 880 nm. The selected area was digitized with a 10 μm pixel size on a Joyce-Loebl Mk4 flat bed microdensitometer as described by Henderson et al. [11]. The final image had 4000 x 4000 pixels, with a pixel size corresponding to 0.278 nm. The image and its Fourier transform amplitude are shown in Fig. 8.
6.2.2. Spectral estimation

As mentioned above, the goal pursued in this section was to study how well the different spectral estimation approaches performed a real problem.

Estimates of the PSD of the real data presented in the previous section were obtained by means of the periodogram, periodogram averaging and autoregressive modelling. Also, radial averaging – a variation of periodogram smoothing – was used since the image could be considered as non-astigmatic.
Although several experiments were carried out with periodogram averaging and autoregressive modelling methods, only the best results obtained with each of them will be presented here.

With regard to the periodogram averaging approach, the best result in terms of the tradeoff variance versus resolution was judged visually to be the one obtained by dividing the image into 225 (15 × 15) sections of 500 × 500 pixels with one half the section length of overlapping.

In relation to the autoregressive modelling approach, the best result in terms of resolution and computation time was the one obtained by using a 60 × 60 AR model.

Fig. 9 shows the radial averages coming from the periodogram estimate (solid line), the best periodogram averaging estimate (dotted line) and the best AR estimate (dashed line) up to 0.7 nm resolution (spatial frequency of 1.43 nm⁻¹). In Fig. 9 four minima can be distinguished from the three graphs. This fact means that the three methods may be valid to detect the CTF zeroes. However, it is obvious that this approach has been possible only due to the non-astigmatic character of the image that allows radial averaging to be used. Otherwise, radial averaging could not have been used at all.

The peaks at low frequencies appearing in the graphs are due to the height of the reflections corresponding to the imaged specimen. This may be explained by the fact that at low frequencies the radial averaging involves a few terms and, as a consequence, the reflections are not hidden by the averaging operation, as occurs with higher frequencies.

In order to facilitate the observation of the CTF oscillations, a two-dimensional gaussian function was fitted to the minima of the radial averages in Fig. 9 and was then used for background subtraction in all the PSD estimates [13].

Fig. 10 shows the results of the periodogram (a), periodogram averaging (b) and autoregressive modelling (c) estimates after background subtraction. The maximum resolution corresponds to 0.7 nm. From Fig. 10a, only one CTF zero may be distinguished. It is clear from this figure that the periodogram is an unreliable estimator of the PSD of the image. However, Fig. 10b and Fig. 10c show spectral estimates that allow up to three CTF zeroes to be clearly separated.

![Fig. 9. Comparison of the radial average of the different spectral estimates for the real image. Solid line is the Periodogram Averaging, dashed line is the Autoregressive Modelling and dotted line corresponds to the Periodogram.](image-url)
It is also worth mentioning that from Fig. 10b and Fig. 10c the reflections of the imaged specimen located at low frequencies are easily visible. However, their location is very coarse because of the low spectral resolution of both the periodogram averaging and the autoregressive modelling approaches.

From the results of periodogram averaging (Fig. 10b) and autoregressive modelling (Fig. 10c) spectral estimates, the defocus degree used to obtained the image was estimated by using the algorithms explained in Appendix A. The average defocus degree was found to be 880 nm underfocus, with no consideration of astigmatism. From this defocus degree, the theoretical CTF that is affecting the image is completely specified (see Section 2). Fig. 11 presents the radial averages of the periodogram averaging and autoregressive modelling spectral estimates coming from Fig. 9 after background subtraction. Fig. 11 also presents the PSD of the theoretical CTF (squared CTF, because CTF is a deterministic signal) affecting the image. A scale factor had to be used in order to compare the CTF PSD to the spectral estimates of the image.

It is clearly observed that the four minima found in the spectral estimates match the four first zeroes of the theoretical CTF PSD. This confirms the reliability of the procedures presented in this work to estimate the defocus degree and, therefore, to determine the theoretical (without envelope function effects) CTF of the microscope.

The radial averaging shown in Fig. 11, likewise with Fig. 9, would be unuseful if the image under consideration were astigmatic, and, as a result, the observation of the spectral estimates would be the only method to determine the CTF zeroes. In this sense, Fig. 10 clearly shows that the periodogram is a poor PSD estimate, and periodogram averaging and autoregressive modelling result in more reliable PSD estimates.

7. Discussion

We address in this work the very general problem that appears in Electron Microscopy when we aim at studying the underline image aberrations introduced by the Contrast Transfer Function (CTF) of the microscope with the help of the Fourier transform of the image calculated in a computer after digitizing the image.

Usually, Fourier transform amplitudes are digitally calculated and squared as a way to estimate the power spectrum of the experimental image. The rational behind this approach is that this estimation of the power spectrum of the experimental image will provide us with the information needed to characterize the CTF. However, experience tells us that power spectrum calculated in this way may be rather noisy, precluding a proper CTF characterization.
In this work we have analysed this behaviour from the generalized perspective of power spectrum estimation in multidimensional signal processing, reaching the conclusion that power spectrum calculated by a direct elevation to the square of the amplitudes of the Fourier transform of a discretized image are rather poor estimator of the power spectrum of the experimental image.

In an effort to obtain better estimation, we have studied different methods to calculate more accurate estimates of the power spectrum of the original image. The results of the comparison between several of these approaches have been presented in the previous sections. This contribution comes timely in a moment in which this problem starts being realized in the field, and more or less ad hoc procedures for better power spectrum estimation are being proposed. A review of them is presented in the following.

Recently, Avila-Sakar et al. [12] and Zhou et al. [14, 15] have implemented a way to estimate the CTF of the microscope by means of averaging the amplitude corresponding to Fourier transforms of fragments of the images. Certainly, their realization that spectral estimation methods had to be reevaluated has been vital. In a sense their approach is “similar” to the periodogram averaging method presented in this work, but the important mathematical difference is that for a proper estimation of the power spectrum of a stochastic process, the magnitude to be averaged has to be the periodogram (squared amplitude) and not amplitude itself (this has been studied in depth in the context of signal processing, and the reader is referred to [29] for a full explanation). In the context of electron microscopy this was first realized by Zhu and Frank [13]. Obviously, the zeroes of the functions resulting from averaging either the amplitude or the squared amplitude
are going to be placed in the same locations. However, the shape of the resulting averaged curves themselves are going to be different.

In this work we have placed the problem of spectral estimation encountered in electron microscopy within a rigorous signal processing framework. Within this context we have made evident that a direct computation of the power spectrum from the squared amplitudes of the Fourier transform of the digitized image – usually referred in the signal processing field as the "periodogram" – is a rather poor estimate of the power spectrum of the original image because it has a high variance. We have then analysed two different approaches for a proper spectral estimation, each of them with their relative merits in terms of potential frequency resolution and computational complexity.

In general, two tradeoffs are found in power spectrum estimation. The first one appears in virtually all aspects of signal processing, and it refers to whether we are going to incorporate some form of prior knowledge about the image formation system into the power spectrum estimation process or not. The second one specially applies to the case in which no prior information is introduced in the system, and it relates in an inverse way the terms resolution and variance. That is, a gain in resolution translates into a higher variance, while a reduction in resolution translates into a smooth estimate.

We have analysed three different methods: periodogram averaging, periodogram smoothing, and autoregressive methods. The first ones do not introduce any a priori assumption into the signal formation model. With regard to the adaptation of the periodogram smoothing method presented in this work, it requires to know either that the image has no astigmatism, or the amount of astigmatism. Autoregressive methods certainly introduce certain assumptions. However, the experimental results obtained both for computed simulated images and for real ones indicate that these general assumptions do generally apply in practical cases.

As a consequence, our results indicate that the improvements obtained when more reliable power spectrum estimators are used instead of the periodogram are extremely noticeable. In general, the simple method of periodogram averaging works in general very well while being quite fast from a computational point of view, since it exploits the efficiency of the FFT algorithms. This approach is expected to work better when the amount of samples gets larger. However, if the discretized images are not very large, then autoregressive methods may be the solution of choice, in spite of being very computing-intensive. In all cases, the adaptation of the periodogram smoothing presented here also improves significantly the results, although it necessarily needs prior knowledge on the astigmatism.

8. Conclusion

In this work, we have placed the problem of CTF estimation within the spectral estimation framework. We have analysed periodogram averaging, periodogram smoothing and autoregressive methods, obtaining far better results in all cases than using the amplitude of the Fourier transform alone, as it is common practice in this field up to now.

Certainly, we hope that the methodologies introduced in this work will help other researches in their work towards a better understanding of the informational content of the electron microscopy images.

Acknowledgements

Work in this topic was prompted by a fruitful discussion held with Dr. Joachim Frank in November 94, since then the two groups have independently approached the problem of spectral estimation in similar ways, sharing their final results prior to publication. Dr. Jose-Maria Valpuesta provided the real data presented in this work. José Jesús Fernández is a predoctoral fellow of the CSIC. This work has been possible in part thanks to a grant from the Spanish CICYT (grant number BIO95-0768).
Appendix A. Algorithms to derive the defocus degree from the CTF rings

A.1. Algorithm I

Algorithm to estimate the defocus degree from the positions of the CTF zeroes that can be distinguished in the Fourier transform of the recorded image. The algorithm starts with \( n = 1 \).

1. Find the CTF zero of order \( n \) in the Fourier transform of the recorded image.
2. Compute the spatial frequency \( \rho_n \) corresponding to the CTF zero being processed.
3. Compute the defocus value for the current CTF zero, denoted by \( \Delta f_n^0 \), by using the following formula:

\[
\Delta f_n^0 = \frac{C_s \rho_n^4 \lambda^4 + 2n}{2 \rho_n^2 \lambda},
\]

where \( C_s \) and \( \lambda \) are microscope parameters, \( \rho_n \) has been calculated in step 2, and \( n \) is the order of the CTF zero.

The formula has been derived directly from Eq. (2) by considering the CTF term as zero (no amplitude contrast is considered):

\[
H(\rho) = -\sin \chi(\rho) = 0 \Rightarrow \chi(\rho) = \pi n.
\]

4. Repeat steps 1–3 for \( n + 1 \), if there still is any visible CTF zero in the Fourier transform of the recorded image.
5. Calculate the average defocus value by means of linear least-squares fitting of \( \{ \Delta f_n^0 \} \).

A.2. Algorithm II

Algorithm to estimate the defocus degree from the positions of the CTF peaks that can be distinguished in the Fourier transform of the recorded image. The algorithm starts with \( n = 1 \).

1. Find the CTF peak of order \( n \) in the Fourier transform of the recorded image.
2. Compute the spatial frequency \( \rho_n \) corresponding to the CTF peak being processed.
3. Compute the defocus value for the current CTF peak, denoted by \( \Delta f_n^1 \), by using the following formula:

\[
\Delta f_n^1 = \frac{C_s \rho_n^4 \lambda^3 + 2n - 1}{2 \rho_n^2 \lambda},
\]

where \( C_s \) and \( \lambda \) are microscope parameters, \( \rho_n \) has been calculated in step 2, and \( n \) is the order of the CTF peak.

The formula has been derived directly from Eq. (2) by considering the CTF term as unit (no amplitude contrast is considered):

\[
|H(\rho)| = |\sin \chi(\rho)| = 1 \Rightarrow \chi(\rho) = \frac{2n - 1}{2} \pi.
\]

4. Repeat steps 1–3 for \( n + 1 \), if there still is any visible CTF peak in the Fourier transform of the recorded image.
5. Calculate the average defocus value by means of linear least-squares fitting of \( \{ \Delta f_n^1 \} \).

Appendix B. Derivation of variance reduction factor for periodogram averaging

In this appendix the main steps in the derivation of Eq. (13) are specified. This formula relates the variance of the periodogram estimator and that of the periodogram averaging estimator for the case in which the
maximum overlap is one-half the section length in both directions. It follows the steps used by Welch [31], but in the two-dimensional case.

From Eq. (11) we can conclude that the variance of the averaged periodogram estimator is

\[
\text{VAR}[\hat{F}_{xx}^\text{APER}](\omega_1, \omega_2)] = \frac{1}{K_1^2 K_2^2} \text{VAR} \left[ \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \hat{F}_{x_1 x_2}^\text{APER}(\omega_1, \omega_2) \right].
\]  

However, when overlapping among the sections is considered, the sections are not uncorrelated nor independent. As a consequence, the individual periodograms would not be uncorrelated nor independent, and this fact precludes the derivation shown in Eq. (12).

The following property has to be taken into consideration:

\[
\text{VAR}[a + b] = \text{VAR}[a] + \text{VAR}[b] + 2 \text{COV}[a, b],
\]

where COV represents covariance.

The following term in Eq. (B.1),

\[
\sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \hat{F}_{x_1 x_2}^\text{APER}(\omega_1, \omega_2),
\]

can be broken down into:

\[
\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2) + \left( \sum_{k_1=1}^{K_1-1} \sum_{k_2=0}^{K_2-1} \hat{F}_{x_1 x_2}^\text{APER}(\omega_1, \omega_2) + \sum_{k_1=0}^{K_1-1} \sum_{k_2=1}^{K_2-1} \hat{F}_{x_1 x_2}^\text{APER}(\omega_1, \omega_2) \right).
\]

Let us define the term SUMRES (it stands for sum residual) as

\[
\text{SUMRES}[n_1, n_2] = \sum_{k_1=n_1+1}^{K_1-1} \sum_{k_2=0}^{K_2-1} \hat{F}_{x_1 x_2}^\text{APER}(\omega_1, \omega_2) + \sum_{k_1=0}^{K_1-1} \sum_{k_2=n_2+1}^{K_2-1} \hat{F}_{x_1 x_2}^\text{APER}(\omega_1, \omega_2).
\]

Then, Eq. (B.4) can be written as

\[
\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2) + \text{SUMRES}[0, 0].
\]

By substituting Eq. (B.6) into Eq. (B.1), and applying the property expressed in Eq. (B.2), the following result is obtained:

\[
\text{VAR}[\hat{F}_{xx}^\text{APER}](\omega_1, \omega_2)] = \frac{1}{K_1^2 K_2^2} (\text{VAR}[\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2)]
\]

\[
+ \text{VAR}[\text{SUMRES}[0, 0]] + 2 \text{COV}[\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2), \text{SUMRES}[0, 0]]).
\]

Then, the term SUMRES[0, 0] as defined by Eq. (B.5) can be broken into

\[
\text{SUMRES}[0, 0] = \hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2) + \text{SUMRES}[0, 1].
\]

By substituting Eq. (B.8) into Eq. (B.7),

\[
\text{VAR}[\hat{F}_{xx}^\text{APER}](\omega_1, \omega_2)] = \frac{1}{K_1^2 K_2^2} (\text{VAR}[\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2)]
\]

\[
+ \text{VAR}[\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2)] + \text{VAR}[\text{SUMRES}[0, 1]]
\]

\[
+ 2 \text{COV}[\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2), \text{SUMRES}[0, 1]]
\]

\[
+ 2 \text{COV}[\hat{F}_{x_0 x_0}^\text{APER}(\omega_1, \omega_2), \text{SUMRES}[0, 0]]).
\]
By breaking down the successive terms \( \text{SUMRES}[.\) , the following result is obtained:

\[
\text{VAR}[\hat{\beta}_{xx}(\omega_1, \omega_2)] = \frac{1}{K_1 K_2} \left( \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \text{VAR}[\hat{\beta}(\omega_1, \omega_2)] \right.
\]

\[
+ 2 \sum_{n_1=0}^{K_1-2} \sum_{n_2=0}^{K_2-2} \text{COV}[\hat{\beta}(\omega_1, \omega_2), \text{SUMRES}[n_1, n_2]] \right) \tag{B.10}
\]

At this point, we have to remember that for a stationary process, we can write the covariance function as

\[
d(i, j) = \text{COV}[\hat{\beta}(\omega_1, \omega_2), \hat{\beta}(\omega_1, \omega_2)]
\]

independent of \( k_1, k_2 \). The first term in Eq. (B.10) is equal to

\[
\sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \text{VAR}[\hat{\beta}(\omega_1, \omega_2)] = \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} d(0, 0) = K_1 K_2 d(0, 0). \tag{B.12}
\]

On the other hand, it is easily shown that

\[
\sum_{n_1=0}^{K_1-1} \sum_{n_2=0}^{K_2-1} \text{COV}[\hat{\beta}(\omega_1, \omega_2), \text{SUMRES}[n_1, n_2]]
\]

\[
= 2 \sum_{i=1}^{K_1-1} \sum_{j=1}^{K_2-1} (K_1 - i) (K_2 - j) d(i, j) + K_1 \sum_{i=1}^{K_1-1} (K_2 - j) d(0, j) + K_2 \sum_{j=1}^{K_2-1} (K_1 - i) d(i, 0). \tag{B.13}
\]

By substituting Eqs. (B.12) and (B.13) into Eq. (B.10), we have

\[
\text{VAR}[\hat{\beta}_{xx}(\omega_1, \omega_2)] = \frac{1}{K_1 K_2} \left( d(0, 0) + 2 \sum_{j=1}^{K_2-1} \frac{K_2 - j}{K_2} d(0, j) + 2 \sum_{i=1}^{K_1-1} \frac{K_1 - i}{K_1} d(i, 0) \right.
\]

\[
+ 4 \sum_{i=1}^{K_1-1} \sum_{j=1}^{K_2-1} \frac{K_1 - i}{K_1} \frac{K_2 - j}{K_2} d(i, j) \right). \tag{B.14}
\]

Further, if

\[
\rho(i, j) = \frac{d(i, j)}{d(0, 0)} = \text{CORR}[\hat{\beta}(\omega_1, \omega_2), \hat{\beta}(\omega_1, \omega_2)], \tag{B.15}
\]

where CORR denotes correlation, then

\[
\text{VAR}[\hat{\beta}_{xx}(\omega_1, \omega_2)] = \frac{d(0, 0)}{K_1 K_2} \left( 1 + 2 \sum_{j=1}^{K_2-1} \frac{K_2 - j}{K_2} \rho(0, j) + 2 \sum_{i=1}^{K_1-1} \frac{K_1 - i}{K_1} \rho(i, 0) \right.
\]

\[
+ 4 \sum_{i=1}^{K_1-1} \sum_{j=1}^{K_2-1} \frac{K_1 - i}{K_1} \frac{K_2 - j}{K_2} \rho(i, j) \right) \tag{B.16}
\]

At this point, let us remember that

\[
d(0, 0) = \text{VAR}[\hat{\beta}_{xx}(\omega_1, \omega_2)] \tag{B.17}
\]

independent of \( k_1, k_2 \).

By applying Eq. (10), then

\[
d(0, 0) \approx P_{xx}^2(\omega_1, \omega_2) = \text{VAR}[\hat{\beta}(\omega_1, \omega_2)] \tag{B.18}
\]
Hence,

\[
\frac{\text{VAR}[\hat{P}_{x1}(\omega_1, \omega_2)]}{\text{VAR}[\hat{P}_{x1}(\omega_1, \omega_2)]} \approx \frac{1}{K_1 K_2} \left( 1 + 2 \sum_{j=1}^{K_1-1} \frac{K_2 - j}{K_2} \rho(0, j) + 2 \sum_{i=1}^{K_1-1} \frac{K_1 - i}{K_1} \rho(i, 0) \right) + 4 \sum_{i=1}^{K_1-1} \sum_{j=1}^{K_2-1} \frac{K_1 - i}{K_1} \frac{K_2 - j}{K_2} \rho(i, j). \tag{B.19}
\]

This is the most general formula in the sense that it considers all possible amounts of overlapping among sections. If the maximum overlapping among the sections is fixed to one-half the section length, in both directions, then each section overlaps at most with its adjacent sections, and as a consequence:

\[
\rho(i, j) = d(i, j) = 0 \quad \text{for } j > 1 \text{ or } i > 1. \tag{B.20}
\]

Then, the final formula that relates the variance of the periodogram estimator and that of the periodogram averaging estimator for the case in which the maximum overlap is one-half the section length in both directions is

\[
\frac{\text{VAR}[\hat{P}_{x1}(\omega_1, \omega_2)]}{\text{VAR}[\hat{P}_{x1}(\omega_1, \omega_2)]} \approx \frac{1}{K_1 K_2} \left( 1 + 2 \frac{K_1 - 1}{K_1} \rho(0, 0) + 2 \frac{K_2 - 1}{K_2} \rho(0, 1) \right) + 4 \frac{K_1 - 1}{K_1} \frac{K_2 - 1}{K_2} \rho(1, 1). \tag{B.21}
\]

References